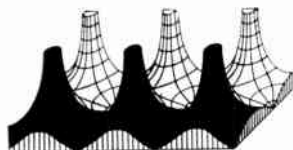


Proceedings of the IRE



Poles and Zeros



Another Special. This issue of the PROCEEDINGS is dedicated to the achievements of the various Government research laboratories of the United States and Canada. It thus becomes an opportunity to recognize the work of people which too often goes unrecognized, as well as to point out the magnitude and diversity of the contribution which governmental support makes to our field. The issue is diverse in technical content and thereby deviates from the past patterns of special issues on specific technical areas of development. Some of you may like it simply because of this diversity, while others may like it because of the applied nature of some of the papers, or because of the opportunity to hear of works too often buried in government project reports.

Dr. A. N. Goldsmith, inspired by the establishment of the National Bureau of Standards Boulder Laboratories in 1956, suggested the desirability of such an issue. Since that time the importance of governmental support of basic and applied research has continued to grow, culminating in the appointment last year of Dr. J. R. Killian as Special Assistant to the President for Science and Technology, and it is he whom we have asked to write the guest editorial which follows.

The field of work described was deliberately limited to that done by government-employed scientists and engineers in government-owned laboratories. Acknowledgment for valuable aid and suggestions should go to Captain Chris Engleman, former chairman of the Professional Group on Military Electronics, and to Dr. John T. Henderson, our Senior Past President. The 39 papers represent contributions from 21 U.S. and 3 Canadian laboratories. "Scanning the Issue" well demonstrates the diversity of their work.

One word of caution: bibliographies should not be accepted as complete, since in many cases classified papers and reports could not be listed.

Will the Supply Be There? In the fall we pointed out that contrary to publicity efforts and expectations, the engineering freshman class in our colleges and universities had fallen, and the confirmed figure is now set at 11.1 per cent below 1957, and 1957 was in fact only 1.2 per cent above 1956. Preliminary data seems further to indicate that the freshman engineering class for 1959 will be no larger and may be as much as 10 per cent smaller than 1958. In contrast to these figures are the enrollments in the physical sciences which almost universally went up last fall.

Many suspect this represents a shift of students' interest to basic science, possibly as a result of past years' publicity on our need for greater basic science knowledge, to which was not sufficiently coupled our need for engineering research to apply this basic knowledge. In any case it appears we may well be in for manpower surplus troubles in some of the

smaller physical sciences, since a field cannot readily assimilate a step function increase of 100 per cent in its manpower input, and this may happen in some of the science fields.

On top of the engineering freshman decrease there seems superposed a disturbing dropout rate. Discussion with a few engineering college administrators places this difficulty on the inability of many students to compete at the college level—the high schools seem to have failed to ready these students for the college level of output and quality. There seems to be a real lack of desire to *want* to be engineers among many of these students—are they hangers-on brought in by the tales of the engineering shortage? If they had not been so brought in to what even lower figure would our engineering enrollment have sunk?

To counteract these trends we see the need for a very well planned campaign to demonstrate the challenge to the good high school student of the unsolved engineering problems for the world ahead—those which our upcoming generations of engineers will have to solve if we are to win the race in space. Such problems include the linking of the human and automatic control systems, regenerative air, water, and food supplies, problems imposed by meteor damage, temperature control during flight and re-entry, prevention of radiation damage, astrogation, communication, and escape and rescue from a malfunctioning craft. These could be added to the extensive list of immediate needs put forward monthly by the employment advertising pages of these PROCEEDINGS.

Our modern high school products do not doubt the romance and challenge of this field when properly presented to them. Especially is this true of the field of electronic engineering, which is destined to receive a sizable portion of the total engineering enrollment if our colleges provide adequate staffs and modern facilities and curricula for this study.

What is required in addition is presentation to our secondary schools of the need for better prepared graduates, and the importance to the graduates of having been expected to present high quality work in high school. Having played in the high school band may make a boy a more interesting person, but it is not likely to help him pass calculus. Engineering education begins with ninth grade algebra—let us see that it *does* begin for the good students, and that they are taught that energy must go into engineering.

More Roses. Last month we pointed out the number of Section publications aiding in the important job of IRE communications. Now comes to hand an additional publication from a new level—a monthly circular "Electronic Production" produced by the Boston Chapter of the Professional Group on Production Techniques.

"And all the leaflets have littler leaflets and so on, ad infinitum."—J.D.R.



James R. Killian, Jr.

*Special Assistant to the
President
for Science and Technology*

James B. Killian, Jr., Special Assistant to the President for Science and Technology, was born in Blacksburg, S. C. on July 24, 1904. From 1921–1923 he attended Trinity College (Duke University), and in 1926 received the B.S. degree from Massachusetts Institute of Technology. He holds in all twenty-five honorary Doctor's degrees in law, science, and engineering from American and Canadian colleges and universities.

From 1926–1927, Dr. Killian served as Assistant Managing Editor of *The Technology Review* (M.I.T.); he was Managing Editor from 1927–1930 and then Editor from 1930–1939. Since 1939, he has successively served as Executive Assistant to the President of M.I.T., Executive Vice-President, Vice-President, and Tenth President of M.I.T. He is currently Chairman of the Corporation of M.I.T.

Since 1950 he has served on numerous Presidential and Governmental advisory committees concerning management, communications, science, and education. Since 1957 he has been Chairman of the President's

Science Advisory Committee. At the present time, he is also Director of the Research Corporation, and member of the Board of Trustees of the Nutrition Foundation, the Alfred P. Sloan Foundation, and the Board of Trustees of the Atoms for Peace Awards, Inc.

Among the awards received by Dr. Killian are the President's Certificate of Merit (1948), the Certificate of Appreciation of the Department of the Army (1953), the Decoration for Exceptional Civilian Service, Department of the Army (1957), the Public Welfare Medal of the National Academy of Sciences (1957), The Gold Medal Award of the National Institute of Social Sciences (1958), the World Brotherhood Award of the National Conference of Christians and Jews (1958), and the Award of Merit of the American Institute of Consulting Engineers (1958). In 1957 he was made an officer of the French Legion of Honor.

He is a Fellow of the American Academy of Arts and Sciences, and a member of Sigma Chi. He is an honorary member of Phi Beta Kappa and Tau Beta Pi.

*Guest Editorial**

James R. Killian, Jr.

By devoting this issue to Government research, the PROCEEDINGS spotlights an important factor in our total national research and development effort. About one-half of the Government expenditure for research, engineering and development will be spent this year in Government-owned and operated laboratories. Similarly, the increase in the number of publications originating in Government (and Government-sponsored) laboratories is one of the striking aspects of American science in the past decade. The growth of Government-financed science is one of the most remarkable accomplishments of the post-war period.

This issue has contributions from a wide array of Government laboratories. The range of the Government's interests, and the extent of its support of research and development, is indicated by the fact that some 38 different Government departments and agencies have research programs. This issue of the PROCEEDINGS is, of course, concerned principally with electronics and

advances in electrical engineering, but many other areas of Government research could be impressively represented in issues of professional journals such as this.

Even though much Government research is wide-ranging and of high quality, it must be recognized that the Government faces many problems as it supports and seeks further to encourage "in-house" research. The recently established Federal Council for Science and Technology, which resulted from a study of Government laboratories and operations conducted by a panel of the President's Science Advisory Committee, is intended to provide the Government with a better means to deal with some of these difficulties.

Let me cite some of the comments and recommendations from the report, "Strengthening American Science," which led to the establishment of the Federal Council:

There are excellent private and public laboratories working on programs for the Federal Government. The choice between these two types of operations should be based on careful judgments of where the

* Original manuscript received by the IRE, March 31, 1959.

work can best be accomplished. Both Government and industrial firms have facilities and experience that would be difficult for the other to duplicate. Sometimes a university, with its academic strength, provides the more suitable environment. Still, however, some Government researchers have tended to lose heart as the number of contract-operated laboratories has grown over the last decade. Government laboratories are vital national assets whose activities will need to keep pace with the growing magnitude of Federal research and development programs. Undue reliance on outside laboratories in placing new work of large scientific interest and challenge could greatly impair the morale and vitality of needed Government laboratories. This is an important reason why it is necessary to understand those factors in Government laboratories which can improve morale and add significantly to their capabilities.

Inadequate personnel policies and practices can inhibit the healthy growth and development of the Government's technical activities. Morale will suffer if Government salaries cannot be made more competitive with industrial research salaries. Congress and the Civil Service Commission have recently made some welcome salary improvements, but more needs to be done. Attempts should be continued, for example, to make Government salaries more flexible at the beginning levels. Recognizing, however, that it may never be possible, or desirable, to equalize Government and industrial salaries completely, it is imperative that Government departments pay closer attention to cultivating significant research programs and formulating attractive policies that will appeal to imaginative and creative people. For example, scientists and engineers should be given frequent opportunities by their agencies to attend, expenses paid, national and international meetings in their field. It has been demonstrated many times—notably in universities—that a unique challenge and a proper atmosphere can make up for some of the monetary inequalities.

Another basic requirement for making Government research attractive is to provide openings for

scientists and engineers at the highest rungs of the administrative ladder. There are still departments where it is not possible for a technically trained civilian to assume full responsibility for the management of a laboratory. If an individual is denied the opportunity to assume major responsibilities or continue his professional growth, there will be a further reduction in the attractiveness of technical careers for civil service personnel.

Government laboratory operations sometimes suffer from improper financing. A pattern has been established in which a few laboratories receive only a part of their operating budget from the particular department, bureau, or agency to which each reports. The balance is obtained through transfers. This practice not only hinders the director of the laboratory in planning a rational program but also introduces instability among programs and personnel. If fiscal arrangements of this kind were corrected, these laboratories could serve the Nation more efficiently. Departments could benefit in this regard by further emulating the general trend in industry which is moving away from piecemeal financing as it strives for more efficient utilization of technical manpower.

Finally, Government laboratories must be given the freedom by their departments to do the basic research necessary to support their clearly defined missions. They also need authority to negotiate contracts and make research grants necessary for a mutually profitable relationship with private scientific communities. Congress has recently extended the authority of Government agencies with operations in the research field to do both, and this action is most helpful.

The Federal Council for Science and Technology is now in operation. With its membership made up of policy-making officials from those departments with major research programs, and with the President's Science Advisory Committee made up of able representatives of the non-Government scientific community, the two groups working together have an exceptional opportunity to be vigorous protagonists of strong Government science programs and institutions.

Scanning the Issue

The 39 papers in this special issue provide a comprehensive and representative picture of the many types of communications and electronics work being carried on by government laboratories in the United States and Canada. The issue comprises both a report of new results and a review of recent developments of interest and importance to the entire electronics community.

Although the majority of the laboratories that contributed to the issue are connected with military establishments, the scope and significance of their work, as this issue demonstrates, extend far beyond areas that are of purely military concern. The work reported here includes investigations that are highly fundamental in nature (standards of measurement, wave propagation studies, solid-state phenomena), developments that are of major importance to industry (microminiaturization techniques, reliability studies, communication engineering methods), and applications that are of widespread utility (medical electronics).

The importance of the government laboratory as a contributor to the technical progress of our profession, portrayed in piecemeal fashion by the 39 papers in this issue, is brought into sharp focus by the guest editorial on the preceding page. We are privileged to present the words of a man uniquely qualified to speak on behalf of all United States Government scientific activity, Dr. James R. Killian, Jr., Special Assistant to the President for Science and Technology.

The papers that follow, since they cover a wide variety of topics, have been grouped into general areas of common interest. Broadly speaking, the first part of the issue deals with communications and the second part with electronics. The subject matter is further delineated by the section headings below.

MEASUREMENT STANDARDS

Underlying all branches of the communications and electronics field, and indeed all fields of physical science, is the necessity for accurately measuring many physical phenomena. The spectacular scientific advances of the past few decades have in fact resulted largely from improvements in the accuracy with which these measurements can be made. It is appropriate, therefore, to open this issue with a subject that is of such fundamental importance to all the papers in this issue.

Two outstanding papers are presented from the National Bureau of Standards which describe our system of measurements in a manner that will be meaningful and interesting to every reader. The first (McNish, p. 636) gives a clear insight into the basis of our system of units and the standards that have been established to fix and preserve the sizes of the units. A companion paper (Silsbee, p. 643) describes in further detail the interesting measuring procedures and supporting research which must be continually carried out for just one of the many existent units, the ampere.

COMMUNICATION SYSTEMS

The interest of the military services in radio communication is as old as radio itself. It dates back at least to 1899, when the U. S. Army Signal Corps established radio contact over the 12 miles separating the Fire Island Lightship from Fire Island. Army communication requirements have since expanded to the astonishing point where today a so-called Type Field Army needs approximately 1,000,000 miles of

communication channels. It is thus that the Signal Corps has become extensively involved in a large number of related technical areas, including radio equipment and transmission, multiplexing, switching, data processing, systems engineering, radio interference—and now space communication. An excellent review is given of present work in these important areas, setting forth a number of challenging and advanced ideas in the process (Lacey, p. 650).

The very-low frequencies received practically no attention until very recently, but are now taking on considerable importance. The Defence Research Telecommunications Establishment in Canada has prepared an excellent engineering study of VLF long-range communication systems, the likes of which has not appeared in the literature for at least 30 years (Belrose, *et al.*, p. 661). In the HF band (3 to 30 mc), which is more commonly used for long-range communication, it becomes necessary to try to take into account the temporal vagaries of the ionosphere. The DRTE discusses in another paper the means of predicting what frequencies to use to provide radio service between any two points at any time of day, season, and year (Sandoz, *et al.*, p. 681). A recent and important addition to communications technology was the UHF tropospheric scatter communications system. As an extension of the angle diversity technique, a novel multibeam system is proposed by the Rome Air Development Center as a step toward utilizing the less crowded microwave frequencies for scatter communication (Vogelman, *et al.*, p. 688). A discussion of currently interesting communication systems would not be complete without a review of the AM-SSB-DSB-CSSB struggle now in progress, presented here from the impartial viewpoint of the Naval Air Development Center (Nupp, p. 697).

ANTENNAS AND PROPAGATION

The evolution of present-day radio technology has depended to a large degree on a continuing study of both radiated and guided electromagnetic fields. The actual detection and measurement of radiated fields, in turn, play an important part in the understanding and application of electromagnetic waves. From the National Research Council of Canada comes a comprehensive tutorial survey of modern radiation measurement techniques as they apply to electromagnetic wave studies, antenna design, and antenna performance (Cumming, p. 705). From broad fundamentals we turn to a descriptive paper on antennas of a special type that has long been neglected in the literature. The Navy Underwater Sound Laboratory has had principal responsibility for the development of submarine communication antenna systems since 1945, and the fruits of their work in the VLF, HF, and UHF bands are now described in a brief, well-illustrated manner (Turner, p. 735).

In evaluating the performance of a VHF or UHF communication system it has long been customary to account for the important effect of atmospheric refraction by means of the well-known "four-thirds earth" concept, which greatly simplifies the analysis and usually provides a reasonably accurate medium-range model of the refractive index structure of the atmosphere. The National Bureau of Standards has now devised improved models that are more accurate for the increasingly important long-range case, such as is encountered in the design of new long-range radar systems (Bean and Thayer, p. 740). Much useful experimental knowledge concerning radar and propagation conditions is being provided by the Naval Research Laboratory's novel "flying laboratory"

(Ament, p. 756). Ground stations, too, provide important data. They have shown that one of the most distinctive features of VHF radio transmission is the occasional occurrence of sporadic-E propagation, resulting in greatly increased signal power at the receiver. The National Bureau of Standards reports the results of an important five-year study of this phenomenon over an 800-mile path between Iowa and Virginia (Davis, *et al.*, p. 762). The discussions of VHF and UHF propagation end with an account of the noteworthy work that has been done at the Air Force Cambridge Research Center in ionospheric and tropospheric propagation at extreme distances and heights (Ames, *et al.*, p. 769).

NAVIGATION AND DETECTION SYSTEMS

Navigation Aids

This section opens with another outstanding contribution from Canada on a subject of major current interest—a self-contained airborne Doppler radar navigation system developed by the Canadian Defence Research Board Electronics Laboratory. The six-part paper includes an excellent introduction to aircraft navigation and a discussion of the principles of Doppler radar, as well as a description of the circuits, hardware and systems that go to make up the unit. The discussions of microwave antennas, modulation and propagation theory, analog vs digital computers, and data vs mechanical stabilization are a few of the reasons why the paper will enjoy a wide audience (Thompson, *et al.*, p. 778).

The Doppler effect has also been put to interesting use by the Civil Aeronautics Administration (now called the Federal Aviation Agency) to produce an improved VHF omnirange system which promises to be of importance in meeting the increasingly stringent demands of air traffic control operations (Anderson and Flint, p. 808). The Diamond Ordnance Fuze Laboratories proposes a novel distance measuring system that uses random noise to modulate the transmission. Because the nonrepetitive nature of noise avoids a source of error inherent in other systems, permitting distance measurements down to a few feet, this system appears to be very well suited to radio altimeters, especially in connection with blind landing systems (Horton, p. 821).

As noted above, the VLF band is attracting a growing amount of attention. This interest is due not only to the long communication range it affords, but also to the recent realization that much greater phase stabilities exist at these low frequencies. This realization has opened the door to the possibility of building a system that would provide navigation for ships and planes, day or night, under all weather conditions, to a range of about 5000 miles with an accuracy of 1 mile or better—a system believed to be well beyond the state of the art only a short time ago. As a result, the Navy Electronics Laboratory is now working on an important new system, called Radux-Omega, which is intended to operate below 20 kc and to provide world-wide navigational coverage with as few as eight stations. The highly important experimental data included here regarding the phase stability of VLF waves will substantially improve our understanding of VLF propagation (Casselman, *et al.*, p. 829).

Simulation

An important and especially interesting adjunct to air navigation work has been the recent development of electronic techniques for studying and handling air traffic control problems and for training personnel. The Federal Aviation Agency (formerly the CAA) has done a particularly outstanding, unpublicized job of developing and introducing simulation techniques, improved radar displays, and automatic data processing systems to keep pace with the growing air traffic problem (Astholz, p. 840). Novel simulation equipment has been de-

veloped also by the Naval Training Device Center for training pilots, using a pulsed analog computer (Herzog, p. 847).

Tracking

The space age has added new problems to the important business of tracking and determining the paths of objects, and again the computer is playing an important role. Members of the National Aeronautics and Space Administration, Naval Research Laboratory, and National Bureau of Standards have developed a computer program which permits a rapid determination of a satellite orbit from only a small number of radio tracking observations (Harris, *et al.*, p. 000). When tracking is done by radar, the size and complex shape of a target introduce wave-interference phenomena, resulting in scintillation noise which affects the accuracy of the results. The Naval Research Laboratory is shedding a great deal of light on this source of tracking error, which is becoming increasingly important as the speed and distance of targets increase (Dunn, *et al.*, p. 855).

Underwater Sound

A ten-year program of underwater acoustic research by the Navy Underwater Sound Laboratory has led to a dramatic ten-fold improvement in the capability of submarines to detect and track target vessels by means of passive sound-detection systems (Ide, p. 864). Underwater sound transmitters, too, have undergone extensive study. The Navy Underwater Sound Reference Laboratory has developed an audio-range sound transducer that represents a substantial advance in the art (Sims, p. 866). An interesting parallel between sonar and radar has been explored by the Navy Electronics Laboratory, in which the results of recent investigations into echo detection theory for radar are shown to have application to sonar as well (Stewart and Westerfield, p. 872).

ELECTRONIC DEVICES AND EQUIPMENT

Readers are forewarned that this section, while dominated by developments in the solid-state art, also delves into a number of other areas, such as ultrasonics, information theory, and reliability.

One of the most exciting new horizons for solid-state technology is in the microminiaturization of electronic equipment. This means a good deal more than just making existing components and assemblies smaller. It involves the creation of totally new circuit elements, each of which may combine in a single solid the functions of many standard components and circuits, and perhaps some new functions, as well.

The electronics industry has been looking forward to a report of the important work being done in this area by certain government laboratories. Their wishes are now fulfilled. The Diamond Ordnance Fuze Laboratory has prepared a major contribution which describes the fabrication of a 14-part transistorized binary counter on a $\frac{1}{2}$ -inch square of steatite-ceramic, only $\frac{1}{50}$ inch in thickness. The unusual techniques employed involve photolithography, thin film deposition, conductive adhesives, and ultrasonic drilling (Prugh, *et al.*, p. 882). Highly important work is also being done at the Army Signal Research and Development Laboratory. Several microminiature circuits are discussed in a second important paper, revealing the latest industry-wide thinking on the subject (Danko, *et al.*, p. 894).

Photoemissive materials can be coupled with fluorescent materials to form image converting and intensifying devices for making infrared and other invisible images visible to the eye. Recent developments, reported here by the Army Engineer Research and Development Laboratories, have led to devices which when perfected will permit night observation under starlight illumination with no projected radiation

whatsoever. These devices will be of substantial interest to astronomers, radiologists and nuclear physicists, as well as the military (Klein, p. 000). Photoconductive, as well as photo-emissive materials are important to the infrared detection problem. An excellent survey of work being done on this subject, and also on the applications of correlation methods to a variety of signal processing problems, is presented by the Naval Ordnance Laboratory (Scanlon and Lieberman, p. 910). The investigations at Picatinny Arsenal of the amount of charge released by ceramic ferroelectric materials under various conditions of temperature and pressure provide information of special interest in the development of new sources of power (Doremus, p. 921). Other interesting uses of ferroelectric materials, summarized here by the Naval Air Development Center, include sonar transducers, piezoelectric hydraulic pumps, and intracardiac phonocatheters for recording heart sounds from within the heart (Wallace and Brown, p. 925).

The Naval Ordnance Laboratory is engaged in a variety of interesting electronic subsystem development problems, including underwater influence mechanisms using the recently developed Solion, rocket-sonde meteorological equipment, missile-target miss measurements, and field test sets (Beach, *et al.*, p. 929). An overriding problem faced by all equipment designers, and especially those dealing with airborne or missile equipment, is that of reliability. This excellent description of the approach taken by the Rome Air Development Center will be of wide interest to readers (Naresky, p. 946).

INSTRUMENTATION

The advent of the jet airplane has intensified the study of the origin of aircraft noise and its elimination or reduction. In this connection, the National Aeronautics and Space Administration has developed several instruments designed to process noise measurement data with greater reliability and speed (Carlson, *et al.*, p. 956). Aircraft engineers, together

with missile engineers, are interested also in testing components at supersonic speeds and large accelerations. The Naval Ordnance Test Station built a research track for this purpose, on which a sled which carries the components is propelled at high velocities, necessitating the development of a special telemetry system for obtaining precise velocity data (Stirton and Glatt, p. 963). Turning from the air to the sea, we find that the Naval Underwater Ordnance Station has explored a wide variety of novel instrumentation methods for tracking torpedoes (Barry and Formwalt, p. 969).

It is especially fitting that an issue devoted to the work of the government should end by showing how this work is being directly applied to the general health and welfare of its citizens. For example, one of the most costly catastrophes that annually befalls our nation is floods, and an important weapon for controlling them is the flood-control reservoir. The effectiveness with which this weapon can be used has been greatly increased by the development of telemetering networks by the Army Engineer Waterways Experiment Station for quickly gathering vital data on the flood stages of main and tributary streams above and below the flood-control dam (Hanes, p. 977). In the field of medicine, the Naval Medical Research Institute has developed a novel control system for injecting a radio-opaque dye into the arteries at the proper instant so that an X-ray picture may be taken of heart action at the desired moment during the coronary cycle (Eicher and Urschel, p. 984).

Sanitary engineering is a field most readers will not associate with electronics. Yet within the past decade electronics has been an increasing influence both in the research and applied fields at the Robert A. Taft Sanitary Engineering Center. The extent of this influence in air pollution, radiological health, milk and food sanitation, and water supply and pollution studies (Geilker and Kramer, p. 986) and the electronic instrumentation developed for studying air pollution (Nader, p. 988) is briefly outlined in two companion papers that close the issue.

The Basis of Our Measuring System*

A. G. McNISH†

Summary—The measuring system used for scientific work affords a means of making physical measurements with great precision and accuracy. The best measurements can be made of the quantities taken for the basis of the system. A decrease in both accuracy and precision arises in measuring quantities which are related to them in a complicated way. The standards which fix the magnitudes of the units on which the system is based appear to be very constant. Some improvement in the system may be obtained by substituting physical constants for these standards. This has already been done for the standard of temperature, and it can be done advantageously for the standards of length and time; but there seems to be no way to replace advantageously the standard for mass.

INTRODUCTION

THE most satisfactory kind of information that we can secure about phenomena is developed through measurement. If the phenomena are not susceptible to measurement, the information we can obtain of them is rather unsatisfactory and incomplete. It is the high measurability of the phenomena involved which sets the physical sciences apart from the others and causes them to be called the exact sciences.

Corollary to this, the more accurate the measurements, the better is the information which we obtain. To a considerable extent, the spectacular advances we have achieved in the physical sciences during the past few decades are the results of more accurate measurements. Rough descriptions of phenomena are no longer satisfactory. A few centuries ago, it was adequate to know that the rate of free fall for a body is independent of its mass. Later, the knowledge that the path of a freely falling body with some initial motion describes an approximate parabola sufficed. Today we recognize that the ideal path is actually an ellipse; and, if the gravitational field is not completely specified by the inverse square law, perturbations of the elliptical path must be allowed for, as in the case of an artificial satellite.

In order to carry out accurate measurements, we have established a system of units and a system of standards to fix and preserve the sizes of the units. What are the requirements of such a system? How many units and how many standards do we need? And how many of them are what we call "basic"?

Let us consider several possible systems. For example, we might take for the unit of length, the length of some arbitrarily selected bar; for the unit of mass, the mass of some specified object; the unit of electromotive force, the open circuit EMF of a particular Clark cell; the unit of current, the short-circuit current of a particular grav-

ity cell; the unit of resistance, the resistance of a particular piece of wire, etc. For each quantity in physics, we would have an independent unit for measuring that quantity and the magnitude of each unit would be fixed by the standard embodying it.

What are the faults of such a system? One of its disagreeable aspects is that, with units so chosen, it would be a rare case of good fortune if we found that the equation $E = IR$ holds. We would find, of course, $E \propto IR$. Each physical law expressed in the form of an equation would require a different factor of proportionality to relate the sizes of the units involved. However, if we have selected our standards for such a system with good judgment, we would have an excellent set of standards and we could measure physical quantities with great accuracy in terms of the unit embodied by the standards. Also, we could select the standards so that all units would be of convenient size for practical use.

We can exercise a little parsimony in our choice of independent standards by selecting only a few to fix the units of several quantities and have the units for other quantities fixed by physical equations. Thus, for example, having independent units for electromotive force and resistance, we can write $I = kE/R$ to define the unit of current. We can simplify things by setting k equal to unity. How far can we go in our parsimony? What is the smallest number of arbitrary, independent units and standards which are required for a measuring system?

The founders of the metric system sought to reduce the number of independent standards by defining the gram as the mass of one cubic centimeter of water, leaving the meter as the one independent unit and adopting a constant of nature, the density of water, to fix the unit of mass. For reasons to be discussed later, they found this unsatisfactory. Had it worked we would have found it convenient for it would be necessary for measuring laboratories to compare only meter bars. Each could then set up consistent units of area, volume, and mass, depending only on the meter bar.

We can see that with a system such as this an entire system of measuring units can be built up based upon a single arbitrary, independent standard. The unit of time could be the time of swing for a one-meter pendulum, or the time light takes to travel one meter. Other units could be defined by various equations as we did before for the unit of electric current.

If we set all of the constants in our defining equations equal to unity, we will find this system very convenient for theoretical work, for most equations are in their simplest form. But we will not find the system good for experimental work because many of the quantities most frequently measured in physics would not be measurable

* Original manuscript received by the IRE, December 9, 1958; revised manuscript received, February 19, 1959.

† National Bureau of Standards, Washington, D. C.

with sufficient accuracy in terms of the units prescribed for them.

Thus we see that we can have a multi-unit multi-standard system in which all units and standards must be regarded as "basic." Also, we have seen that we can have a single-standard system in which only one quantity is taken as basic. Is this the limit to which our parsimony can go? Can we have a consistent system of measurement with *no standard at all*? The answer to this is, yes, and a moment's reflection will show that it can be done in a number of ways.

To illustrate one of the many ways this can be done, suppose, we rewrite the electromagnetic wave equation $\partial^2\phi/\partial t^2 = \partial^2\phi/\partial x^2$ instead of $\partial^2\phi/\partial t^2 = c^2\partial^2\phi/\partial x^2$, the quantum equation $E = \nu$ instead of $E = h\nu$, the gravity equation $F = m_1m_2/r^2$ instead of $F = Gm_1m_2/r^2$, and the molecular energy equation $E = (3/2)T$ instead of $E = (3/2)kT$. Now with provisional measuring units of arbitrary size we can perform various experiments involving these equations, assuming that other equations of physics are written in their conventional forms. We then solve the results obtained and find what values we must assign to our provisional units. If these sets of experiments are performed by different people, they will be in agreement on the sizes of their units.

The necessary experiments already have been performed, using our conventional units of measurement, so that we can write down the equivalence of these in our new units as shown in Table I. To put it another

TABLE I
APPROXIMATE VALUES OF CONVENTIONAL UNITS IN TERMS
OF UNITS DERIVED BY SETTING c , G , h , k AND
 ϵ_0 EQUAL TO UNITY

One meter	2×10^{34} length units
One second	8×10^{42} time units
One kilogram	2×10^7 mass units
One degree Kelvin	2×10^{-33} temperature unit
One watt	2×10^{-53} power unit
One coulomb	2×10^{17} charge units

way, what we have done is to assign the value unity to the speed of light c , Planck's constant h , the universal gravity constant, G , and Boltzmann's constant k . If we also set μ_0 and ϵ_0 , which we call the permeability and permittivity of space equal to unity, as we may do consistently since c is unity, further interesting results are obtained. We now find that the units of the electrostatic system and the electromagnetic system as defined by Coulomb's equations are of identical magnitude; the unit of electric charge is about 4.8×10^{-18} coulombs, which is approximately equal to the nuclear charge for the elements where the percentage mass defect is a maximum; the electronic charge is given exactly by $e = \sqrt{\alpha}/2\pi$, where α is the well-known Sommerfield fine-structure constant, approximately equal to $1/137$, Eddington's magic number! Also the unit of energy is exactly the energy released when one unit of mass is annihilated in accordance with Einstein's equation $E = mc^2$.

In spite of a number of appealing features of this system we cannot recommend it for practical use. Because of the uncertainty involved at the present time in relating experimentally the gravitational constant to other quantities of physics, the magnitudes of the units so defined would have an uncertainty in their practical realization of about 1 part in 1000. We can, of course, use some other constant instead of G and thus reduce the uncertainties, but they would still be much greater than we can afford to have them.

We now ask ourselves if we have really eliminated standards from our system by this procedure. Have we not made c , h , G , k , and ϵ_0 our new standards? Are these not our new "basic" quantities of physics? This is all a matter of point of view.

Some of these concepts may seem heterodox to those who have been indoctrinated in the trinity of mass, length, and time as fundamental units. They have been touched on here because some smattering of them is necessary to understand the system of units and standards currently used by physical scientists of all nations of the world. They serve to illustrate the requirements of a system of units and standards for an adequate measuring system, requirements which are very well met by our present system.

OUR PRESENT UNITS AND STANDARDS

The authors of the metric system intended that the meter should be one ten-millionth of the length of the north polar quadrant of the Paris meridian, demonstrating a desire that a measuring system be based on some natural magnitude which would remain constant, a desire which still exists with metrologists. But they soon found that they could compare two meter bars with each other with greater precision than they could relate them to the earth's quadrant.

Similarly, they found that the masses of two metal cylinders weighing about one kilogram each could be compared with each other more precisely than either could be related to the mass of 1000 cubic centimeters of water which was supposed to define them. Thus, the pioneers in precise metrology found that greater precision in measurement could be achieved if they adopted some readily measurable artificial things for standards than if they adopted less readily measurable things of nature.

Accordingly, the International Commission of the Meter, meeting in 1872, resolved to take the meter in the Archives of Paris "as is" (*dans l'état où il se trouve*) for the standard of length. Similarly, they adopted the platinum-iridium kilogram of the Archives as the standard of mass, "considering that the simple relationship, established by the authors of the metric system between the unit of weight and unit of volume, is represented by the actual kilogram in a sufficiently exact manner for the ordinary uses of industry and science . . . and that the exact sciences do not have the self-same need of a simple numerical relationship, but only

of a determination as perfect as possible of this relationship." Three years later many of the leading nations of the world signed the treaty of the meter which created a procedure for coordinating the standards of measurement for the scientific world through an International Bureau of Weights and Measures and a General Conference of Weights and Measures.

For many years no new standard for time was adopted, the ancient definition of the second as 1/86,400 part of a mean solar day being retained. A separate and independent standard for temperature measurements was adopted and changed a number of times until, in 1954, the thermodynamic temperature scale was based on the triple-point temperature of water as 273.16°K. Other units were defined in terms of the units embodied by these standards by agreed-on equations of physics.

Thus, our system of units and standards does not follow either of the extreme types of systems described in our introduction. Four, and four only, independent standards have been adopted to which we attach the name prototypes. The reasons for this are clear. They are standards for quantities which can be measured very accurately and they are standards which we believe can be preserved or *re*-produced, as in the case for the standard of temperature, very accurately. Furthermore, standards for other quantities can be constructed from them with adequate accuracy in accordance with the defining physical equations.

FROM PROTOTYPES TO ELECTRICITY STANDARDS

As pointed out before, we could establish our standards of electricity and magnetism with the same degree of arbitrariness that we used for the prototypes. However, it is not desirable that we do this, for then the units of electrical and mechanical power would not be the same unless we cluttered up the equations of electricity and magnetism with an unnecessary number of numerical constants. Furthermore, standards for the more useful electric quantities may be established in terms of the prototype standards with great accuracy, as shown by Silsbee.¹

In the early days of research in electricity and magnetism, there was no way to measure these phenomena accurately. The strength of a current was measured by the deflection of a compass needle which the current produced, but, since the deflection depended on the strength of the earth's magnetic field, two observers in different places could not compare their results. Variations in the strength of the geomagnetic field from place were measured by timing the oscillations of a "permanent" magnet which, even in those days, was known to be far from permanent.

C. F. Gauss first showed how magnetic and electric quantities could be measured accurately in terms of the units used for mechanical quantities which are embodied in relatively invariant standards. This is the well-known magnetometer experiment which every student

of physics and electrical engineering must have performed. That Gauss thoroughly understood what he was doing we must presume. That he anticipated the reverence and mysticism with which future generations of scientists would regard the units of mass, length, and time is not likely. It is clear from the writings of those who followed him that they recognized a duplicity in electric units (esu) and magnetic units (emu) and that they could not simultaneously set ϵ_0 and μ_0 in these systems equal to a dimensionless unity and still retain the generally agreed-on unit of time as well as the metric units of length and of mass.

Thus, in the early days of electromagnetic science, the systems of units were not on a satisfactory basis since electric and magnetic units were conflictingly defined. Nor was the situation remedied by combining the esu and emu systems into that bifurcated system which falsely is called Gaussian. A further disagreeable feature is that some of the units in both systems are of inconvenient size, and are not even approximately equal to the units for voltage, current, and resistance which had been adopted by the communications engineers of that day.

To resolve these difficulties, Giorgi proposed that one more arbitrary unit and a standard for it, this time for an electric quantity, be added to the system. By suitably choosing this new unit we can make the MKS units and the practical electrical units form a consistent system. What we mean by this is that equations of the type $W = I^2 R$ would be valid if I and R were expressed in the electrical units and W is expressed in MKS units. Of course, Coulomb's equations for the force between unit electric charges and between "unit magnetic poles" would contain constants of proportionality other than unity just as does the equation for the gravitational attraction between two-unit masses, as it is ordinarily written.

Giorgi's proposal found such increasing favor among scientists that in 1946 the International Commission on Weights and Measures adopted it in principle. But instead of adopting a new independent unit and a standard for it, they redefined the ampere¹ in accordance with another suggestion of Giorgi in such a way as was equivalent to rewriting the equation for the force between two infinitely long current-carrying conductors. The old emu equation for force per unit length for unit separation of the conductors was $F = 2I^2$, where F is in dynes and I in abamperes. The new form of the equation became $F = 2\mu_0 I^2 / 4\pi$, where F is in newtons, I in amperes, and μ_0 has the value $4\pi \times 10^{-7}$, if rationalized equations are used. It follows from this that the equation for the force between two charges is $F = Q_1 Q_2 / 4\pi \epsilon_0 R^2$ and, if F is in newtons, Q_1 and Q_2 in coulombs, and R in meters, the equation is correct if ϵ_0 is given by $1/\mu_0 c^2$, where c is the speed of light in meters per second.

The present-day system of electric units is known as the MKSA system. It is important to recognize that the unit for electric current, the ampere, does not oc-

¹ F. B. Silsbee, "The ampere," this issue.

copy quite the same position in it as the meter, kilogram, and second, for they are prototype units and their values are fixed by their independent, proper standards. On the other hand, the value of the ampere is fixed in terms of them as given by an equation involving an arbitrarily adopted constant of proportionality. Thus the position of the ampere today is somewhat like that of the kilogram in the scheme proposed by the founders of the metric system.

We can measure all physical quantities directly in terms of the prototype standards, but following this procedure in all measurements is inconvenient, inaccurate, and imprecise. For this reason we construct standards for various derived physical quantities and assign values to them in terms of the corresponding unit as derived by experiment from defining equations involved and the units fixed by the prototype standards. It is clear that, however carefully we perform the experiments in deriving these standards, they can never have the exactness inherent in the prototype standards. How these errors enter our derived standards is discussed in detail by Silsbee¹ and Engen.²

Among the most accurate of the derived standards are those for some of the electric quantities. Because of this high accuracy and the convenience of making electric measurements, these standards are often used for measuring other quantities. We measure the heat of combustion of fuels, for example, by comparing the heat evolved when they are burned in a calorimeter with the known heat evolved by the passage of an electric current through a resistor. The results are given directly in joules, the internationally agreed-on unit for heat, but many chemists foolishly divide the results by 4.1840 to convert them to calories.

We can see how well our measuring system works in practice by examining how well various physical quantities can be measured in terms of the prototype standards. Several typical quantities are listed in Table II with estimates of the uncertainties involved in the best measurements of standards for these quantities at the National Bureau of Standards. The uncertainties are expressed in two ways, those of accuracy and those of precision. Under precision are placed the uncertainties in comparing two nominally identical standards for the quantity involved. Under accuracy are placed estimates of the uncertainties in relating the derived standards to the prototype standards. The uncertainties correspond approximately to "probable errors." No entries under accuracy are given for the meter, the kilogram, or the triple-point temperature of water since they are "accurate" by definition. The quantity, time, is not included in the table because it involves some considerations which will be treated at length later.

The table illustrates a number of interesting characteristics of our measuring system. The greatest precision is attained in measuring the quantities which

have been selected as the basis of our system. In all cases illustrated except one, maximum precision is achieved when the magnitude of the quantity is the unit for the quantity. In all cases there is a decrease in accuracy for standards which embody a multiple or submultiple of the selected unit. Derived standards are subject to considerable inaccuracy, and this inaccuracy increases with the experimental complexity involved in relating them to the prototype standards. However, derived standards may be compared with similar ones with a precision far greater than the accuracy of the particular standard involved.

TABLE II
ESTIMATES OF ACCURACY AND PRECISION IN MEASURING PHYSICAL QUANTITIES

Physical quantity	Device	Magnitude	Uncertainty in parts per Million	
			Accuracy	Precision
Length	Meter bar	1 meter	—	0.03
	Gage block	0.1 meter	0.1	0.01
	Geodetic tape	50 meters	0.3	0.10
Mass	Cylinder	1 kilogram	—	0.005
	Cylinder	1 gram	1	.03
	Cylinder	20 kilogram	0.5	0.1
Temperature	Triple-point cell	273.16°K	—	0.3
	Gas thermometer	90.18°K	100	20
	Optical pyrometer	3000°K	1300	300
Resistance	Resistor	1 ohm	5	0.1
	Resistor	1000 ohms	7	1
	Resistor	0.001 ohm	7	1
Voltage	Standard cell	1 volt	7	0.1
	Volt box-standard cell	1000 volts	25	10
Power DC	Standard cell-resistor	1 watt	11	1.5
	Wattmeter	10-1000 watts	100	50
	Microcalorimeter	0.01 watt	1000	100

The maintenance and establishment of standards for all kinds of physical measurements is the basic responsibility of National Bureau of Standards. Fulfillment of this task requires development of precise measurement techniques and prosecution of basic research in most fields of the physical sciences. For a standards program is not a static program. Growing technology requires more and more standards, extension of the range of existing standards, and improvement in accuracy. For example, the rather poor accuracy associated with measurement of microwave power is not due to difficulties in the experiment alone, but partly to the fact that accurate measurement of microwave power is a recent requirement of our technology. It is interesting to note that the old "International" Ohm as embodied in the standard specified for it by the International Conference on Electrical Units and Standards in 1908, and used until 1948, differed from its theoretical value by about 500 parts in 1,000,000.

² G. F. Engen, "A refined X-band microwave microcalorimeter," to be published.

PHYSICAL CONSTANTS AS STANDARDS

The desire of metrologists to have their units of measurement embodied in indestructible, immutable standards was evidenced by the founders of the metric system. There are good practical reasons for this desire apart from its aesthetic appeal. If realized, we would be assured that measurements made at one time would be strictly comparable with those made at another, unless the substances and the laws of physics are themselves changing with time. But most important is the circumstance that each adequately equipped laboratory could have its own set of standards known to be identical with those of a similar laboratory without the need for regular intercomparisons.

We have seen how, in the early days of the metric system, it was necessary to abandon this desire in order to achieve the greatest precision in measurements. To what extent are we now able to substitute physical constants for the artificial standards of our measuring system without impairing accuracy? We should like, of course, to adopt physical constants of the most basic nature to set the scale of our measuring system, such as G , h , c , and k , but, as we saw, precise measurements cannot be made in terms of them. Improvement can be obtained by abandoning G and using another constant instead of it, but the improvement is not enough.

We can gain considerable improvement in precision, though with a sacrifice of elegance, by selecting less general physical constants as the standards for our measuring system. In fact, this already has been done for one physical quantity, temperature, since the number 273.16 is assigned to the triple-point temperature of water on the thermodynamic Kelvin scale. We can see what an advantage this is, because anyone versed in the art of temperature measurements can construct his own triple-point cell and establish the standard of temperature which can be realized with a precision of two or three ten-thousandths of a degree. Furthermore, this standard can be re-produced at any time in the future with complete confidence that the standard will be the same. (The definition is still deficient in that the isotopic composition of the water is not specified, but this fault will probably be corrected in the near future.) Pleasing to think about, although of no immediate advantage, is that a man on Mars, or even in some other solar system or galaxy, could establish the same standard for temperature measurement as we use on the earth.

If establishment of a standard for the temperature scale were all there were to temperature measurements, our task would be simple and dull. We must make temperature measurements over a wide range, and we must be able to express our measurements over this range with as great an accuracy as possible in terms of our standard. This led to extensive research in the Bureau on phenomena involving temperature and the atomic constants which are needed for measurement of temperatures by Planck's radiation law. Outgrowth of this work involved some of the pioneer precise measurements of the

energy levels of excited atoms in the days when Bohr's theory was still an hypothesis.³

By taking a physical constant to define our temperature scale we have gained a great deal and lost nothing in precision of temperature measurements. There does not appear to be any better way to establish a standard for temperature. This is not true for all the other prototype standards.

Can we replace the standard for length by a natural standard, say, the wavelength of some chosen spectral line? Babinet proposed this in 1827, but 65 years passed before the first measurements were made to reduce this to practice by Michelson and Benoit in 1892-93. However, the use of a wavelength for a standard of length instead of the meter bar is not without fault.

By precise spectroscopy we can compare wavelengths of two highly monochromatic spectral lines with a precision of a few parts in 10^9 . Thus, if a suitable spectral line is accepted as a standard of length, this would also be the approximate accuracy with which good spectral lines can be measured. But the task of measuring a material standard in terms of a wavelength standard is more difficult. Since the pioneer measurements of Michelson and Benoit, eight determinations of the relationship between the wavelength of the red line of cadmium and the length of the meter have been performed. From concurrence of the results, the probable error of the precision of a single determination is calculated to be 1 part in 10^7 . This is somewhat inferior to the precision with which two meter bars can be compared. Also the wavelength measurements may be subject to systematic errors affecting all determinations alike. The Bureau has graduated a meter bar directly using the accepted value for the wavelength of the cadmium line as a standard. Subsequent comparison of this meter bar with others indicates a discrepancy in the graduation consistent with the above results.

The cadmium red line originally used by Michelson is not the best line for precision spectroscopy. To obtain a better line, Meggers, Chief of the Bureau's Spectrographic Section, developed a new lamp.⁴ Earlier, Michelson had suggested the use of the mercury green line as a wavelength standard, but this was found to have too much fine structure because mercury in nature consists of seven isotopes, two of which have nonzero spin. To cure this fault Meggers used mercury 198, which has zero spin. This was produced in pure form by bombarding gold, which has only one stable isotope, 197, with neutrons. This lamp gives much clearer interference patterns, but, to excite the radiation, it is necessary to have some argon in the lamp. The wavelength of the mercury radiation depends on the argon pressure, so that leakage of argon will affect the standard.

³ P. D. Foote and F. R. Mohler, "Determination of Planck's constant h by electronic atomic impact in metallic vapors," *J. Opt. Soc. Am.*, vol. 2-3, pp. 96-99; 1919.

⁴ W. F. Meggers, "A light wave of artificial mercury as the ultimate standard of length," *J. Opt. Soc. Am.*, vol. 38, pp. 7-14; 1948.

Since wavelengths emitted depend on conditions of excitation, experiments have been conducted with lamps containing krypton 86 held at constant pressure by immersion in liquid nitrogen at its triple point. Various national laboratories have found this wavelength standard so satisfactory that the International Committee on Weights and Measures in October last year recommended that the General Conference, which meets in 1960, redefine the meter in terms of wavelength of a specified krypton line, and suggested a number for this equivalence based on the measurements of the cadmium wavelengths and comparisons of the wavelengths of the cadmium and krypton lines.

The effect of this redefinition, if it is adopted by the General Conference, will not be marked. The meter has never been related directly to the chosen krypton line, but the interrelations are sufficiently well known that any differences will probably be within the uncertainties of measurement now existing. It will have the effect of embodying the unit of length in what we believe to be an immutable standard and thus fulfilling an old aspiration. Meter bars, gage blocks, etc., will continue to be used as standards for the kinds of measurements for which they are suited. Every once in a while the length of meter bars will be redetermined in terms of the wavelength of light, instead of determining the wavelength of light in terms of the international meter bar.

The possibility of embodying our unit for time in a physical constant is even more attractive. We can measure time, and its reciprocal, frequency, with the greatest precision of any physical quantity. For example, we may compare the ratio of the average frequencies of two oscillators over concurrent time intervals with as great a precision as we choose. The limit is set by how long the oscillators will operate and how many cycles we wish to count, but such comparisons are pointless if the frequencies of the oscillators are not relatively stable over the interval involved. Conversely, we may measure time with equal precision by counting cycles of a particular oscillator, assuming its frequency is constant.

There is a serious problem in measuring time accurately. We can lay two meter bars side-by-side and compare their lengths. If, by subsequent comparisons, we find that their lengths have not changed relatively we have confidence that our length standards have not changed. But there is no way to lay two time intervals side-by-side; we must rely on the stability of an oscillator to compare time intervals. Man-made oscillators show drifts in frequency with respect to each other, and since oscillators are not passive things like meter bars and kilogram weights, we expect them to drift.

To obtain a good standard for time and frequency we adopted first an astronomical constant, the rotational frequency of the earth which had been regarded as constant since the days of Joshua. Man-made oscillators were used to interpolate for shorter intervals of time and the second was defined as 1/86400 of a mean solar day. Since the apparent solar day varies throughout the

year due to eccentricity of the earth's orbit, astronomers kept track of time by observing star transits, in relation to which earth's rotation is much more uniform.

Precise astronomical observations revealed that this standard was not good enough. The frequency of rotation of the earth is changing with respect to the revolutions of the moon about the earth and the earth about the sun, when allowance is made for perturbations of the revolution time. All planetary motions are in substantial accord. In addition to a gradual slowing down, which is to be expected from tidal friction, there are erratic fluctuations in rotational speed. For this reason astronomers carry out their more precise calculations in ephemeris time, which is based on planetary motions.

With the improvement of quartz-crystal oscillators, seasonal fluctuations in the earth's rotation with respect to the stars have appeared. Though the oscillator frequencies drift, they drift monotonically, allowing us to measure these seasonal fluctuations which amount to 1 part in 10^8 . Correcting for this seasonal fluctuation, astronomers⁵ have established a more uniform time scale, called UT2, good to 1 part in 10^9 , tied in with the earth's rotation, and hence subject to effects of long term changes.

So we see that even the smoothed rotation frequency of the earth is not good enough for a standard. Accordingly the second was redefined in 1956 by the International Committee on Weights and Measures as 1/31556925.9747 of the tropical year 1900 at 12 hours ephemeris time. Why this strange definition? Why not take the sidereal year or the anomalistic year? The lengths of all these years change in known, highly regular ways, so that specification of any epoch was necessary. We chose the tropical year, which is the time between two successive passages of the center of the sun across the celestial equator in the same sense, because accurate tables were already available for its variation, based on the epoch 1900.

The need for a better standard of time became urgent during the past decade with the improvement in microwave techniques. Microwave terms in the spectra of molecules and atoms were being measured with increased precision. The Bureau began to explore these phenomena as the basis for constructing more stable oscillators.⁶ Before 1952 Lyons and his co-workers at the Bureau⁷ had measured the microwave resonance in the ground state of the cesium atom with a precision of 1 part in 10^7 . Essen and his co-workers⁸ at the National Physical Laboratory a few years later increased this precision to a few parts in 10^{10} . It is likely that greater

⁵ U. S. Naval Observatory, "The Naval Observatory Time Service," Circular No. 49; 1954.

⁶ B. F. Husten and H. Lyons, "Microwave frequency measurements and standards," *Trans. AIEE*, vol. 67, pp. 321-328; 1948.

⁷ J. E. Sherwood, H. Lyons, R. H. McCracken, and P. Kusch, "High frequency lines in the hfs spectrum of cesium," *Phys. Rev.*, vol. 86, p. 618; 1952.

⁸ L. Essen and J. V. L. Parry, "The caesium resonator as a standard of frequency and time," *Phil. Trans. R. Soc., London*, vol. 250, pp. 45-69; 1957.

precision can be attained in measuring the cesium frequency and also other atomic frequencies such as those of rubidium, as was indicated by recent work of Bender and Beaty of the Bureau and Chi of the Naval Research Laboratory.⁹ Since the frequencies of these resonances depend on energy levels of the atoms involved, and since, in the case of the cesium, independent experiments have agreed to within the limits of precision, we may presume that they can serve well as standards for time and frequency.

How can we relate these resonance frequencies to the defined unit of time, the ephemeris second? Since the second is defined by an event which occurred over 50 years ago we must measure the resonance frequencies in terms of current values of the UT2 second, and then, through observations on the moon, relate the UT2 second to the ephemeris second. This was done¹⁰ recently, based on four years of observation of the moon, showing that the ephemeris second corresponds to $9,192,632,770 \pm 20$ cycles of the cesium frequency. This is much less precise than the defined value of the second or the precision with which the cesium resonance can be observed.

We have no hope of relating the atomic resonances to the ephemeris second with much greater precision in the near future. We thus are faced with the fact that atomic constants are much better standards for time and frequency than astronomical constants. Furthermore, as standards they are much more accessible than the astronomical constants which require long years of observation to compare them precisely with other quantities. Clearly, we are able today to improve our standard for time by selecting one of the atomic resonances and defining the second in terms of it, making the definition such that the new definition will agree as closely as feasible with the present one.

The frequencies broadcast by the Bureau's stations WWV and WWVH are now monitored and kept as constant as possible by reference to the cesium resonance. The intervals between the seconds pulses are maintained in the same way. Therefore the seconds pulses gradually get out of step with mean solar time. When the difference becomes great enough the pulses are shifted by exactly 20 milliseconds to bring them back in. Thus we are already using two kinds of time, atomic time—that's for the scientists—and mean solar time—that's for the birds and other diurnal creatures.

These attempts to improve time and frequency measurement may seem a quest for precision for precision's own sake, a futile pushing of the decimal point. But this is not correct. It is an attempt to establish standards so that we may learn what physics lies beyond the decimal point. For such things have Nobel Prizes been awarded.

⁹ P. L. Bender, E. C. Beaty, and A. R. Chi, "Optical detection of narrow Rb⁸⁷ hyperfine absorption lines," *Phys. Rev. (letter)*, vol. 1, pp. 311-313; 1958.

¹⁰ W. Markowitz, R. G. Hall, L. Essen, and J. V. L. Parry, "Frequency of cesium in terms of ephemeris time," *Phys. Rev. (letter)*, vol. 1, pp. 105-107; 1958.

We have already learned that the earth turns irregularly on its axis. Now we ask, do time scales based on astronomical, atomic, and molecular processes change with respect to each other as some think they might? (The ammonia maser depends on molecular processes for its frequency stability.) Perhaps from these newly achieved precisions and those soon to be achieved we may even be able to resolve experimentally the famous clock paradox of relativity!

We see that, of the four prototype standards, one is already embodied in a physical constant. Another seems about to be, and a third one is ready to be. What about the standard of mass? We can see no way of embodying that in a physical constant at present without detracting from the accuracy of our system. We seem to be "struck" with the platinum-iridium kilogram. And is that bad?

Table II showed that the kilogram is the most precisely measurable of our prototype standards, that uncertainty in comparing it with other masses is about as small as the uncertainty in comparing the cesium second with the ephemeris second. What could cause the kilogram to change? Any damage to it which would remove 1 in 10^9 of its mass by a scratch or nick would be perceived readily by the naked eye. The oxidation of platinum at normal temperatures is so slow it has never been measured and no sign of an oxide coating on the metal has ever been noticed. We know of one calculable change which will take place in it. One of the isotopes of platinum, Pt 190, is radioactive. It undergoes α -decay with a half-life of about 10^{12} years. Since the abundance of this isotope is only 0.012 per cent it will take about 10^8 years for it to produce a change as great as the imprecision of measurement. We do know that the Arago kilogram, produced during the first half of the last century, has exhibited a loss of mass of the order of a few milligrams. This kilogram was forged from sponge platinum and must have had inclusions of gas which escaped. The kilograms of today are of fused metal and free from changes of this nature. There appears to be no good practical reason for replacing the kilogram with a physical constant.

Physical constants are extensively used for derived standards and for standards embodying magnitudes differing greatly from the defined unit. The International Temperature Scale, which is an approximation to the Thermodynamic Scale, is based on the equilibrium temperatures of various substances. Higher temperatures are measured in terms of the constants in Planck's radiation law. The speed of light is employed in the measurement of large distances in optical and radio surveying. The gyromagnetic precession frequency of the proton,¹ which has been measured with great accuracy at the Bureau, affords a means of comparing two different magnetic fields with high precision. Many other physical constants, too numerous to mention, find employment in extending the range of physical measurement. The determination of the values of these many

constants with the greatest possible accuracy in terms of our prototype standards is one of the main tasks of the Bureau.

CONCLUSION

Our measuring system has developed over the years in such a way as to provide the maximum of precision and accuracy in measuring the many quantities of physics. The standards chosen from time to time to preserve the magnitudes of the units were sufficiently measurable and durable to meet the needs of their days. As scientific advancement required better standards, the persons responsible for preservation of the

system have been ready to adopt new standards for the old ones. Thus we see a transition taking place now from a system based on man-made standards to a system based mainly on physical constants, as far as this is possible.

The prototype standards seem to need little improvement. The advances which we must make in our metrology are in extension of the range over which accurate measurements can be accomplished. Particularly, we need to develop techniques to measure accurately new quantities and new aspects of old quantities to meet the requirements of our expanding technology.

The Ampere*

F. B. SILSBEE†

Summary—The purpose of this paper is to supplement the preceding paper¹ by describing in some detail the various measuring procedures and supporting research which must be carried on by a national standardizing laboratory to meet its responsibilities related to a single one of the many units of measurement on which modern science, engineering and industry are based. The example chosen is the ampere, the unit of electric current in the MKSA system. The tasks involved naturally fall into five successive stages, namely; the definition, establishment, maintenance, extension and dissemination of the unit.^{2,3}

DEFINITION OF THE AMPERE

ALMOST a century ago, a committee of the British Association for the Advancement of Science, under the chairmanship of Clerk Maxwell and guided by the ideas of Weber and Gauss, defined an electrostatic and an electromagnetic system of electric and magnetic units based upon the centimeter, the gram, and the second, as the fundamental mechanical units. For the electromagnetic system, the further assumption was made that the permeability of empty space was to be regarded as a dimensionless quantity numerically equal to unity. (In the electrostatic system, the permittivity of empty space was regarded as a dimensionless quantity numerically equal to unity.) The resulting electromagnetic units of voltage and resistance were found to be inconvenient in magnitude;

and in the 1890's, units larger by factors of 10^8 and 10^9 and named the volt and the ohm, respectively, came into use as the basis for the so-called "practical" electrical units. To retain the coefficient in Ohm's Law at the convenient value of unity required that the practical unit of current, the ampere, should be 1/10 of the cgs electromagnetic unit.

At that time it was considered important that the electrical units be so defined that they could be reproduced in any laboratory with a minimum of inconvenience. In the case of electric current, there was wide use of the coulometer, in which the amount of metal deposited from an electrolytic solution in a measured time is taken as the measure of the average current. In 1908, an international electrical congress was held in London, at which the "International Ampere" was defined as the unvarying current which would deposit silver at the rate of 0.00111800 gram per second from an aqueous solution of silver nitrate.

With the establishment of national standardizing laboratories in the industrial nations, the need for ready reproducibility of the ampere became much less important. The succession of studies at the National Bureau of Standards of the silver coulometer and of the iodine coulometer, which had been suggested as an alternative basis for the definition of the ampere, had shown disconcerting discrepancies and sources of error; there were, in particular, errors from the inclusion of solution in cavities in the deposited silver, the presence of complex ions, and later the recognition of the isotopic complexity of silver. Other studies at the national laboratories showed that methods of measuring the ampere in terms of length, mass, and time, while still difficult

* Original manuscript received by the IRE, December 9, 1958; revised manuscript received February 19, 1959.

† National Bureau of Standards, Washington, D. C.

¹ A. G. McNish, "The basis of our measuring system," *Proc. IRE*, this issue, p. 636.

² F. B. Silsbee, "Establishment and Maintenance of the Electrical Units," *NBS Circular* 475; June, 1949.

³ F. B. Silsbee, "Extension and Dissemination of the Electrical and Magnetic Units by the National Bureau of Standards," *NBS Circular* 531; July, 1952.

and time consuming, were capable of fully as great an accuracy as that obtainable by a coulometer. Accordingly, effective January 1, 1948, the International Conference on Weights and Measures, which by international treaty has authority to coordinate the standards of measurement in the field of electricity as well as of length and mass, promulgated as the definition of the ampere:

"The ampere is the constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular sections, and placed 1 meter apart in a vacuum, will produce between these conductors a force equal to 2×10^{-7} mks unit of force per meter of length."

ESTABLISHMENT OF THE AMPERE

The particular geometrical arrangement described in the foregoing definition is excellently adapted to concise and precise wording and therefore to forming the basis for legislative actions. On the other hand, the mechanical force involved is so small that the simple arrangement as described is not appropriate for the accurate experimental realization of the ampere. Use is therefore made of more complicated circuit arrangements, in which the force produced electromagnetically between two coils connected in series and carrying the current to be measured is measured mechanically. With such an arrangement, the mechanical effect is proportional to the square of the current and to the derivative of the mutual inductance between the two coils with respect to some parameter such as a) the position of one coil with respect to another coaxial with it, or b) the angle between the axes of two concentric coils.

In the former (type a), usually called a "current balance," one of the coils is suspended with its axis vertical from one arm of a precision balance. The current is adjusted until the change in the mechanical force resulting from a reversal of the direction of the current in one coil is balanced by the addition or removal of a known mass in a pan on the same arm of the balance. A balance of this type was suggested by Lord Rayleigh, in which the fixed coil consists of two parts, each of which is itself a circular coil, the cross-sectional dimensions of which are small compared with its radius. The single smaller moving coil is of similar shape and hangs midway between the two halves of the fixed coil. The coils can be wound with numerous turns of fairly fine insulated wire, thus yielding a conveniently large mechanical force. A clever feature of this arrangement is that the major term in the mathematical expression relating the force to the current involves only the ratio of the radii of the small coil to that of either large coil, and this ratio of the effective radii may be determined experimentally by bucking the magnetic effects of the two coils. The best insurance against systematic errors lies in the comparison of measurements made by methods as different as possible. This is especially important at a national laboratory which has no source of higher accuracy to con-

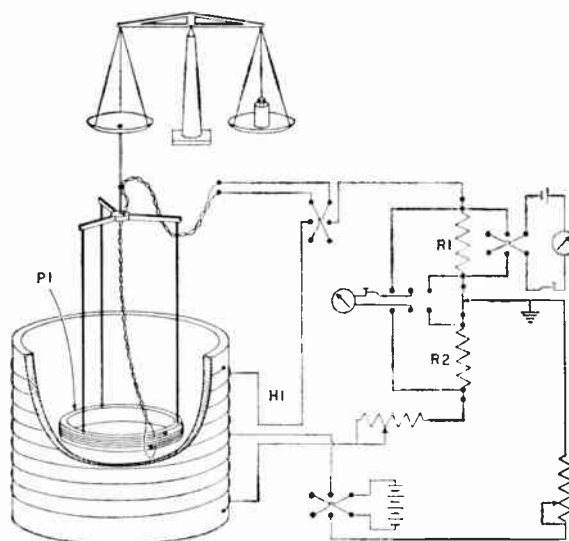


Fig. 1—Diagram of NBS current balance. The movable coil hangs in the plane of the tap at the center of the single layer primary coil. One galvanometer serves for adjustment of equality of currents in the two halves of the fixed coil. The other galvanometer balances the emf of the standard against the voltage drop in resistor R_1 .

sult. Hence, in this case, experiments with different combinations of coil sizes were made and found to show small but significant inconsistencies which were apparently attributable to uncertainty in the exact location of the current-carrying conductors. Accordingly, this type of balance has been abandoned for one in which the two parts of the fixed coil are the upper and lower halves of a single-layer solenoid energized through a tap at the center of its winding, while the moving coil consists of a short single-layer solenoid concentric and coaxial with the fixed coil. (See Fig. 1.) With this arrangement the mechanical force is undesirably less (1.5 grams) with the balance of Driscoll and Cutkosky,⁴ but the position of the current-carrying parts can be measured with much greater accuracy, and the force per ampere still can be computed though more complex mathematics is required.

The apparatus of type (b) is usually called an electro-dynamometer. Such apparatus, in a form originally suggested by Pellat, has recently been used⁵ to corroborate the results by the current balance. In this apparatus a relatively small (11.6 cm diam., 9 cm long) rotatable coil, with its axis vertical, is placed concentric with a long horizontal coil (28 cm in diam.) A lever arm (25 cm long) attached to the movable coil can be loaded with a weight (approx. 1.5 gram) sufficient to balance the change in turning moment produced when the current (1.02 ampere approx.) in the fixed coil is reversed.

In either of these types of apparatus a force is meas-

⁴ R. L. Driscoll and R. D. Cutkosky, "Measurement of current with the National Bureau of Standards current balance," *J. Res. NBS*, vol. 60, pp. 297-305; April, 1958.

⁵ R. L. Driscoll, "Measurement of current with a pellet-type electro-dynamometer," *J. Res. NBS*, vol. 60, pp. 287-296; April, 1958.

ured by balancing it against the force, MG , of gravity acting on a known mass, M . It is estimated that the uncertainty in the generally accepted value of G^6 may contribute at least 3 ppm to the over-all uncertainty in the measurement of a current by such a method. Work is now in progress in several laboratories in an effort to obtain a better accuracy for g . Alternative methods have been proposed in which the electromagnetic torque is balanced by the torsional stiffness of a suspension. The stiffness of the suspension in turn is evaluated by timing the periods of a torsional pendulum, which uses the same suspension, before and after its moment of inertia is increased by a definite and computable amount.

In October, 1946, the International Committee on Weights and Measures had available data from experiments at the British and American national laboratories, using current balances of two types. On this basis it recommended that each national laboratory should, on January 1, 1948, shift the assigned values of its standards by such an amount as would bring them all into agreement with a new value of the ampere equal to 1.00015 times the mean International Ampere as it had been previously realized. For the NBS this factor was 1.000165.⁷

It is hoped that the various laboratories will maintain their standards at the values thus assigned into the indefinite future. If and when data derived by new and better methods of establishing the ampere on the basis of the "absolute" definition have become available, the International Committee may recommend some further adjustment in its value. In the meantime, the invention and application of such new and more accurate methods remains an important and continuing challenge to each national laboratory.

Recent results at the NBS^{4,5} by both Pellat and the NBS types of apparatus have shown an agreement to 5 ppm. The weighted mean by the two methods indicates that the ampere as currently certified by the NBS on the basis of the international agreement of 1946 is probably 1.000010 ± 0.000005 absolute amperes.

MAINTENANCE OF THE AMPERE AND VOLT

It is, of course, not practical, even after the ampere has been established with the utmost accuracy, to lay a standard current aside on a shelf for future reference. (It should be noted that by the use of superconducting circuits maintained at liquid helium temperatures such storage of a definite current is not only theoretically possible but has actually been accomplished.) Current procedure, therefore, is to arrange for the measured, definitely known current to flow in a standard resistor of known value and to compare the electromotive force of a saturated standard cell with the voltage drop produced in the resistor by the measured current. This

process enables one to assign the correct numerical value in absolute volts to the emf of the standard cell. The establishment of the proper value to use for the standard resistor is yet another long story, for which space is not here available.

For the short term maintenance of the ampere (or more strictly speaking, for the maintenance of the ability to reproduce it readily at any time as desired) use is made of reference groups of standard cells and of standard resistors. The saturated cadmium standard cell exhibits an electromotive force which settles down a few months after construction to a value which thereafter normally remains constant for many years to one or two parts per million. At the NBS, the reference group is one of 50 such cells, manufactured at various times and in some of which deuterium oxide has been substituted for the water of crystallization and for the water used as a solvent. Two or three times per year they are readily intercompared with a precision of 1 in 10^7 . If any cell begins to show an abnormal rate of drift relative to the mean of the others, it is removed from the group and replaced by another which previously had shown good stability while it was stored in the reserve stock. The ohm is similarly maintained by a group of 10 standard resistors of the double wall type.⁸

Intercomparisons with the units maintained by the standards of emf and resistance of other national laboratories are made periodically at intervals of two or three years at the International Bureau of Weights and Measures. The precision of these international comparisons is slightly less than that between members of a reference group because of the possibility of changes in the standards during shipment. Nevertheless these comparisons with groups maintained in different countries under quite different circumstances gives a valuable check on the constancy of the units. For example, the best available estimate of the spread which developed between the American ampere and the British ampere between 1910 and 1948 is 15 ppm. These values were readjusted to a common basis on January 1, 1948.

There is, of course, always the possibility that the standards of all the nations are drifting together in the same direction at the same rate. Such comparisons would not detect the drift. Attempts to check the constancy of the ohm by the use of standards of pure materials such as platinum or mercury have not proved satisfactory, largely because of the relatively large temperature coefficient of resistivity of all pure metals. The development of new forms of standard cell involving radically different chemical reactions and therefore not liable to the same rate of drift as the present cadmium cell, has received some attention and currently is a challenging line for future development.

Of course, the repetition of experiments with any form of current balance will give a definite assurance that the unit has not drifted by an amount exceeding the accu-

⁶ H. L. Dryden, "A reexamination of the Potsdam absolute determination of gravity," *J. Res. NBS*, vol. 29, pp. 303-314; November, 1942.

⁷ E. V. Condon, "Announcement of changes in electrical and photometric units," *NBS Circular 459*; May, 1947.

⁸ J. L. Thomas, "Stability of double-walled manganin resistors," *J. Res. NBS*, vol. 36, pp. 107-110; January, 1946.

racy with which the absolute measurement can be made. Unfortunately, this accuracy (estimated at 5 ppm) is not as great as that (about 0.1 ppm) with which standard cells or standard resistors can be intercompared. Hence the repetition of absolute measurements at intervals of a decade, while essential, is not completely satisfactory. It is to be noted that such repetition of an absolute ampere experiment with the same apparatus at a later time is capable of checking the constancy of the combination of standard cells and resistors with accuracy greater than that with which the absolute voltage of the cell can be determined initially, because truly constant errors in the theory and design of the current balance, being the same at the beginning and the end of the interval, would cancel out in the computation of the drift. The most recent such check in 1958⁹ shows an apparent decrease of 6 ppm in the ampere since the previous measurement in 1942.⁹ This is so small that it may well be the result of experimental error.

It will be noted that during the period from 1908 to 1948, while the ampere was defined in terms of the rate of deposition of silver, the faraday¹⁰ was a matter of definition as being 96,494 international coulombs per gram equivalent, and experiments with the silver coulometer were the means for establishing the ampere. Since January 1, 1948, the current balance has become the means for establishing the ampere and the faraday has become a natural constant to be determined experimentally. Since it is now inherently a constant of nature, a pair of experiments by which it was measured on two successive occasions would constitute a check on the constancy of the standards for the ampere between those dates. Here also, constant errors, such as those arising from uncertainty in the isotopic ratio of the supply of silver used, would cancel out. A project currently nearing completion at the NBS on the determination of the faraday¹¹ has as an anticipated by-product such a check on the constancy of the ampere and the volt. Unfortunately, experimental difficulties would seem to limit the accuracy of this check to about 10 ppm.

Recent work has developed a method for the measurement of the gyromagnetic ratio of the proton in relatively weak fields (e.g., 12 gauss) with an accuracy approaching 3 ppm. This type of experiment seems to offer possibilities for a valuable check on the constancy of the ampere.¹² In this method a sample of water is placed in the magnetic field produced inside a large single-layer solenoid of accurately measured dimensions. The sample is enclosed in a thin spherical glass or lucite container about 2 cm in diameter. It is initially polar-

ized in a field of 5000 gauss produced by a permanent magnet. It is then shot pneumatically into position in the center of the single-layer solenoid. The magnet is at a distance of over 40 feet and so oriented that it does not affect the field at the solenoid. A short pulse of RF magnetic field at approximately the precession frequency is then applied perpendicular to the main solenoid field. This causes the proton magnetization to spiral out from the solenoid axis and to start precessing around it in synchronism. A signal is thus induced in a pick-up coil close to the sample. The frequency (about 52 kc) of this signal is proportional to the field and is readily measured with ample accuracy by electronically timing the interval during which a predetermined integral number of precession cycles occur. Once the proportionality factor is known, the magnetic field at any later time can be computed from the observed precession frequency and the current can then be computed from the dimensions of the coil. If the current is also measured in terms of a standard cell and a standard resistor, a comparison of the two values of current will check the combined constancy of the cell and resistor. Work is currently under way looking toward the simplification of such experiments as a check on the stability of these standards in the future.

Still another method for checking the constancy of the ampere can be based on the resonance frequency at which the rubidium atoms undergo a quantum transition in a magnetic field.¹³ An absorption cell containing rubidium vapor is placed at the center of a single-layer solenoid which carries the current to be measured. A circularly-polarized beam of light from a rubidium vapor lamp, filtered to pass only the 7947 Å line, passes through the cell parallel to the axis of the solenoid and is monitored by a photo cell. Under these conditions, as a result of "optical pumping," the Rb atoms tend to line up with the field and absorb less light than if randomly oriented. When a coil, with its axis perpendicular to the field, is energized with RF of a frequency corresponding to the Zeeman transition, the atoms will change their orientations and the absorption will increase. The frequency for the lowest transition is about 700 kc per gauss. By sweeping the RF through its resonance value and measuring the frequency at the minimum light transmission, the field, and from it the current, can be derived. It has been found possible to determine the absorption frequency to better than 1 cps.

EXTENSIONS AND APPLICATIONS OF THE AMPERE

The ability to reproduce accurately a current of one ampere at any time is, of course, merely the first step in the problem of the measurement of electric current. On this foundation, the next stages are to progress upwards and downwards to larger and smaller direct currents, and also to transfer the process of measurement to

⁹ R. W. Curtis, R. L. Driscoll, and C. L. Critchfield, "An absolute determination of the ampere, using helical and spiral coils," *J. Res. NBS*, vol. 28, pp. 133-158; February, 1942.

¹⁰ A faraday is defined as the electric charge required for an electrochemical reaction involving one chemical equivalent.

¹¹ D. N. Craig and J. I. Hoffman, "Determination of the Faraday Constant by the Electrolytic Oxidation of Oxalate Ions," *NBS Circular* 524; August, 1953.

¹² P. L. Bender and R. L. Driscoll, "A free precession determination of the proton gyromagnetic ratio," *IRE TRANS. ON INSTRUMENTATION*, vol. I-7, pp. 176-17; December, 1958.

¹³ T. L. Skillman and P. L. Bender, "Measurement of the earth's magnetic field with a rubidium vapor magnetometer," *J. Geophys. Res.*, vol. 63, pp. 513-515; June, 1958.

alternating currents, and thence to extend the methods of measurement to higher frequencies and to pulses.

The extension to smaller currents involves the development of more sensitive galvanometers¹⁴ and procedures for calibrating them by sending a tiny current through a large enough resistance to produce a voltage comparable with the emf of a standard cell. At the extreme of the range the construction of a stable and properly insulated resistor becomes a serious problem.

An alternative calibration procedure available at the very low ranges in reality constitutes an independent absolute measurement of current. It is based upon the rate at which the charge transferred by the current raises the voltage developed at the terminals of a known capacitor. In this process the value of the capacitor is determined in terms of frequency and the unit of resistance, while the voltage measuring device is checked against the standard cell.

Another line of approach to the measurement of small currents is to introduce a "chopper" or reversing switch by which the tiny direct current is converted to an alternating current which can then be amplified electronically and measured at a higher level. Much research and development is required to minimize the numerous possible sources of error in such an arrangement.

Several other developments to extend the range of measurement to very small currents are based on a relay principle. The slight motion of a galvanometer coil in response to the current, while still too small for visual detection, can be detected and measured by a sensitive pick-up arrangement, which may take the form of a mirror light-beam and photo cell, of a change in capacitance, or of a change in inductive coupling. The resulting amplified signal is then easily read or recorded. The art has now progressed along these several lines to the limit set by the particulate nature of matter, usually described as "Johnson noise" or "Brownian motion."

The extension of measurement based on the ampere to larger direct currents is relatively simple. It involves the design and construction of standard resistors of the 4-terminal type of adequate size and adequately cooled so that they can carry the large current to be measured without undue rise in temperature and consequent uncertainty in resistance. Such equipment at the NBS can measure 10,000 amperes with an accuracy approaching 0.01%. Using such equipment, it is possible to accurately determine the resistance of shunts of even lower resistance and greater current capacities up to 60,000 amperes. This extreme of the range is also subject to insidious sources of error, such as irregularities in the distribution of current in the heavy terminals and effects from mechanical strains developed by differential thermal expansion of parts of the resistor caused by temperature gradients.

¹⁴ F. Wenner, "General design of critically damped galvanometers," *Bul. NBS*, vol. 13, pp. 211-244; August, 1916.

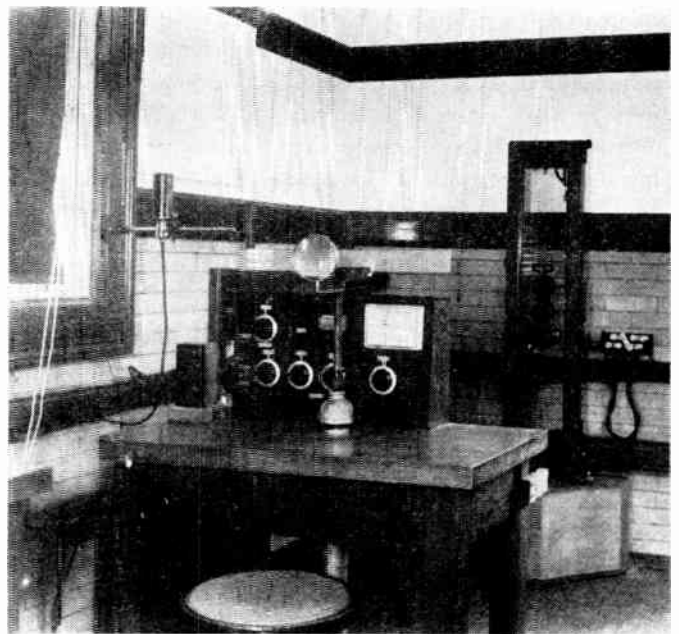


Fig. 2—Electrodynamic transfer instrument. Fixed coils, embedded in plastic, are supported in glass slabs. Mirror on moving coil reflects indicating spot on glass scale at distance of 2 meters.

The step from the direct current standards discussed hitherto to standards for measuring alternating currents of power frequencies which form the basis of the ten-billion-dollar electric power industry rests jointly on the use of electrodynamic and electrothermic instruments. The extension of current measurements to the higher frequencies used in audio and radio communication rests primarily on the thermal converter.

At the NBS an electrodynamic instrument,¹⁵ the fixed coils of which are rated at 2.5-5-10 amperes and the moving coils of which are rated at 0.1 ampere, is arranged with a known standard resistor in series with its fixed current coil (See Fig. 2). The movable coil is connected in series with an appropriate series resistor and to the potential terminals of the standard resistor. The instrument thus becomes, in effect, a wattmeter, the deflection of which is a measure of I^2R -loss in the resistor and thus of the current. The performance of this instrument has been checked both by comparison with electrothermic instruments and by using it as a wattmeter to measure the loss in a large capacitor, the loss factor of which was known from ac bridge measurements. This check measurement being at a very low power factor constitutes a much more severe test of the transfer instrument than is needed to justify its operation as an ammeter at unity power factor. At frequencies above 3000 c the torque produced by eddy currents induced in the aluminum damping vane begins to be appreciable and limits the frequency range of the instrument.

Thermal converters consisting of a thermocouple, the

¹⁵ J. H. Park and A. B. Lewis, "Standard electrodynamic wattmeter and ac-dc transfer instrument," *J. Res. NBS*, vol. 25, pp. 545-579; November, 1940.

measuring junction of which is in thermal contact with an electric heater and the electromotive force of which is measured by a potentiometric circuit, offer a much more flexible basis for the transfer of the ampere from dc to ac which can be used with an accuracy approaching 0.1% up to 30 kc. Equipment is now available at the NBS¹⁶ for measuring currents up to 50 amperes by this means. Here again, errors may arise from lack of symmetry in the thermoelement with its resulting errors from unbalanced Peltier and Thomson effects. For still higher frequencies, thermocouple instruments can be used, but with steadily decreasing accuracy, up to frequencies at which the concept of electric current ceases to have meaning and the phenomena are more suitably described in terms of microwave propagation.

At power frequencies and, more recently, at frequencies of 400 and 1000 c, the measurement of alternating current must be extended from the range of 5 a, at which it is transferred from dc, to the much higher currents used in electric energy transmission. This is done by the use of current transformers, which have the valuable property of reproducing in the instruments connected to their secondary terminals a current which is closely similar in phase and reduced in scale by a definite ratio from the much larger current in their primary circuit. The usual way to accurately determine the ratio of transformation and the small phase displacement in a current transformer is to use two known resistors connected in series with the primary and with the secondary windings, respectively. The impedance drops in the resistors are opposed through a detector, and the secondary resistor is adjusted for a balance. The effect of any phase displacement is balanced and measured by a mutual inductor, the primary of which is energized by the secondary of the transformer, while the secondary of the inductor is in series with the detector. It is essential that the 4-terminal resistors used have small and definitely-known self-inductances and small or negligible skin effect. Above about 2000 amperes the construction of such resistors becomes impracticable. The range of ac current measurement can be extended further by using a standard current transformer which is provided with a number of magnetically equivalent sections of primary winding. If the transformer rating, with its primary sections connected in series, is less than 1000 amperes it can be calibrated for a given frequency and burden over its full range. It can then be reconnected for a higher rating by using more sections in parallel and fewer in series. At the same frequency and burden, but with the new connection, it will have the same phase angle but will have a ratio inversely proportional to the numbers of primary turns in the two connections. Measurements of 60-cycle currents are regularly made at the NBS by such a standard transformer to 12,000 amperes (See Fig. 3). The method is

¹⁶ F. L. Hermach, "Thermal converters as ac-dc transfer standards for current and voltage measurements at audio frequencies," *J. Res. NBS*, vol. 48, pp. 121-138; February, 1952.

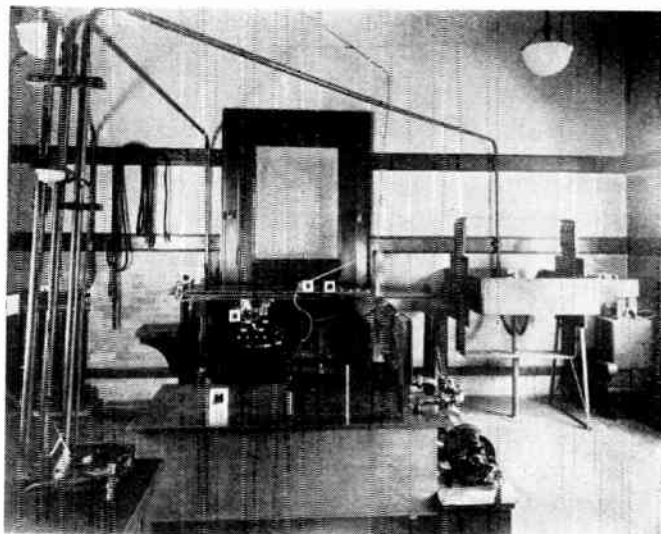


Fig. 3—Equipment for testing current transformers to 12,000 amperes. The transformer under test is mounted at center of "cage" formed by the 4 parallel 6" X 1" bus bars, and compared with the standard transformer S. This, in turn, is calibrated on 1200-ampere connection in terms of oil-cooled 4-terminal ac resistors in tank M.

apparently applicable even to considerably larger currents.

The ampere is extended and applied in still another field in the measurement of large currents of short duration, such as are met with in the testing of lightning arresters and more recently in particle accelerators and perhaps in the future in nuclear fusion apparatus. (See Fig. 4). Here again, the measurement of the current in terms of the ampere must be based upon a measurement of the voltage drop across a resistor, or alternatively, an inductor (self- or mutual-). In the latter case the directly observed quantity is, of course, the time derivative of the current, but in studying transient phenomena this is often of greater significance than the current itself. In the measurement of such transients, the duration of which is measured in microseconds, the voltage must be recorded by a cathode ray oscillograph, so that it is desirable to have an IR -drop of several thousand volts in contrast to the fifty-millivolt drop which is common in the measurement of sustained direct currents. The design of appropriate standard resistors, which must be compact so as to minimize inductance and at the same time sufficiently massive to absorb the energy dissipation, involves much research and development. Nevertheless, measurements can readily be made up to 200,000 amperes on surges lasting for 10 to 50 microseconds.¹⁷

DISSEMINATION OF THE AMPERE

The ability of a national standardizing laboratory such as the NBS to measure currents of these ranges of magnitude and frequency in its own laboratories would be useless without the provision of arrangements by which the ampere can be disseminated so that measurements may be made in terms of it at all of these

¹⁷ J. H. Park, "Shunts and inductors for surge current measurements," *J. Res. NBS*, vol. 39, pp. 191-212; September, 1947.

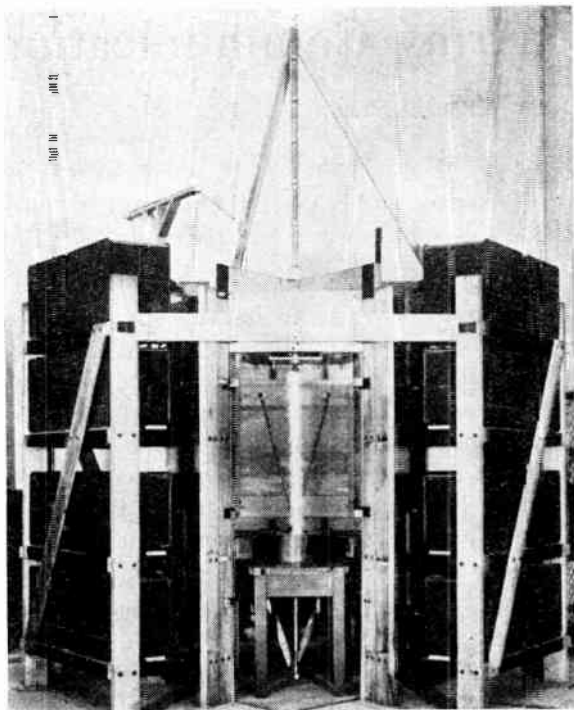


Fig. 4—A 100,000-ampere arc. This 50,000-joule discharge from the capacitor bank of 10 microfarads initially at 100,000 volts is measured by the concentric tubular shunt below the table.

levels throughout the nation.¹⁸ This dissemination process is carried on by cooperative action between the NBS on the one hand and a large number of electrical standardizing laboratories on the other.¹⁹ Some of these laboratories are in the factories which manufacture measuring apparatus (standard cells, resistors, potentiometers, shunts, ammeters, current transformers, etc.). The working standards of such laboratories form the basis for the calibration of the apparatus made and sold by their respective firms, and directly supply the purchasers with consistent calibrations of their new apparatus. A much larger number of laboratories connected with universities, electric power companies, other manufacturing concerns and Government agencies, also derive standards of measurement from the NBS and use them in their periodic checks on the shop instruments and measuring apparatus used in their respective organizations. Still other "primary reference laboratories," either governmental or commercial, having derived their standards of measurement from the NBS, use them to calibrate the reference or interlaboratory standards of "secondary" laboratories at a lower echelon.

The dissemination of the ampere and its multiples from the NBS is accomplished primarily by the testing of standard cells and standard resistors. Each year about 1000 standard cells are submitted for test and are later used with appropriate resistors and potentiometers

to establish currents of known value over a wide range. Until recent years most standard cells used in industry were of the unsaturated type, which have a temperature coefficient so small as to be negligible for most purposes, but which show a gradual decrease in emf of about 80 microvolts per year. Only the two manufacturers of such cells and a very few of the largest electrical manufacturing and power companies troubled to maintain banks of standard cells of the more permanent saturated type in thermostatted enclosures. In the last few years there has been a marked shift in practice and, in the last year, over 30 companies submitted groups (usually of 6) of saturated standard cells for test.

At a lower level of accuracy some organizations derive their ampere by submitting for test high grade self-contained ammeters, or "shunts" (4-terminal standard resistors of large current-carrying capacity) to be used with millivoltmeters to measure large currents. About 50 interlaboratory standard ammeters and 75 shunts are submitted to the NBS per year.

In the field of large alternating currents, the NBS provides a calibration service²⁰ to the power industry by determining with an accuracy of 0.05% and 1 minute in phase angle the value of the ratio of transformation and the phase displacement in standard current transformers. It is relatively easy to compare one such transformer with another of the same nominal ratio, and apparatus based on developments at the NBS is widely available for this measurement. It is therefore necessary for power companies to submit to the NBS only one laboratory standard transformer for each current range and, with it as a basis, to calibrate the rest of their standard and working current transformers in terms thereof. Annually, about 75 transformers, each embodying on the average 9 different ranges of primary current, are tested. This service is available up to 12,000 amperes at 60 c and to 200 amperes at 400 c.

The foregoing paragraphs have shown that the responsibility for one unit of measurement, the ampere, is inextricably tied in with other units, particularly the ohm and the volt; that its definition involves official international cooperation; that its establishment involves difficult and expensive instrumentation based directly on the most fundamental standards; that its maintenance involves research in apparently remote fields such as electrochemistry, nuclear spin theory, hyperfine spectroscopy; that its extension involves engineering skills ranging from the handling of the tiny current in an atomic beam to the surges of an artificial lightning stroke. To insure that the ampere shall be consistent with the other units of science and engineering, that it is constant through the years and uniform throughout the nation, involves responsibilities which can be met only by a single central governmental laboratory.

¹⁸ F. B. Silsbee, "Standards for electrical measurement," *Phys. Today*, vol. 4, pp. 19-23; January, 1951.

¹⁹ F. B. Silsbee, "Suggested Practices for Electrical Standardizing Laboratories," NBS Circular 578; August, 1956.

²⁰ F. B. Silsbee, R. L. Smith, N. L. Forman, and J. H. Park, "Equipment for testing current transformers," *J. Res. NBS*, vol. 11, pp. 93-122; July, 1933.

Progress and Problems in U. S. Army Communications*

RAYMOND E. LACY†, SENIOR MEMBER, IRE

Summary—A review is given of communication progress under the cognizance of the U. S. Army Signal Research and Development Laboratory since World War II. The strong dependence on and close cooperation maintained with non-Government institutions is emphasized. Operational and technical areas are discussed and equipment shown to illustrate the present state-of-the-art in meeting communication requirements for U. S. Army preparedness.

INTRODUCTION

AS the primary research and development establishment of the U. S. Army Signal Corps., the U. S. Army Signal Research and Development Laboratory has a large interest in all aspects of communications. Its successful technical history is usually known to those in the communication professions. Today it faces a still greater challenge. The demands for communications, made by the armies which must not only fight under atomic-age conditions but also be prepared for space-age implications, place definite and tangible technical problems directly on the doorstep of the Communications Department of the USASRD.

The communication requirements for a so-called Type Field Army, which is an operational organization element, have increased from 216,000 miles in 1948 to 976,000 miles in 1958.¹ Projected requirements estimate that a Type Field Army will require 71,000,000 miles of communications by 1970. It is desired that these objectives be met while simultaneously reducing sharply applicable research-and-development-cycle time, by systems which will not involve any increase in weight, or any increase in the number of personnel required, and by systems having capabilities for eliminating direction-finding and jamming by opposing forces.

The hopeful part of these expectations is that the exponential law of communication increase followed during the last decade previous to 1958 was not only met, but was met with a 3 per cent reduction in monetary cost of equipment, with a 52 per cent reduction in the weight of the equipment, and with a 73.5 per cent reduction in the number of communications personnel required in connection with it. This sets an operational pattern; it also indicates that research and development can be expected, based upon past performance, to meet the demands of the future. Let us view research and development details directed toward meeting these somewhat awesome but very tangible objectives.

* Original manuscript received by the IRE, December 12, 1958; revised manuscript received, February 17, 1959.

† U. S. Army Signal Res. and Dev. Lab., Fort Monmouth, N. J.
¹ "Staff Study on Quantity of Communications (U)," U. S. Army Electronic Proving Ground, Fort Huachuca, Arizona, Proj. 54-57-0015, USAEPG-SIG 940-36; July, 15, 1958.

THE POSITION OF USASRD RELATIVE TO NON-GOVERNMENT INSTITUTIONS

USASRD is a Government Laboratory whose *modus operandi* is to use all its allocated resources to focus the maximum and best technical efforts anywhere available on the problems involved in providing the best possible communication equipment for the Army. As a result, such equipment from basic principles to the final, reliable bolts and nuts, is contributed as appropriate by many persons in many organizations, both Government and non-Government, universities and other nonprofit institutions, and/or in the profit-making highly-competitive industrial organizations. All these institutions not only play a major and an essential part in the research and development of the equipment, but also in the shaping of policies and procedures used throughout its early research, later development, and eventual mass production.

ATOMIC-AGE REQUIREMENTS

The atomic age has required a completely different concept of Army communications.² Briefly, this has greatly affected two areas: the area of communications in which the soldier operates, and the area in which the highest-level commander operates. Dispersion and mobility require that tactical vehicular radio equipment, for instance, be operable over transmission distances far beyond those expected of any municipal police network. Also, before an atomic weapon is fired, a top-level decision may be required to be communicated between the echelon doing the firing and the high-level installation which is authorized to make such a firing decision. As a result, long-distance communications must be viewed as a part of the old-time front-line-type tactical network, and no delay in transmission tolerated, even though the distances covered may span continents and oceans.

BREAKDOWN INTO TECHNICAL FUNCTIONS

For convenience in this presentation, the technical problems involved are broken down into seven major functional areas and several minor areas. Three of these major technical areas are related to radio equipments and transmission; one to multiplexing; one to switching; one to data processing, and the last to systems concepts.

Forward Area Tactical Radio Sets

Just previous to World War II, two fateful technical

² Lt. Gen. A. G. Trudeau, "Army Research and Development in C and E," *Signal*, pp. 19-22; June, 1958.

decisions were made. One was to use FM for forward-area tactical Army radio sets. The other decision was to utilize quartz crystal as the radio-carrier frequency-stabilizing elements in these equipments as a basis for the large number of channels needed for the newly evolved, highly mobile armor operations.³ Subsequent events all through World War II, where such crystal-controlled-FM equipments saw action in all services of the Armed Forces of the Allies, showed the correctness of these decisions. An efficient and quiet noise-free network-type of operation was provided. Amazing as it seems, considering that this was less than 20 years ago, the radio-carrier-frequency spectrums then used (20 to 50 mc) were almost totally unoccupied. Also, it was sufficient in most military operations for such equipment to net with itself. The Army operated in units that were relatively concentrated to hold all ground occupied, leaving no gaps between units. The communication coverage of radio sets with ranges of less than 15 miles was usually sufficient.

A program for continued improvement of the FM systems making full use of current transistorization and other miniaturization is continuing. Typical of such equipment is the radio set shown in Fig. 1. This is a transceiver which provides a choice of any four preset channels in the 30-70-mc spectrum. It utilizes transistors completely and has a nominal distance range of one mile.

A new concept, though, has evolved within the Army Division. Dispersal and mobility have extended the distance requirements for such forward-area tactical radio sets to 25 and to 50 miles rather than the 5 to 15 mile limit needed in World War II. This dispersal has placed an even greater emphasis on radio and required many more channels per unit. This, together with the increasing congestion in the RF spectrum, has resulted in the Army's joining the general trend toward the use of single sideband modulation as a means for more economical utilization of the RF spectrum.⁴

Even so, this single-sideband equipment is viewed as an interim approach to the operational problems, as it contains no provisions for reducing intentional interference or for insertion of voice-security techniques.

Radio Relay

In the area of radio relay, technical history again was achieved during World War II. A whole new vista was apparent from the success of what was then a new radio operation which provided a wire-type communication, but without the wire. The term "Antrack" (AN/TRC) became almost synonymous for VHF-radio relay.⁵ Also the first microwave-radio relays which were evolved dur-

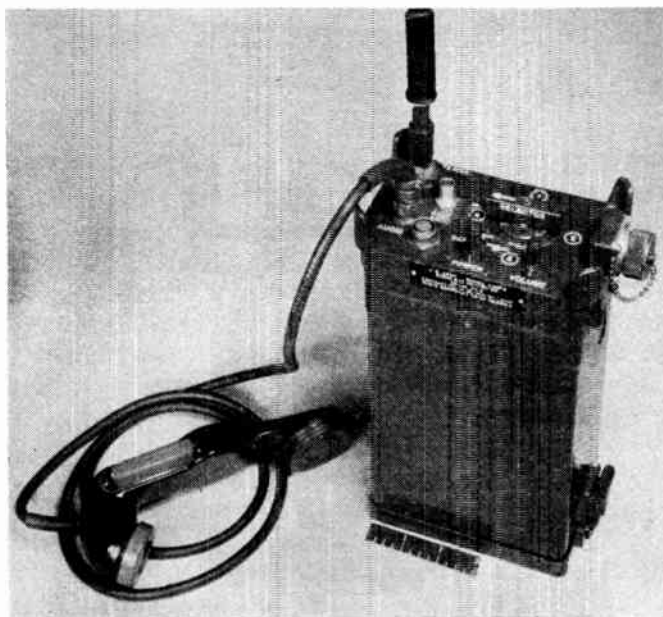


Fig. 1—Radio set, AN/PRC-36 (Devel. model).
Manufactured by RCA.

ing World War II were the forerunners of the radio-type long-lines systems that have subsequently interlaced the nation.⁶

Post-World War II research and development have devoted much effort to making radio relay an even more useful adjunct to Army communications. A whole family of new equipments has been developed and produced. Techniques are currently being investigated to optimize radio-relay systems for handling digital traffic.

A great reliance by the Army on radio relay has been necessitated by increased mobility. As a consequence, development of smaller equipment with increased channel capacity (Fig. 2) and lightweight quickly erectible antenna systems has been necessitated. Also automatic-switchover devices (Fig. 3) for controlling unattended stations and eliminating circuit outages due to equipment failures have been provided.

A tropo-scatter equipment operating in the frequency range of 360 to 600 mcs has been developed (Fig. 4), but here again the main limitation on the use of tropo-scatter phenomena is the equipment mobility.

Similar to the vehicular and portable combat-area radio-communication equipment trend, Army radio-relay development is attempting to employ single-sideband modulation to provide greater range and higher traffic capacity at carrier frequencies well up into the X-band microwave ranges.

Long-Range Ionospheric Radio

Most long-distance radio transmissions have been using the ionosphere to reach around the curvature of the Earth. The result is that this particular portion of

³ R. B. Colton, "Army ground communication equipment," *Elec. Eng.*, vol. 64, pp. 173-179; May, 1945.

⁴ A. M. Creighton and D. B. Reeves, "A Single Sideband Radio Central to Replace Military Wire Lines," *Signal*, pp. 33-37; March, 1958.

⁵ W. S. Marks, Jr., O. D. Perkins, and W. R. Clark, "Radio-relay communications systems in the United States Army," *Proc. IRE*, vol. 33, pp. 502-522; August, 1945.

⁶ R. E. Lacy, "Two multichannel microwave relay equipments for the U. S. Army communications network," *Proc. IRE*, vol. 35, pp. 65-70; January, 1947.

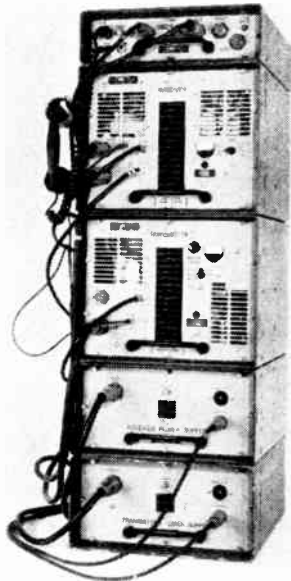


Fig. 2—Radio set AN/GRC-66 (Devel. model). Manufactured by Federal Telecommunication Labs., Nutley, N. J. Front $\frac{3}{4}$ view, showing complete assembled equipment.

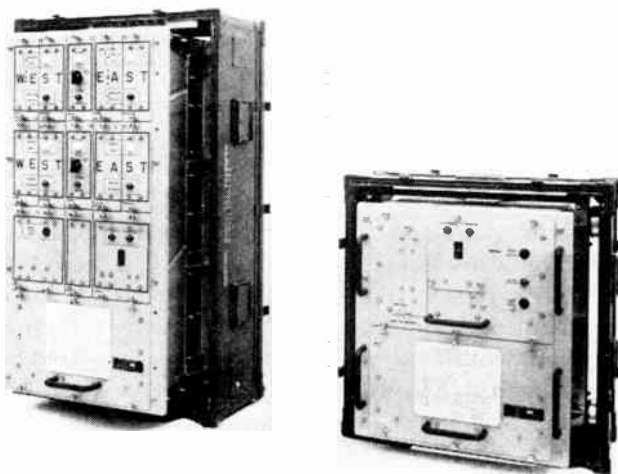


Fig. 3—Control Indicator Group AN/TRA-11 (Devel. model). Manufactured by Federal Telecommunication Labs., Nutley, N. J. Group view, showing equipments with covers removed.

the RF spectrum (2 to 30 mcs) is especially crowded. Consequently, most research and development devoted to this type of radio communications is directed toward more effective use of this limited natural resource. The exception to this is the advent of ionospheric-scatter and meteor-burst transmissions. Even these must contend with the high congestion experienced with radio-carrier frequencies, although these transmissions do show promise of being operable slightly above the Maximum Usable (Ionospheric Supported) Frequencies (MUF).

Also the first equipment embodiment of information theory has been passed to the U.S. Army Signal Research and Development Laboratory from the Lincoln Laboratory of M.I.T. Fig. 5 shows a view of this formidable attempt to trade bandwidth for signal-to-noise.

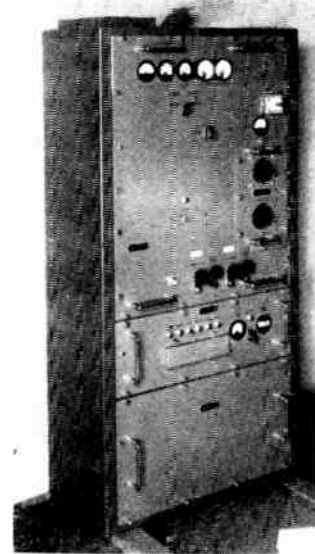


Fig. 4—Tropo-scatter equipment operating in the frequency range of 360–600 mc. Manufactured by National Co.

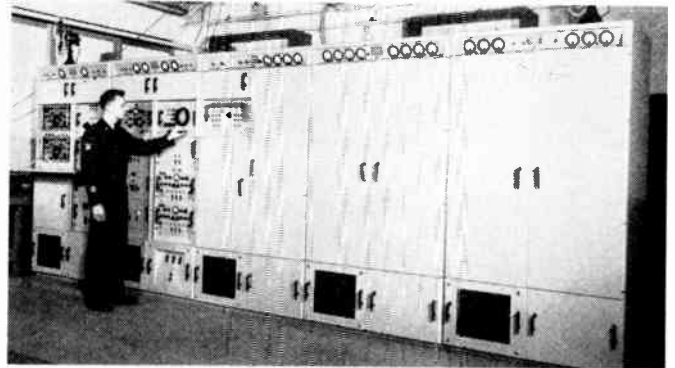


Fig. 5—Lincoln F9C-A Radio teletype system (Exper.). Manufactured by Electronic Defense Lab. Front view of receiving terminal equipment.

An extensive program of equipment development to place all these old and new techniques in the hands of troops is in various stages of progress. This has of necessity required equally extensive supporting technique investigations in the circuitry, propagation, and frequency stabilization fields. Also much effort has been applied to receiving and analyzing data transmitted from the artificial satellites as well as that reflected from the natural satellite of Earth. The implications of these devices are being closely considered for long-distance communication, and equipment development expedited wherever possible. Transmitters with RF power outputs as high as 300 kw and mobile receivers with frequency accuracy and stability at least 100 to 1 improved have recently been achieved for the HF or ionospheric range.

Multiplexing

Previous to World War II most, if not all, multiplexing of different intelligences was done on the basis of the frequency dimension. Time division was first used by the British during World War II in a military radio-

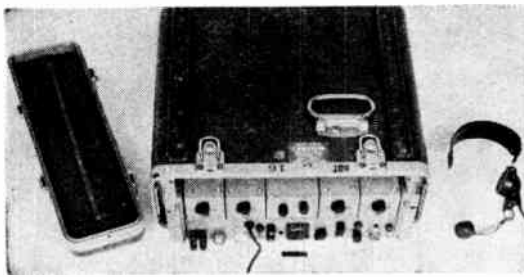


Fig. 6—Telephone terminal AN/TCC-24(XC-2), transistorized (Devel. Model). Manufactured by Lenkurt Electric Co. Group view, showing equipment in case, with cover removed, spare-parts container, and head set.

relay set.⁶ Their initial equipment used width modulation of a pulse to carry the intelligences. American versions of the British multiplex instituted the first use of position modulation of the pulse, followed closely just before the war ended by the first pulse-code-modulated equipments. These represented the first physical embodiments of what is now classical information theory. They even included auto-correlation detection for synchronization purposes.

Post-war developments have made great strides in miniaturization of multiplexing equipment, thanks again to wonderful progress in semi-conductor research. Fig. 6 illustrates a four-channel frequency division telephone-terminal multiplex, and Fig. 7 illustrates a seven- or 14-channel pulse-position-modulated unit. By the application of pulse-code-modulation voice, security is readily provided in similar terminal equipment for application to transmission means where the frequency bandwidths required are available.

Voice compression to save bandwidth is one of the favored goals in this area of voice multiplexing. It is in the terminal-multiplexing equipment that the additional equipment complexity penalty for inclusion of such a technique must be paid, but all such research efforts are being closely scrutinized and applied as they become feasible.

It is interesting to see the telemetering art beginning to approach the technical sophistication of an almost classical nature reached by the communication-multiplexing art a decade ago.⁷ Although the telemetry approach has been via the mechanical commutator, the electronic analogy with the fine art of multiplexing evolved by the communication profession is almost complete. It is now the communication art that is at times receiving in turn techniques from the highly progressive telemetry art that is working so closely with space-age objectives.

Switching

The Army in the field has in the past used manual telephone switchboards. While being relatively flexible, such a system is not compatible with the speed and ef-

⁷ V. D. Landon, "Theoretical analysis of various systems of multiplex transmissions," *R. C. A. Rev.*, vol. 9, pp. 287-351; June, 1948; pp. 433-482, September, 1948.

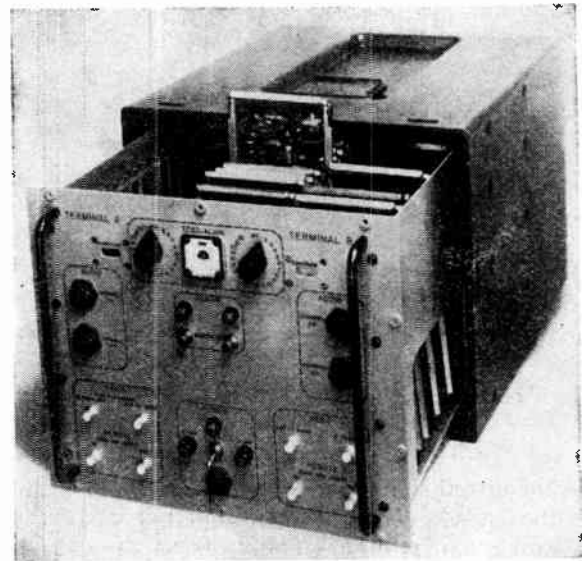


Fig. 7—Multiplexer set AN/TCC-27. Manufactured by RCA. Two-7-channel or one 14-channel terminal or reshaper repeater with D/I functions.

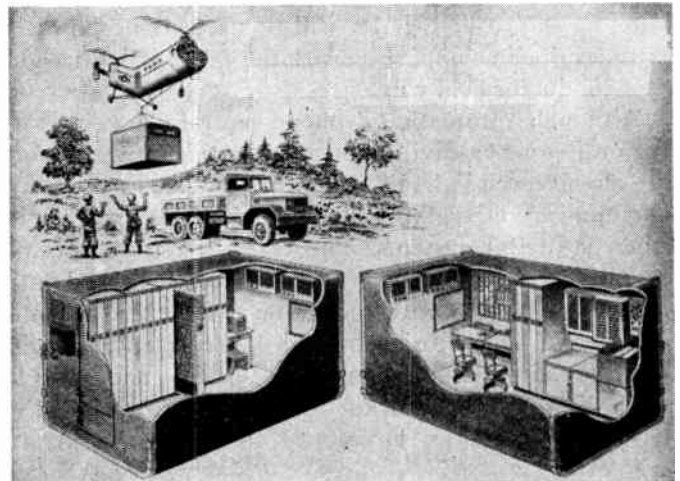


Fig. 8—Typical electronic automatic telephone central office. Approximately 300 lines.

iciency required to implement the communication networks serving tactical needs. As a result a whole new system is being evolved (see Fig. 8).

Flexibility is the prime requirement in new Army switching operations. To be able to tap in anywhere in the system, yet retained the individual subscriber's identity independent of location, is the type of requirement which makes the military switching system so much more difficult to achieve than its domestic counterpart.

Four-wire operation with its inherent circuit gain as well as freedom from feedback instability as compared to two-wire operation will be utilized. The usual multi-frequency dial system is planned along with all its complex of signaling from seizure to release of appropriate circuits. The signal system as a whole is interlocking. That is, when a signal is sent from one switchboard to another, the receiving switchboard must acknowledge



Fig. 9—Telephone set TA-341/PT. Manufactured by Stromberg Carlson

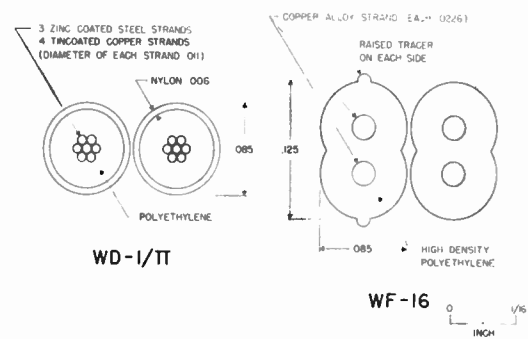


Fig. 10—Comparison of 2-conductor field wire with new 4-conductor field wire.

the transmitted signal correctly before the transmitting switchboard will send the next signal.

The telephone (Fig. 9) which is designed to terminate four wires is powered by dry-cell batteries housed in a fiberglass case which also houses the signal-detector circuits, amplifiers and necessary oscillator circuits for control and supervision. A new four-conductor field wire, as shown in Fig. 10, will be used. It is shown in cross section in comparison with the existing lightweight two-conductor field wire.

Although automatic telephone central offices operating on frequency-division basis and utilizing transistors are already well in hand, it appears that future military communication traffic will demand much more from time or digital systems. The transition from specific intelligence handling, such as voice, to generalized data requires a digital-switching system (Fig. 11). Various commercial laboratories are also laboring on the technical advancements necessary for such systems, and it might be appropriate for their information to outline the basic characteristics envisioned for the U. S. Army system:

1) All information will be handled in digital form throughout the system. Each channel will be represented by a binary signal transmitted in serial form at a standard speed of transmission.

2) Different speeds of transmission will be tolerated. A possible speed distribution might be

One in the teletype range,

One in the range of present data speeds,

One high speed for digitalized voice, facsimile, computer outputs, etc; bit rates will be standardized to obtain interoperation with existing and proposed data-transmission systems.

3) Switching will be automatic and automatic alternate-trunk routing will be employed which will equalize the traffic load on the system. A model of such a system is being designed and constructed to test concept feasibility and to obtain an idea of what can be accomplished with presently available components. The model will also be of assistance in determining the direction in

which further investigation should be conducted. It is possible that a digital-switching system such as that represented by the model will ultimately supersede the 4-wire system described earlier. Certainly there is the possibility that such a system will have a place in the ultimate universal switched-communications system.

Data Processing

Data processing is all that the name implies. The basic reason for the immense volume of paper work involved in seemingly every operation today is the transport and analyzing of information data of some type, or in some form. The only apparent method of effectively and efficiently coping with this appears to be by the use of machines. The most obvious are the computer-type machines which are relieving man of such routine, and even nonroutine effort. In military operations it is most important that full use be made of such machines to provide command and control of operations most effectively, efficiently, and in a minimum of time. Fig. 12 shows a tactical operation center (TOC) making use of existing computer equipment.

As mentioned under the above section on switching, the evolution of communication to more generalized data than the written or spoken word throws that type of intelligence into the general field of data. If we wish to generalize further, we process rather than merely communicate the data. In any case, the relative newness of such generalizations warrants a brief indication of the USASRD's approach to the data-processing field. Actually the Army has already very successfully applied computer techniques to logistics, and has an extensive computer development program well under way in connection with all aspects of combat operation.⁸ While the emphasis on the computer proper is to reduce size and complexity while increasing its capabilities, effort is also focused on the programming problem. The programming ease and program availability will no

⁸ W. F. Luebbert, "Combat computers," 1958 NATIONAL IRE CONVENTION RECORD, pt. 4, pp. 292-295.

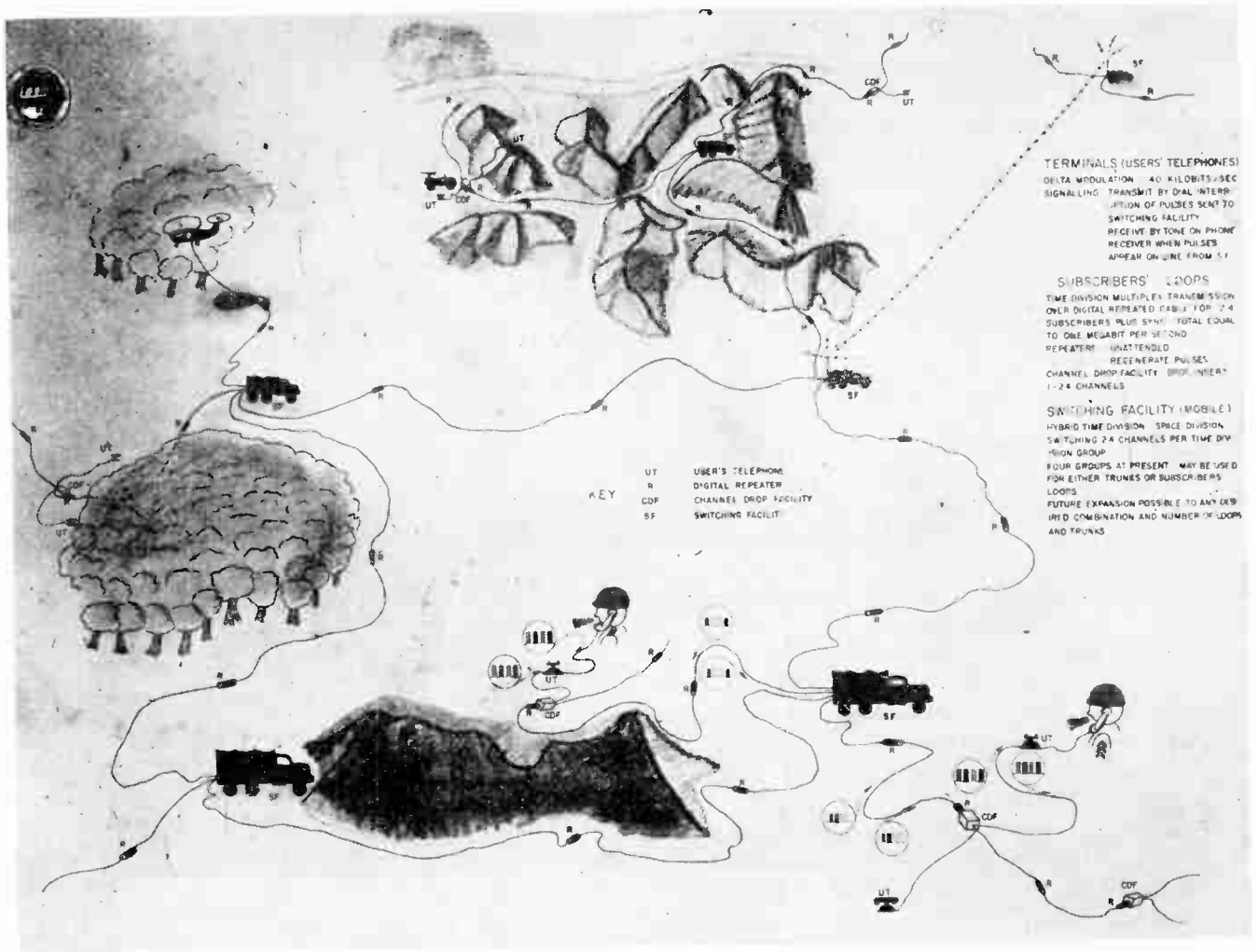


Fig. 11—Digital communication system ("Digicom").

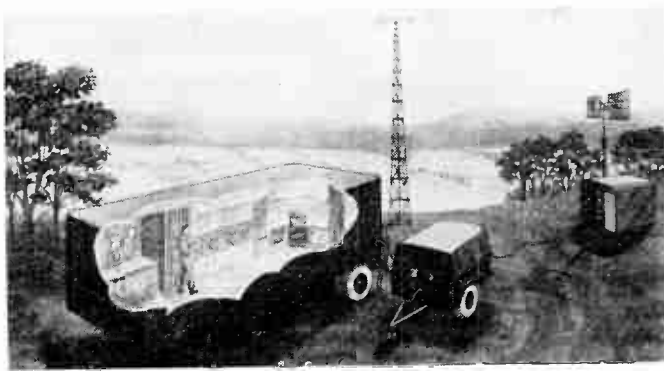


Fig. 12—Automatic data-processing facility (TOC mobile).

doubt be the determining factor as to whether or not computers may be used for most Army combat-data processing. Many subareas such as visual presentation are receiving much research and development attention with regard to the Army application objectives, which run into the hundreds.

The special-purpose computer versus the more universal computer is typical of the decisions required of

echelons involved in formulating operational and the accompanying development policy and equipment. There has to be a compromise made between the logistics and other economics of universality with its additional cost and complexity, and the economics of small, more mobile computers designed especially for one job. Combat-area radio sets are a good example of this. It is foreseeable that a family of such equipment mounted in vehicles, on a soldier's back, or even in his helmet or pocket, will incorporate computer elements to perform many electronic operations necessary to maintain secure communications under combat conditions in a highly-congested-frequency spectrum. In this case the computers will be unique to the operations, and miniaturized elements of semi-conductor dendrites together with closed, error-correction circuits automatically make use of muscular-generated potentials coupled from the soldier's sense organs. For the present, though, the program is to concentrate available effort on the immediate development problems concerned with adapting known universal-type computer techniques to many Army data-processing problems.

Systems Concepts

Some concept of the systems problems in general can be illustrated by an analogy. Take the commercial world-wide communication stations; add the police-radio nets, both municipal and state-police nets; include all private-transportation radio systems, both vehicular and pipe-line, whether mobile or fixed-microwave radio relays; and to confuse further but not to complete the picture, include untold miles of overhead and underground wire and cable lines. This, together with all exchanges and outside plant, is to be arranged into a system that can be installed in minutes, or hours at most. After a few hours, or less, of carrying its maximum capacity in traffic, it is to be dismantled, for the most part, moved a few hundred miles and re-installed for a repetition of the same operation. If a terrific hurricane is also present during the operation, the scene is almost analogous to that expected to be presented by the U. S. Army communications under war conditions.

The USASRDRL has recently defined the term "technical concept"⁹ This definition states in part that a technical concept is considered to mean an undeveloped idea or a new technical method of employing equipment which has potential capability of providing the required signal support for an operational objective.

This definition in one sense is limited, yet in another sense relatively broad. It denotes equipment employment, yet calls for ideas not yet developed. Here is an area of research and development very close to the defense of your home and mine. Brilliance is required, yet because most of our more brilliant, disciplined minds gravitate to relatively narrow academic pursuits, it is questionable as to whether new systems concepts are receiving the concerted attention given to the more subtle details that require deep knowledge of science. The rise of operational research is the exception that proves this statement.

Data-processing systems are typical of this need for new concepts. The mere routing sequence which uses the term "buffer" in an almost completely new descriptive manner, as compared to the vacuum tube buffer application familiar to radio engineers, is indicative of the new systems evolved and being evolved based on new concepts.¹⁰ This could bring up the whole question of dial, or equivalent signaling, so lightly mentioned previously. In spite of the wonderful achievements made commercially, many new concepts are needed to meet the heterogeneous requirements in this military-system area, if all subscribers are to have access to a closed network where anyone must be able to communicate almost any form of intelligence to anyone also with the least effort and in the shortest time.

This is not to indicate that the U. S. Army does not

have this situation under control. Rather, it is indicative that engineers and scientists are more aware of the problems unsolved than those solved; which is as it should be. The Army must currently contend with the need for new equipment concepts, though, through re-orientation of their operational concepts. That is one of the major reasons for the various "experimental armies" which are being tested. Whole catalogs of new equipments in each of the communication areas discussed above, could be and are listed in equipment programs which are destined to outfit armies of 1960, 1970, 1980 Even so, the technical state of the military communication art is limited by professional technical desire to vouchsafe to the design of new system concepts just as surely as it is limited by our ability to advance to new frontiers of basic scientific knowledge.

In this system area the USASRDRL is working very closely with the United States Army Electronic Proving Ground at Fort Huachuca, Ariz. The number and extent of new system concepts has required the use of the 70,000 acres of flat and mountainous terrain for test purposes in such an area, because it is relatively free of electromagnetic-radiation interference. The size of the test area shows the stress which the U. S. Army has placed on system tests of new signal concepts.

RADIO INTERFERENCE ANALYSIS AND CONTROL

The USASRDRL has achieved wide recognition for its radio-interference-control work dating back to the beginning of World War II. The mechanization of the U. S. Army initially required attention primarily to vehicular radiation suppression. Subsequently a continuing program of interference analysis and control has aided all services, both military and otherwise, in solving their electromagnetic-transmission-interference problems. Standards specifying spurious emission and interference criteria have been formulated and widely disseminated by specialists working in the Communications Department, U. S. Army Signal Research and Development Laboratory, although their interests necessarily encompass all areas of equipments which involve electromagnetic radiations. A large segment of their current research, for instance, is devoted to radar-interference problems.

One of this Division's primary assignments is to develop any equipment that is necessary for application of their standards. In this connection, Figs. 13 and 14 illustrate typical interference-measuring sets provided by them for different frequency ranges. The development of instrumentation for ever higher frequencies is a continuing activity. In the research category they have a continuing program for obtaining new and improved suppression components. More recently, major problems have been associated with mutual interference among electronic equipment, caused primarily by their ever-increasing density in power. Current effort is being placed on lossy-type filters and the use of nonlinear devices as suppression components.

⁹ "Formulation of New Technical Concepts," Headquarters, Signal, Memo No. 82, October 1, 1958.

¹⁰ R. E. Montijo, Jr., "An Automatic Communications Switching System," *Signal*, pp. 36-44; September, 1958.



Fig. 13—Radio interference-measuring set AN/PRM; (mock-up). (Lab. Design.-Constr.) Frequency range: 75-500 mc. Overall $\frac{3}{4}$ view, showing personnel demonstrating operation of model.

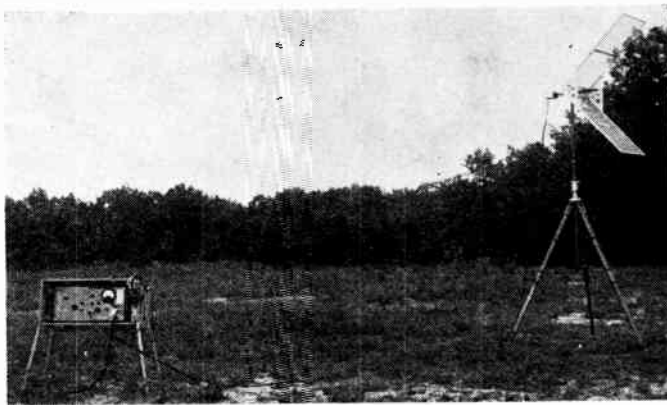


Fig. 14—Radio interference-measuring set AN/URM-85 (Service-test model). Manufactured by Empire Devices. Frequency range: 0.15-1000 mc. Overall view, showing equipment assembled in operating position for field-strength measurement in frequency range 400-1000 mc.

MISCELLANEOUS ASPECTS

There are many miscellaneous aspects of communication research and development required of USASRD by the U. S. Army. While not all are earth shaking, some literally approach this description. For instance, the auger shown in Fig. 15 is a necessary adjunct to the com-



Fig. 15—Auger, earth, truck-mounted LC-245(XC-2)/M (Devel. model). Installed in cargo space of truck, $2\frac{1}{2}$ ton, 6X6, cargo M34. Manufactured by Highway Trailer Co. Roadside rear $\frac{3}{4}$ view, showing rack bar and guide tube in raised position, ready for digging.

munication officers' tool supply. Also, as shown in Fig. 16, the communication engineer must provide a cabinet, sometimes complete with air conditioning, to house the electronic heat sources.

Then there are the myriad of test items needed to keep a system in order. Fig. 17 typifies such items.

On the acoustical side of the ledger there are such items as those shown in Figs. 18 and 19.

The transmitter of Fig. 20 utilizes acoustical interference methods for reducing the masking effect of ambient noise surrounding a listener in Army vehicles, tanks, and aircraft.

While still in the acoustical vein, the equipment shown in Fig. 21 is worthy of note because it reproduces speech and other signal frequencies without the use of conventional high-power amplifiers. The equipment utilizes magnetic particle clutches driven by electric motors to operate at AF and "friction-drive" a loud-speaker diaphragm.

A temptation will be resisted at this point to include a multitudinous group of figures illustrating everything from complex multiple-call reperforators to progress on 3000-wpm all-electronic page printers. The result of viewing such illustrations would probably be the same type of feeling one experiences in walking through the displays at the annual IRE NATIONAL CONVENTION show in New York City, N. Y.

SPACE AGE IMPLICATIONS

Much has been written and proclaimed regarding the effects that satellites and other space events will have on communications, relative to the Earth as well as to space in general.¹¹ There is no object in listing the various, usually grandiose schemes here; nevertheless, the impact on communications of these elevated platforms portends to be considerable. Overriding priorities as well

¹¹ R. W. Johnson, J. B. Medavis, *et al.*, "The Army in the space age," *Army Information Digest*, vol. 13, pp. 1-65; October, 1953.

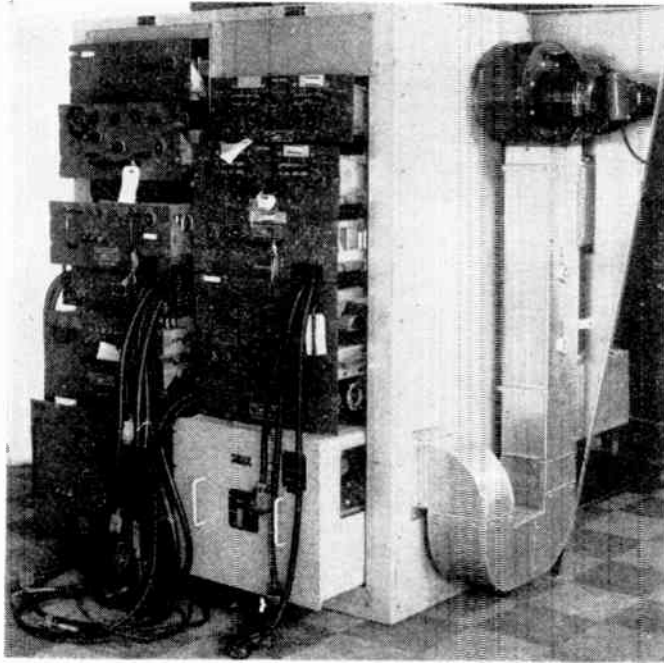


Fig. 16—Cabinet, electrical equipment CY-2295/U (Engineering test model). Manufactured by Craig Systems Inc. Front $\frac{3}{4}$ view, showing installation of complete telephone terminal AN/TCC-7 (less one telephone) modem TA-219/U and combination heating and air-conditioning unit in cabinets, electrical equipment CY-2295/U and external blower and ductwork.

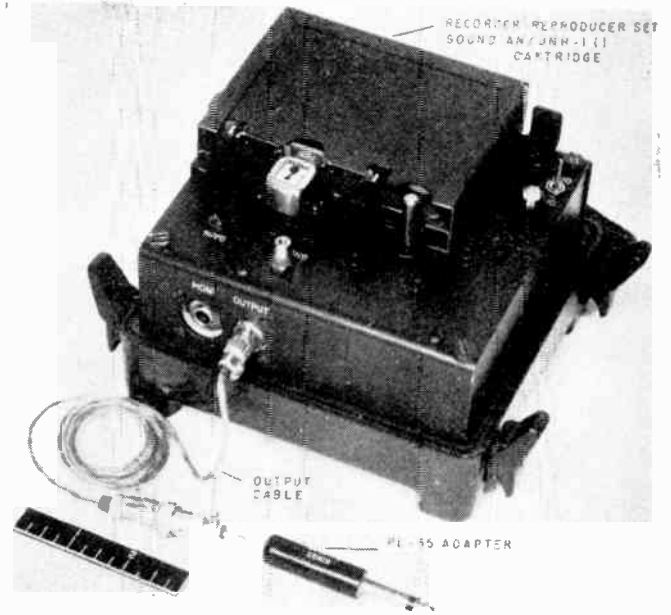


Fig. 19—Reproducer, sound RP-116/U (Service test Model). Used with recorder-reproducer set, sound AN/UHH-1. Manufactured by Mohawk Business Machines Corp. Overall front view, cover removed, showing complete unit, ready for operation.

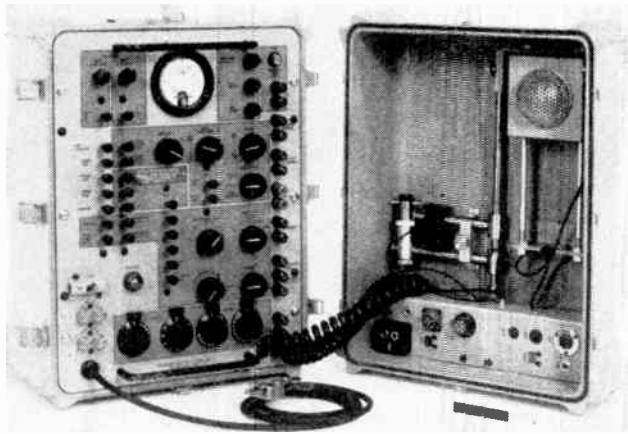


Fig. 17—Telephone test set TS-716/U (Devel. model). Manufactured by RCA. Front interior view. Operating panel, with cover off, showing acoustical equipment.



Fig. 20—Handset H-138/GR (Devel. model). Manufactured by Electro-Voice, Inc. Used with new FM radio sets. Overall $\frac{3}{4}$ view, showing complete assembled unit.

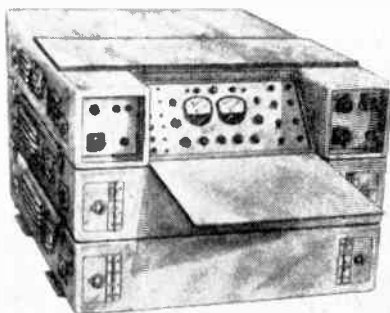


Fig. 18—Public-address set AN/UH-3. Manufactured by Stromberg Carlson.

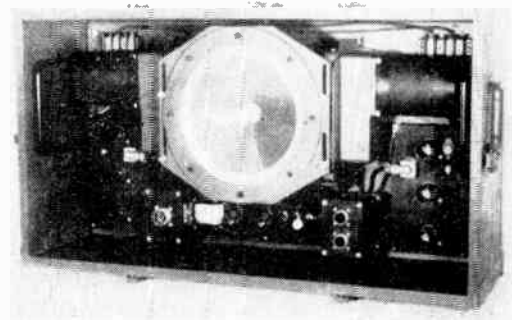


Fig. 21—Public address set AN/VIQ-1 (XC-2). Manufactured by Raymond Engineering Co.

as high-security classifications, in some instances, have been attached to communication-satellite programs assigned to the Communications Department of USASRDL. The SCORE-ATLAS communication experiment is indicative of such programs.

Immediate developments as well as very-long-term research programs have been instituted. Detailed proposals have been received from most of the prominent communication and electronic firms of the U. S. A., and equally detailed systems analyses have been made, even previous to the responsibility-assignment decision regarding Project Vanguard. Indeed, it is these 1955 systems analyses that resulted in the U. S. Army Signal Corps' receiving the assignment to establish the worldwide monitoring and tracking network initially set up in connection with Vanguard, which is now also seeing yeoman service for the ensuing array of Sputnik and Explorer satellites.

Even in today's pace of electronic and communications progress, it is relatively easy for a Government laboratory to keep abreast. In fact, it could not be avoided. Every new idea or system is immediately brought to the Government's attention. Usually a major percentage of the time of Government technical personnel is spent examining such items, whether forwarded by the lone inventor or by the industrial colossus. Satellites for communication are no exception to this procedure, so USASRDL has had the benefit of many fine proposals.

Satellites excite the imagination of all communicators, especially radio communicators. Here is the Indian rope trick without the rope. Yet, when all the proposals are examined in detail it is found that the one thing still lacking is an original idea regarding the satellite for communication purposes. This is not meant to detract from the utility of satellites for communications. Rather, it is meant that, as satellites are established in various orbits, both roving and fixed, relative to the Earth, they essentially provide an elevated platform. So the communicator places his equipment on the platform and takes advantage of the height-gain in the same manner he would use a mountain top, or the Empire State Building, for that matter. Of course, it is not that simple. Environmental and reliability factors enter the picture, and here is a job for the research engineer as well as the component and circuit designer. With the vacuum of space, a different concept of components may be possible, but such application specialties are not confined to communications equipment. Also, if the satellite is to be an amplifier-repeater, primary power raises a need for a whole series of such devices. These are described in an accompanying paper.¹²

In addition to the amplifying, or so-called active satellite, is the reflective or so-called passive satellite. Many surveys have been made by competent groups at communication research and development laboratories

to determine the system feasibility of such repeaters. Both have their advantages and disadvantages. All are intended for long-distance communications between various points on earth, although there are also long-term programs which envision broadcast-type stations that could encompass most of the earth through a system of satellites fixed in positions relative to the Earth in their rotation.

In any case, the communication engineers must await the progress of the vehicle or rocket programs. The system actually used, particularly in the immediate future, will depend on the orbits that are achievable. As a result most readily-achievable orbits have been examined relative to their radio line-of-sight coverage of the Earth.

This in turn brings to the fore consideration of two types of message conveyance via satellites. One is the straight-through relay; the other is the storage or postal messenger-carrier type of relay. An orbit which might be best for one is not necessarily the best for the other type of transmission. The straight-through repeater, whether active or passive, depends upon radio line-of-sight-transmission paths simultaneously apparent from both terminals. The postal-delivery type in which the messages are stored over the sending terminal, and released over the receiving terminal, require radio line of sight from only one terminal at a time. In the military versions, of course, it would also be desirable to exclude potential enemies from access to the satellites, either through use of codes and/or orbits.

The next parameter the systems engineer must choose is the radio-carrier frequency to be used. This depends upon channel availability from an allocation viewpoint, and the transmission characteristics from a propagation viewpoint. The former decision is more in the realm of politics, but the latter requires technical knowledge. Much has been learned from monitoring electromagnetic radiations from Sputniks, Explorers, and a Vanguard. Nevertheless, much more needs to be learned.

In examining such recordings, it is relatively apparent that frequencies most unaffected by the ionosphere should be used. Diversity in its fullest aspects should be incorporated, and last but not least, a modulation technique best able to cope with doppler shifts and provide best system gain should be adopted.

It is not the intention here to dwell on conventional system-type engineering details, but the propagation aspects of earth-orbiting satellites are worthy of brief consideration. Most intriguing have been the anomalies experienced with the 20-mc emanations of the Sputniks. Such an ionospheric-affected frequency would be avoided by a communication engineer designing for reliability. Yet even ionospheric-bounded frequencies have their place in satellite systems. The high attenuation that keeps them from radiating out from the Earth too readily also keeps them from coming in toward the Earth. For inter-satellite communication at distances

¹² W. L. Doxey, S. F. Danko, and J. P. McNaul, "Parts and techniques for the new electronics," this issue.

outside the ionosphere concentrations, such frequencies may be the only ones relatively uncongested, if the trend to high-powered television broadcast and radar-type operations continues to increase the occupancy of non-ionospheric-affected frequencies.

This brings up another point in connection with the use of satellites for military communications. Once a satellite-radio repeater is in orbit it is anyone's property. Not only military enemies, but even our own radio amateurs have every legal right, if not a moral right, to use it. Codes can be inserted, but the longer in time the satellite remains in orbit, the more time there will be to break the code. Perhaps the national craze in the not too distant future will be to see who will be the first to discover the satellite's code. With computers this will be a relatively simple task, at least the task of deciphering the code used to allow interrogation of a storage-type repeater.

Much consideration has already been given at USASRDRL to ground-communication equipment required in connection with satellites, whether active or passive. Such ground equipment varies from the simple communication receiver and equally simple transmitter with only a few watts of RF power plus equally simple dipole antennas, to the most complex configuration of scatter-radar type of systems with their immense astronomical-type paraboloidal antennas. In contrast with the latter, the 20-mc Sputnik signals were easily received on amateur-type receivers with merely long-wire antennas. A passive-reflector satellite, though, if used to reflect microwave frequencies, would require as much as 100 kw of transmitter-output power if video frequencies were to be communicated, even though low-noise Maser-equipped receivers were used. Paraboloidal or lens ground antennas of the order of 100 to 200 feet in diameter would be required if a passive reflector-type satellite 100 to 200 feet in diameter were available at satellite heights, comparable with those of the initial Explorers.¹³ Also there are advocates of metalized balloons and other types of reflectors as well as various chemicals to be strewn about the Earth's outer-space areas.¹⁴

The most critical components other than the power sources are recorders. Being a mechanical component, the recorder is likely to fail in a relatively short time compared to electronic elements. The recorder is only likely to last a few hundred hours at best, as compared to the many thousands of hours expected of electronic components. On this basis it would, therefore, appear that no recorder should be placed in a satellite that is intended to last beyond a few years. A storage-type satellite, based upon present techniques and expected-duty cycle of present orbits, might last five years, but trouble-free operation over this period is relatively im-

probable with existing recorder techniques.

This brings up a so-called "burst" type of satellite repeater which could be either active or passive. This system, similar to the meteor-burst systems recently evolved, would keep the recording or other storage devices on the ground.¹⁵ Reliability of these components would not then be a critical factor in the overall system.

In any case, as attested by project SCORE (Satellite Communications via Orbiting Relay Experiment), the U. S. Army will take advantage of space vehicles, particularly as it falls within the responsibility of the U. S. Army Signal Corps to provide vital communications.

ACKNOWLEDGMENTS

There is so much technical work proceeding in the Communications Department of the U. S. Army Signal Research and Development Laboratory that is not included in this article for various valid reasons, that it appears appropriate and just to apologize to the people working in those categories neglected here because of space limitations. For instance, communication research has been bypassed in its entirety, although it is the forte of the writer who has enjoyed the privilege of delving into other peoples' work in order to compile this material. Only a small fraction of the major equipment developments is described; just sufficient are included to fulfill the objective of this type of paper, which is somewhat akin to a "get acquainted" visit to a portion of U. S. Army Signal Research and Development Laboratory. Details of the SCORE success are not included because many other articles have been and are being published dealing with that subject alone.

Illustration captions have had mention of commercial or other laboratories left on them purposely to acknowledge the close partnership-type cooperation mentioned in an earlier section. Furthermore, while the modern rush into complex, technical advances leaves very little room for nostalgia, nevertheless the writer was very much aware of the fact that the current work described together with the engineers and scientists involved, was but a small segment in a long historical continuance having tradition of a type very seldom found in our multitude of electronic laboratories, which have recently sprung up throughout the nation. The U. S. Army Signal Research and Development Laboratory has been singularly fortunate to have lived through the earliest days of electrical and smoke signals, not to neglect pigeon communications. The list of great and small names of the past as well as of the present, closely associated with the USASRDRL or its predecessor organizations, is very impressive. Vail, Squier, Armstrong . . . typify the better known of the past; Everitt, Colton, Goubau . . . of the present; and of the future . . . ?

¹³ J. R. Pierce, "Orbital radio relays," *Jet Propulsion*, vol. 25, pp. 153-157; April, 1955.

¹⁴ American letter, "An artificial ionosphere," *Brit. Commun. Electronics*, vol. 5, pp. 117ff; February, 1958.

¹⁵ R. Eshleman, "Meteors and Radio Propagation," Radio Propagation Lab., Stanford University, Stanford, Calif., Tech. Rep. No. 44, February 1, 1955.

The Engineering of Communication Systems for Low Radio Frequencies*

J. S. BELROSE†, MEMBER, IRE, W. L. HATTON†, MEMBER, IRE, C. A. MCKERROW†,
AND R. S. THAIN‡, SENIOR MEMBER, IRE

Summary—The low radio frequencies are of considerable importance in specific radio communication applications. However, little improvement has been made in recent years in methods of engineering low-frequency systems. In this paper the factors that influence the design of a communication system at low radio frequencies are discussed, and a description is given of some experimental work designed to obtain information concerning optimum values of some of the design parameters.

INTRODUCTION

THE low-frequency band occupies only a very small part of the radio frequency spectrum. Nevertheless, this small band of frequencies has been used for communication since the advent of radio. Although the historical significance of the low-frequency band is well known, the great extent of present usage of this band is not generally appreciated.

In the history of radio communication, the low-frequency transmitting installations of the pioneering era were characterized by their large physical size and by their high construction and maintenance costs. Moreover, then, as now, signal reception at low frequencies was seriously hampered by atmospheric noise, particularly at low geographical latitudes. In addition the increasing demand for radio communications soon resulted in serious congestion of transmissions within the then available spectrum. Thus, with the development of the high-frequency "beamed" systems about 1924, it was found that the high frequencies offered an attractive solution to many long-distance communication problems, and the use of the low frequencies became considered as outmoded. Nevertheless, propagation factors peculiar to the low-frequency band have resulted in their continued use for radio communication. In particular, the reception of low-frequency waves is not adversely affected during periods of ionospheric disturbance when communication at the high frequencies is disrupted. Because of this, there is a particular interest in the application of low frequencies at high geomagnetic latitudes. Since the initial use of radio for communication purposes, considerable advances have been made in improving both the efficiency and the quality of communication by the application of communication theory, radio wave propagation studies, and new techniques. Unfortunately, very little effort has been made

to apply this new knowledge to systems design in the low-frequency band.

Because of the large extent of defense interest and investment in low-frequency installations in Canada, a group was formed at the Radio Physics Laboratory in 1950 to study low-frequency radio wave propagation and applied communication techniques. The objectives of this group were necessarily broad. The primary aim was to assist the defense services in improving the efficiency of the existing circuits. In directing efforts toward this objective it was soon found that the tasks involved were numerous and diversified. A review of the literature indicated that considerable research effort was required, and it became necessary to select the most promising directions for the rather limited research facilities available. Attention was focused on three principal factors which affect the systems design: the propagation medium (including the behavior of the ground wave, ground conductivities, the amplitude and fluctuation of the sky wave as a function of transmission distance, frequency, time of day, season, and the amplitude and characteristic nature of atmospheric noise as a function of frequency); antenna systems (with special emphasis on efficiency and economy); and modulation systems (bearing in mind the conflicting demands of reliability and power economy). Experimental systems were built in accordance with the various conclusions reached, and these systems were tested in operation.

It is the purpose of this paper to present a rather detailed discussion of the factors indicated, to outline the results of the experimental tests, and finally, to illustrate the design phase by a sample calculation.

Before commencing the discussion of the design factors, let us consider certain broad observations which affect the contents of this paper. First, systems designed for radio communication at frequencies below 200 kc may be separated into two categories. One of these, designed to provide world-wide communication coverage, is limited to frequencies below about 30 kc. The second, intended to provide communication over relatively shorter distances, *i.e.*, up to about 2000 km, is more efficient from cost considerations if it is designed for operation between about 80 and 200 kc. The design problems are somewhat different for the two categories and, because the Canadian interest was confined to the shorter transmission distances, the discussion here is limited almost entirely to the higher frequency range, 80 to 200 kc. In addition, the chief Canadian requirement for low-frequency systems is that they provide radio com-

* Original manuscript received by the IRE, November 17, 1958; revised manuscript received, February 16, 1959.

† Defence Res. Telecommun. Est., Ottawa, Ont., Can.

‡ Directorate of plans, Defence Res. Board Headquarters, Ottawa, Ont., Can.

munication "back-up" support for high-frequency circuits, particularly when the latter are disrupted by ionospheric disturbances. These requirements concern conditions common to northern Canada. Under the circumstances, the discussion is aimed primarily at northern conditions with the chief design effort being directed toward the provision of reliable communication, while exercising strict economy of cost and of power. In accordance with this design objective, the discussion of modulation is confined to single-channel systems and signaling rates of 60 wpm maximum.

THE PROPAGATION OF LOW-FREQUENCY RADIO WAVES

Propagation over Short and Long Distances

Before discussing the problems encountered in the engineering of low-frequency radio systems, it is necessary to review briefly our knowledge of the propagation of low-frequency radio waves. This discussion is limited to considerations of the vertical component of the electric field only.

The signal received at a distance d measured along the earth's surface from the transmitter consists of a contribution due to the ground wave and contributions due to skywave components. In the simplified case of a received signal that consists of a ground wave and a once-reflected skywave, the signal received by a loop antenna located at the surface of the earth is proportional to the total field intensity E_N , where

$$E_N = E_0 + E_s \quad (1)$$

E_0 is the amplitude of the ground wave, and E_s is the effective amplitude of the skywave.

Throughout the daytime, the signal amplitude, E_N , describes some form of quasi-cyclic variation as the phase of the skywave vector varies relative to the constant phase of the ground wave. The variation of amplitude as a function of time depends on the reflection heights and reflection coefficients of the skywave, and on the amplitude and phase of the ground wave.

The field strength of the ground-wave contribution at the receiver may be expressed by

$$E_0 = \frac{mE_u}{d} \quad (2)$$

where E_u is the inverse-distance attenuated field intensity at the earth's surface, at unit distance from the radiator, and m is the ground-wave attenuation factor, which in turn is a function of frequency, polarization, distance, earth curvature, and the effective conductivity and dielectric constant of the intervening terrain.

The variation of ground-wave field intensity with distance has been studied theoretically by many workers, including van der Pol, Bremmer, Eckersley, and Millington. Norton¹ has presented a solution in a form

¹ K. A. Norton, "The calculation of ground-wave field intensity over a finitely conducting spherical earth," *PROC. IRE*, vol. 26, pp. 623-639; December, 1941.

that permits simple numerical calculations of the ground-wave field intensities for propagation over a homogeneous and smooth curved earth. Propagation curves based on the method of Norton and the more rigorous analysis of van der Pol and Bremmer have been prepared by Thain,² and Wait and Howe.³ Agreement is still lacking, however, concerning the extent of tropospheric refraction effects at low radio frequencies. The referenced propagation curves were prepared using an effective earth radius equal to 4/3 times the actual radius, on the assumption that air refraction effects are applicable at the frequencies concerned. The CCIR, on the other hand, has recommended that no account be taken of tropospheric effects at frequencies below 10 mc, and has provided propagation curves based on this recommendation.⁴ The net result is that the CCIR curves predict a greater attenuation of ground-wave field intensity as a function of distance from the transmitter, for a given value of earth conductivity, than is predicted assuming refraction effects. The difference is especially marked in the far-field case where, for example, for radio waves at 100 kc propagated over ground of conductivity 10^{-3} mhos/m, the field intensities predicted by the CCIR and by the Norton method differ by a factor of about 1.75 at 1500 km from the transmitter.

The empirical method of Millington⁵ is the easiest and best way to take into account the effect of a mixed path on the ground-wave field intensity. The method can also be extended to estimate the resultant phase of the ground wave propagated over an inhomogeneous smooth earth.

The effective amplitude of the skywave at the receiver, E_s , depends on the characteristics of the transmitting and receiving antennas, and the reflection coefficients of the ionosphere and of the earth. If the transmitting antenna is a short vertical radiator situated on the surface of a perfectly conducting earth, the amplitude of the downcoming skywave, E_1 , can be represented as

$$E_1 = \frac{E_u}{L} \sin i_\theta \|R\| D \quad (3)$$

where

- E_u = the inverse-distance attenuated field at unit distance from the transmitting antenna,
- i_θ = the angle of incidence of the skywave with the ground,

² R. S. Thain, "Ground-Wave Field Intensity Curves, 20 KC to 540 KC," Defence Res. Telecommun. Establ., Ottawa, Ont., Can., DRTE/RPL Rep. No. B3; April, 1952.

³ J. R. Wait and H. H. Howe, "Amplitude and Phase Curves for Ground-Wave Propagation in the Band 200 C/S to 50 KC," NBS Circular 574; 1956.

⁴ Recommendation No. 52, "Ground-wave propagation curves below 10 mc/s," Documents of the VIII Planetary Assembly, CCIR, Warsaw, Poland, vol. 1, pp. 47-59; 1956.

⁵ G. Millington, "Ground-wave propagation over an inhomogeneous smooth earth, (parts I and II)," *J. IEE*, vol. 96, pt. 3, pp. 53-64; January, 1949, and vol. 97, pt. 3, pp. 209-217; July, 1950.

- L = the length of the skywave path,
 ${}_{\parallel}R_{\parallel}$ = the ionospheric reflection coefficient,⁶
 D = a factor introduced by the focusing effect of the spherical ionospheric reflecting surface.

The variation with transmission distance of the angle of incidence of the skywave, and the focusing factor, for an assumed reflection height of 80 km, are shown in Fig. 1. One of the limitations in computing field strengths by geometrical optics, so-called hop theory, occurs when the skywave leaves and returns to the earth tangentially (this will be discussed in more detail later). The geometrically optically computed focusing factor becomes infinite at this limiting range. Wait⁷ has recently examined this difficulty, and has introduced a correction factor, which is a function of frequency, to remove the infinity. His values of the convergence factor for a frequency of 100 kc are shown in Fig. 1 as the solid line.

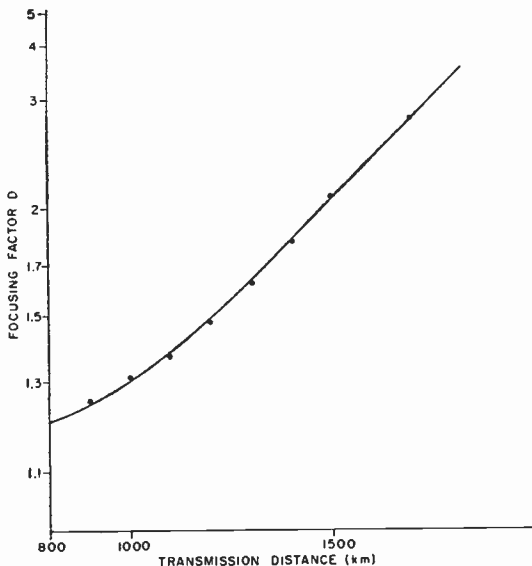


Fig. 1—Angle of incidence at the ionosphere and the ground, and focusing or convergence factor, for a smooth spherical ionosphere reflecting surface located at a height of 80 km. The dashed curve for D shows the convergence factor as derived from geometrical optics, and the solid curve shows the wave-theoretical convergence factor for 100 kc calculated by Wait.⁷

If the receiving antenna is a small loop located on the surface of a flat perfectly conducting earth, the effective amplitude of the skywave, E_s , can be expressed as

$$E_s = 2E_1 \quad (4)$$

$$= 2 \frac{E_0}{L} \sin i_0 {}_{\parallel}R_{\parallel} D \quad (5)$$

The factor 2 in (4) and (5) arises from consideration of

the incident wave and the wave reflected from the surface of the earth at the receiving antenna.

Observations made by Belrose⁸ by recording the field strength of CW signals from European transmitters in the 70- to 272-kc band, at distances between 600 and 1500 km, revealed quasi-cyclic amplitude changes. A study of these features enabled the determination of ionospheric reflection heights and reflection coefficients. The observed fading patterns could be described as though reflection occurred, without a change of phase, from a sharply bounded, mirror-like surface. The reflection height was about 72 km at midday in summer, 75 km at midday in winter, and about 90 to 95 km at night. The apparent reflection height was found to vary in a strikingly regular and predictable manner with the zenith angle of the sun. The seasonal variation of the ionospheric reflection coefficients, and their variation with frequency, have also been described. The variation with frequency and transmission distance may be described in terms of an effective frequency, $f \cos i$, where f is the frequency of the radio wave incident on the ionosphere at an angle i . It has been shown that the logarithm of the reflection coefficient is approximately a linear function of the logarithm of the effective frequency. Fig. 2 illustrates the variation of reflection coefficient with effective frequency for the months of November, 1954, and March and June, 1955. In order to derive these ionospheric reflection coefficients from the experimental data, it was necessary to take into account the effect of ionospheric focusing, and the effect of the finitely conducting earth on the polar diagrams of the transmitting and receiving antennas. This latter effect will be discussed in the next section. The reflection coefficients of Fig. 2 represent reflection as if from a flat ionospheric surface and, if these values are used in conjunction with (5), the effective amplitude of the skywave deduced is that which would be measured by a loop antenna, if both transmitting and receiving antennas were located on the surface of a perfectly conducting earth.

The seasonal variation of the reflection coefficients for 100-kc radio waves over a path of 1475 km (effective frequency, $f \cos i$, approximately 17 kc), and for 191 kc over a path of 1173 km (effective frequency approximately 34 kc), are shown in Fig. 3.

The data for Figs. 2 and 3, obtained for latitudes 40–55° N, are insufficient to indicate any difference between east-west and north-south transmission for propagation at these distances. Very little work has been done in the Arctic regions. Bickel⁹ has studied 135.6-kc transmission in the Alaskan region (60° N) over a 1550-km path. The reflection coefficients he derived are included in Figs. 2 and 3, and appear to be in reasonable

⁶ The notation is the usual, *i.e.*, the first subscript denotes that the electric field of the incident wave is parallel to the plane of incidence, and the second subscript refers in the same way to the electric field in the reflected wave.

⁷ J. R. Wait, "Diffractive Corrections to the Geometrical Optics of Low Frequency Propagation," Natl. Bur. of Standards, Boulder, Colo., Rep. No. 5572; 1958.

⁸ J. S. Belrose, "Some Investigations of the Lowest Ionosphere," Ph.D. dissertation, University of Cambridge, Eng.; 1957. To be published in *J. IEE*.

⁹ J. E. Bickel, "A method for obtaining LF oblique-incidence reflection coefficients and its application to 135.6 kc data in the Alaskan area," *J. Geophys. Res.*, vol. 62, pp. 373–381; September, 1957.

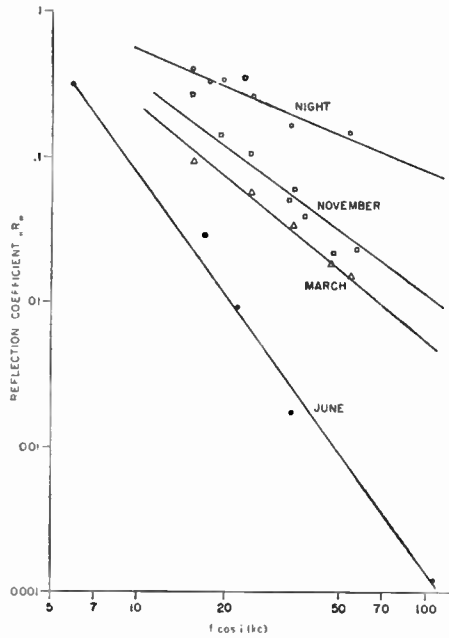


Fig. 2—The variation of ionosphere reflection coefficient for low-frequency waves at oblique angles of incidence. The plotted points have been calculated mainly from data derived from 1954–1955 measurements, and represent the estimated median values. The ionosphere convergence factor and the antenna factors have been removed.

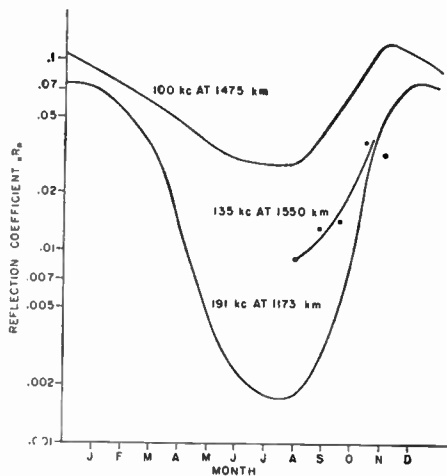


Fig. 3—The seasonal variation of the ionospheric reflection coefficients for low-frequency waves at oblique angles of incidence. The 100-kc data refer to observations in 1951; the 135- and 191-kc data, to observations in 1954–1955.

agreement with those for lower latitudes. The latitude effect, for summer at least, appears to be small.

The Effect of a Finitely Conducting Earth on the Radiation Pattern of Antennas

In the measurement of the amplitude of long waves reflected from the ionosphere, it is often assumed that the earth is a perfect reflector. While such an approximation is valid for the lowest frequencies and for low frequencies reflected from the ionosphere at steep incidence, the effect of a finitely conducting earth can con-

siderably modify the radiation pattern of an antenna system and, accordingly, the apparent reflection coefficient for propagation at very oblique angles of incidence, *i.e.*, radiation at very small angles above the horizon.

The vertical radiation pattern of a short vertical dipole located on the surface of a flat earth of finite conductivity is shown in Fig. 4.¹⁰ The parameter *n* is given by

$$n = \frac{18 \times 10^3 \sigma}{f_{mc} \epsilon_r} \tag{6}$$

where σ is the earth's conductivity expressed in mhos/m, ϵ_r is the earth's dielectric constant, and f_{mc} is the frequency in megacycles. If an average value of 15 is assumed for ϵ_r , and a frequency of 120 kc is considered, then $n = \infty$ represents the case of a perfectly conducting earth, $n = 100$ represents the case for an earth of conductivity 10^{-2} mhos/m (good conductivity), and $n = 1$ is representative of an earth conductivity of 10^{-4} mhos/m (poor conductivity).

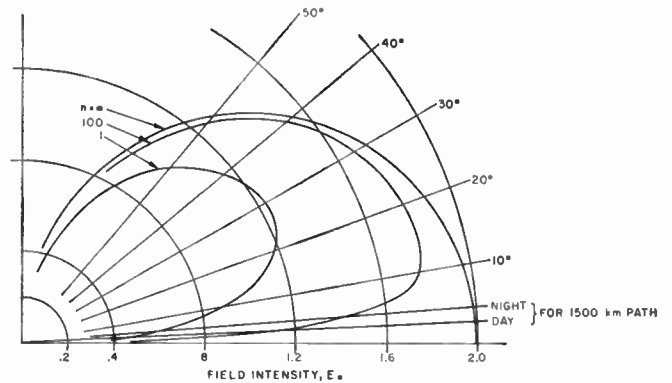


Fig. 4—Vertical radiation pattern of a short, base-fed, vertical dipole located on the surface of a flat earth. The angle of elevation for day and for night skywave propagation over a 1500-km path is shown.

It is apparent that the effect of the finite conductivity of the earth on skywave signal strengths is experienced chiefly when the propagation path describes a low angle with the horizon, and results in a reduction of skywave field intensity. This effect is caused by the rapid change of phase of the ground reflection coefficient for angles of incidence near the pseudo-Brewster angle. At angles of incidence less than the pseudo-Brewster angle the phase of the ground reflection coefficient approximates zero; at greater angles, near grazing incidence, the phase change approximates 180°. The pseudo-Brewster angle is a function of earth conductivity; the effects occur at smaller angles of incidence for poorer conductivities.

For the case of the transmitting antenna and the receiving loop antenna located on earth of finite conduc-

¹⁰ E. C. Jordan, "Electromagnetic Waves and Radiating Systems," Prentice-Hall, Inc., New York, N. Y., p. 621; 1950.

tivity, it is necessary to modify (3) and (5) to include the effect of the ground reflection factors. The amplitude of the downcoming skywave can then be rewritten as

$$E_d = \frac{E_u}{L} \sin i_g \|R\| DF_T \quad (7)$$

where F_T = a factor < 1 , introduced by the reflection coefficient of the ground in the vicinity of the transmitting antenna. The effective amplitude of the skywave, E_s , can then be rewritten as

$$E_s = 2E_d F_R \quad (8)$$

$$= 2 \frac{E_u}{L} \sin i_g \|R\| DF_T F_R \quad (9)$$

where F_R = a factor < 1 , introduced by the reflection coefficient of the ground in the vicinity of the receiving loop antenna.

Considering the earth as a reflector, F_T and F_R are the ratio of the actual vertical electric field at the surface of the earth to the electric field that would be measured if the earth were a perfect conductor. The antenna factors involve the complex ground reflection coefficients which have been studied by many workers for a flat ground.¹¹ Recently Wait and Conda¹² theoretically calculated the magnitude of the vertical electric field at the surface of a spherical earth of finite conductivity, and they have presented their results in a convenient graphical form. The antenna factors for a loop on the surface of the earth at a frequency of 100 kc are shown in Fig. 5. The data for this figure have been taken from the work of Wait and Conda.¹² Two comments should be made. The antenna factors, F_T and F_R are not the same as the so-called cut-back factor F used by Wait and Conda. Their factor F relates the actual field to the field in free-space, whereas our factors F_T and F_R relate the actual field to the field if the earth were a perfect reflector, and thus they differ simply by a factor of two. The significance of the field values for negative angles of incidence will be discussed in the next section. These refer to the diffracted field inside the shadow zone, *i.e.*, the field beyond the optical limit of a once-reflected skywave.

The antenna factors are of fundamental importance in the design of low-frequency radiating systems in the Arctic where the earth conductivity can be very poor. For example, for propagation over a path of 1500 km (*i.e.*, angles of elevation 3° – 4°), the skywave field intensity can be reduced by a factor of 4 in the case of poor conductivity (Arctic tundra), or 10 in the case of very poor conductivity (polar ice), from that produced by an antenna located on a perfectly conducting earth, that is, F_T and/or F_R would be equal to 0.25 and 0.1, respectively.

¹¹ Because of the assumed flatness of the earth, the curves in Fig. 4 are not valid at near grazing angles. At low frequencies the sphericity must be taken into account.

¹² J. R. Wait and A. M. Conda, "Pattern of an antenna on a curved lossy surface," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 348–359; October, 1958.

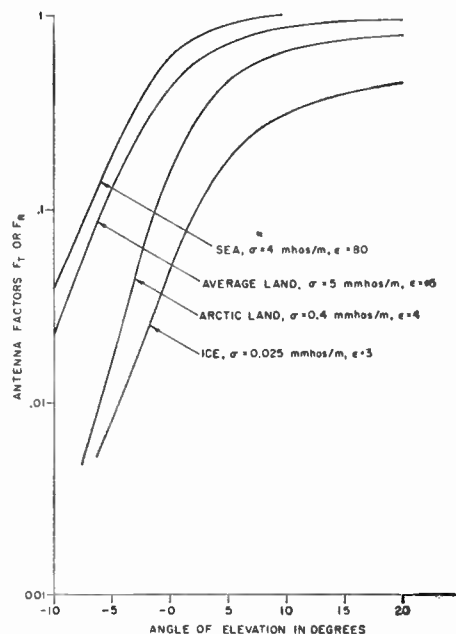


Fig. 5—Antenna factors F_T and F_R for a loop antenna on the surface of the earth at 100 kc.

Propagation Over Great Distances

As the transmission path length is increased to approach the condition for grazing incidence of a one-hop skywave, two oppositely directed effects are introduced. The first is the focusing effect of the ionosphere, which tends to increase the effective signal intensity (see Fig. 1), and the second is the decrease in low angle radiation from the transmitting antenna, which causes a decrease of the signal intensity (see Fig. 4).

The ionospheric reflection coefficients plotted in Fig. 2 represent reflection from a flat ionosphere; *i.e.*, the convergence factor of the ionosphere has been removed. For propagation over distances up to about 1000 km the focusing factor is small, but it became appreciable for greater transmission distances. The full theoretical increase due to focusing at the ionosphere will probably not be realized in practice because of irregularities in the ionosphere. Nevertheless, if correct account is taken of the convergence factor and the antenna pattern factors, (9) can be used to calculate the skywave field strength out to the limiting distance for a one-hop skywave.

For still greater transmission distances the path length exceeds that for a one-hop skywave. The received signal will then consist of skywave components reflected more than once from the ionospheric layer, and at least once from the surface of the earth. Ground-wave field intensities will, in general, be small compared to the skywave at these distances for frequencies between 80 and 200 kc.

If, from considerations of the high day-time absorption, it can be predicted that the received signal will be due predominantly to a two-hop skywave, then the field intensity at the receiving loop antenna can be represented simply as

$$E_s = 2 \frac{E_u}{L} \sin i_\theta \parallel R_{1\parallel} \parallel R_{2\parallel} D^2 D_G \parallel R_{G\parallel} F_T F_R \quad (10)$$

where D_G represents the divergence factor caused by the spherical surface of the earth, $\parallel R_{G\parallel}$ is the effective reflection coefficient of the finitely conducting earth, and L is the total propagation path length of the two-hop mode. $\parallel R_{1\parallel}$ and $\parallel R_{2\parallel}$ are the ionospheric reflection coefficients for the first and second reflection, respectively. In general, these ionospheric reflection coefficients will not be equal because the polarizations of the incident waves are not the same. However, it can be assumed that $\parallel R_{1\parallel} = \parallel R_{2\parallel}$ for propagation at very oblique angles of incidence, as a first-order approximation.

Very little experimental work has been done in low-frequency propagation at distances between 1500 and 3500 km. The available data include observations made by Goldberg and Woodward¹³ during 1950–1951 at a frequency of 100 kc over distances ranging to 4200 km; by Belrose and Thain during August, 1953 at frequencies of 133.15 and 115.3 kc over distances up to 3553 km; and by Crichlow, *et al.*¹⁴ during the summer of 1945 at a frequency of 180 kc over distances up to 2600 km. Measurements have also been made out to similar distances by Goldberg and Woodward, and by Crichlow, *et al.*, using equipment installed in aircraft. With the exception of the 100-kc measurements, the data are restricted to summer observations.

The available data for summer propagation of 100, 133.15, and 180 kc are plotted in Figs. 6–8, in the form of received signal strength, for 1-kw radiated power, as a function of distance from the transmitter. The ground-wave field intensity as shown was obtained from data at the shorter distances from the transmitter; the effective earth conductivity appropriate to these measured field intensities was deduced, and the curve extended to greater distances assuming this conductivity. Thus no account is taken of possible variation in the effective conductivity at greater distances. The discrepancy between field values so predicted and the actual values can be large. For example, the curve shown for the 133.15-kc data depends on the measured value at 810 km, and since the transmitter was located at Matsqui, B. C., and the measured value shown was recorded at Suffield, Alta., the effective conductivity of 10^{-3} mhos/m, as shown, is quite reasonable for propagation across the Rocky Mountains, but the field strength values predicted by this curve will be too small for reception at greater distances across the prairies where the conductivity is high.

It is noted, in Figs. 6–8, that the predicted values of summer daytime skywave field strengths are less than

those of the ground wave to distances of about 1500 km. The predicted values of nighttime skywave appear to be in reasonable agreement with measured values. The somewhat lower nighttime values obtained experimentally for 133.15 kc probably result from the method of observation; field strength values were recorded at 15-minute intervals, which provided only a relatively small statistical sample.

For propagation to distances greater than 1500 km the skywave field strength predominates, and decreases in amplitude exponentially with increasing distance. The outstanding feature of the data is the complete absence of a discontinuity in field strength at the optical limiting range of one-hop skywave. In addition, the measured field strengths at distances greater than 2000 km are greater than would be expected from calculations that assume a two-hop mode as given by (10). The difference between measured and calculated field strength is most marked during summer daytime, that is, when ionospheric absorption is great. For example, consider the propagation of 180-kc waves over a 2600-km path. The optical limiting range for one-hop daytime propagation is about 1880 km, the actual value being somewhat greater than this due to bending of the low-angle radio waves by the troposphere; see, for example, the discussion by Bean and Cahoon.¹⁵ Clearly, however, a 2600-km path extends well inside the shadow zone for a once-reflected skywave. Assuming a twice-reflected skywave to be the predominant mode, and a ground conductivity of 5×10^{-3} mhos/m, the field strength for 1 kw radiated, calculated from (10), is $0.002 \mu\text{v/m}$. The measured field intensity was $0.065 \mu\text{v/m}$, which is obviously very much greater than that calculated for the two-hop mode.

At large ranges a purely ray description of the first-hop skywave begins to lose its validity. Wait and Murphy¹⁶ considered the diffraction by the earth's bulge of the first-hop skywave for VLF propagation. They showed that the first-hop skywave did not cut off at the optical limiting range and, in fact, the field strength inside the shadow zone fell off smoothly, and decreased with distance in a manner not unlike the exponential attenuation of a ground wave in the diffraction region. A detailed study has since been carried out by Wait and Conda¹² of the computation of skywaves for ranges beyond their optical cutoff. The curves given can be used to estimate the attenuation of the field strength inside the shadow zone.

It seems likely that the most appropriate method of calculation would be that for a symmetrical path of propagation having the transmitter and the receiver located within separated shadow zones, since this path is the one of minimum phase. In the corresponding model, the signal would be propagated as a skywave

¹³ O. Goldberg and R. H. Woodward, "Radio wave propagation over long distances at 100 kc," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-2, pp. 40–53; March, 1952.

¹⁴ W. Q. Crichlow, J. W. Herbstreit, E. M. Johnson, and C. E. Smith, "Measurement Technique and Analysis of a Low Frequency Loran System," Operational Res. Staff, Washington, D. C., Rep. No. ORS-P-23-5; May, 1946.

¹⁵ B. R. Bean and B. A. Cahoon, "The use of surface weather observations to predict the total atmospheric bending of radio waves at small elevation angles," PROC. IRE, vol. 45, pp. 1545–1546; November, 1957.

¹⁶ J. R. Wait and A. Murphy, "The geometric optics of VLF sky-wave propagation," PROC. IRE, vol. 45, pp. 754–760; June, 1957.

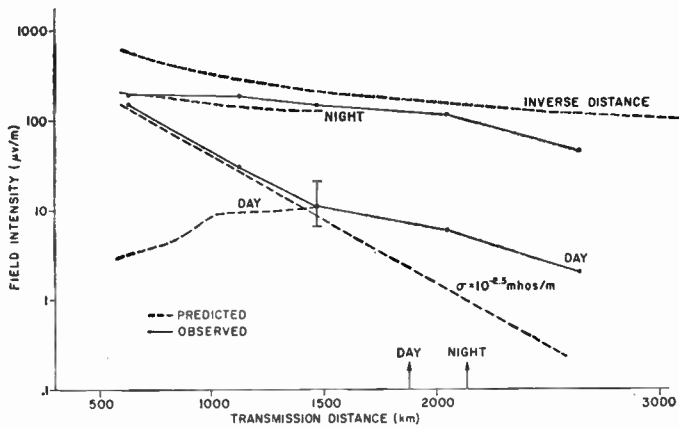


Fig. 6—Variation of field intensity with distance for 100-kc radio waves propagated over land. The values refer to a radiated power of 1 kw and are for observations made during the summer of 1951. Calculated values of ground-wave field intensity ($\sigma = 10^{-2.5}$ mhos/m) and skywave field intensity for summer daytime and nighttime propagation conditions are shown. The short vertical arrows mark the optical limiting range of a once-reflected skywave.

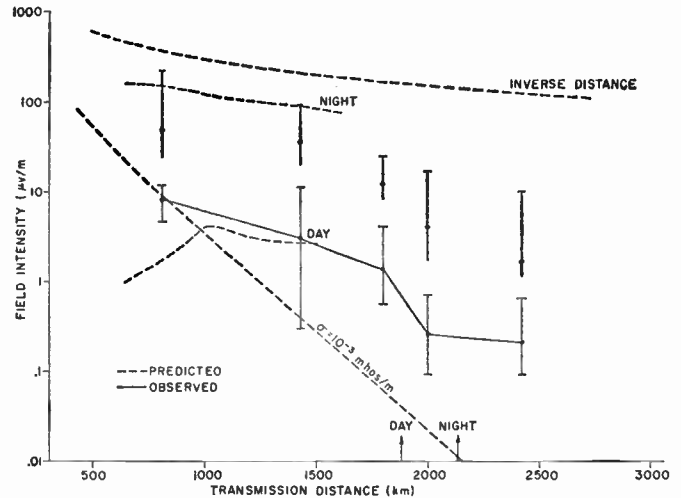


Fig. 7—Variation of field intensity with distance for 135.15-kc radio waves propagated over land. The values refer to a radiated power of 1 kw and are for observations made during August, 1953. Predicted values of ground-wave intensity ($\sigma = 10^{-3}$ mhos/m), and skywave field intensity for summer daytime and nighttime propagation conditions are shown. The short vertical arrows mark the optical limiting range of a once-reflected skywave.

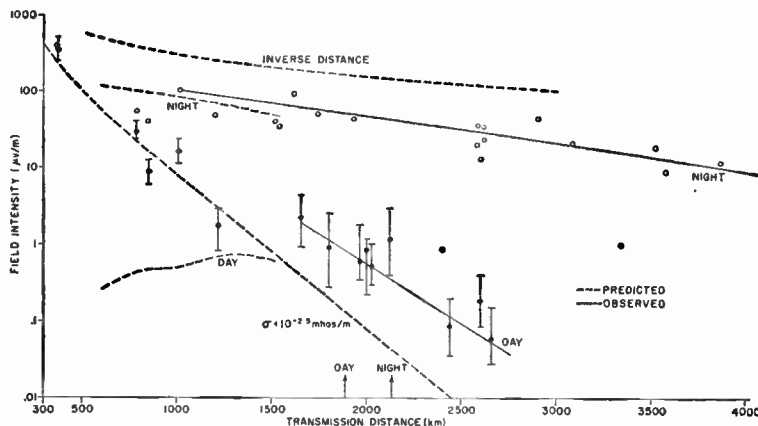


Fig. 8—Variation of field intensity with distance for 180-kc radio waves propagated over land. The values refer to a radiated power of 1 kw and are for observations made during the summer of 1945. Calculated values of ground-wave field intensity ($\sigma = 10^{-2.5}$ mhos/m), and skywave field intensity for summer daytime and nighttime propagation conditions are shown. The short vertical arrows mark the optical limiting range of a once-reflected skywave.

over the optical limiting distance of a one-hop skywave and as a diffracted wave within the shadow zones located at the extremities of the skywave path. Eq. (9) can be used directly if effective values for F_T and F_R are used, and if the other parameters are those appropriate to skywave propagation over the optical limiting range of a once-reflected skywave. For propagation of a 180-kc wave over the 2600-km path discussed previously, the shadow zones would extend for a distance of 360 km from the transmitter and the receiver. If the conductivity within both shadow zones is assumed to be 5×10^{-3} mhos/m, the effective values for F_T and F_R can be estimated¹² and are 0.13. It is estimated⁷ that the focusing factor is 2.8, and the ionospheric reflection coefficient is 0.005 ($f \cos i$ for 180 kc at near grazing incidence at the ionosphere is about 26 kc). Hence (9)

leads to the estimate, $E_s = 0.07 \mu\text{V/m}$. This calculated value of E_s is in better agreement with the measured value of $0.065 \mu\text{V/m}$ than the value previously calculated on the basis of a two-hop mode. Ionospheric roughness tends to reduce the effective value of the focusing factor D , and so reduces the calculated field strength. Inclusion of atmospheric refraction increases the limiting distance of the once-reflected skywave, and so increases the factors F_T and F_R , which increase the calculated field strength.

The suggestion is, in fact, that low-frequency waves reflected once from the ionosphere contribute to the field intensity at distances far beyond the geometrical optical limit for these signals, in fact to distances of 2600 km and greater. During a summer day this can be the predominant mode.

The diurnal variation of field strength at a frequency of 133.15 kc is shown in Fig. 9 from data obtained at various receiving locations. These data are included to show the diurnal variation of field strength and the manner in which the variation changes with transmission distance.

The Effects of Skywave Fading

The discussion of the preceding section has been concerned with mean values of signal amplitude only. In practice, the received signal has a statistical variation of amplitude about the mean as a result of the characteristic fading. In the engineering of radio systems the received signal is considered customarily in terms of its long-term amplitude variation. Unfortunately, complete statistics are not available for low radio frequencies and, although the factors introduced by signal fading are of considerable importance, this discussion must be limited to a few general remarks.

The fading speed and the depth of fading depend on the frequency, the transmission distance, and the time of day. Only the skywave fades, so the fading of it alone must be first considered, then the resultant fading of the combined ground and skywaves can be calculated. The speed of fading is least during the daytime and greatest during the nighttime. The daytime skywave appears to be predominantly specular. The amplitude is substantially constant and the phase change is regular. The rate of change of phase and the total change can be predicted, because they depend on the diurnal variation of the apparent height of reflection. The fading during nighttime is much more irregular. The phase fluctuations contain a random component, and a component due to bodily movement of the reflection height. The amplitude fluctuations are approximately Rayleigh. The nighttime fading speed increases during geomagnetic storms. The normal fading speed, which is to a first approximation independent of frequency and distance, has a substantially constant quasi-period of about 7 minutes. Superimposed on this slow fade is a small component whose fading speed is more rapid. The fading speed of this small component varies directly with the effective frequency for the transmission path, *i.e.*, $f \cos i$. The lower the effective frequency, the smaller the depth of fading.

For propagation over short distances, when the daytime signal is due predominantly to the ground wave, amplitude fluctuations are small. During the nighttime, the skywave contribution will cause shallow fluctuations in received signal amplitude, and the resultant amplitude-probability distribution is of the displaced-Gaussian type. The standard deviation of the amplitude fluctuations represents the mean value, which in this case is the same as the rms value, of the skywave contribution.

For propagation over long distances the skywave amplitude, even during daytime, will be comparable to or greater than the ground-wave amplitude. The day-

time fading will be mainly due to phase interference between the ground and skywaves. Signal fading is then at a slow rate, and the statistical analysis is hardly applicable because of the small number of fading cycles. For propagation over great distances, the skywave predominates and the signal fluctuations can be described statistically.

The amplitude fluctuations which have been most extensively studied are those of nighttime. Ebert¹⁷ has shown that the long-term amplitude probability distribution of momentary values of field strength is almost exactly Rayleigh distributed for radio waves 164 to 1466 kc propagated to great distances. The quasi-maximum and quasi-minimum values of field intensity (*i.e.*, 10 per cent and 90 per cent relative probability values) are approximately 4.8 db greater than and 6.3 db less than the 50 per cent, or median values. However, King¹⁸ has shown the probability distribution of amplitude of 845-kc waves propagated over a distance of about 1500 km is more nearly log-normal than Rayleigh. This implies that the amplitude assumes large values more often than would be expected on the Rayleigh distribution. Crichlow, *et al.*¹⁴ have plotted day and night summer values of 180-kc rms field intensity exceeded for a given percentage of time for a large number of transmission paths, ranging to 5000 km. Some of their plots show good log-normal distributions, whereas the distribution of others, averaged over the same period of time, is closer to Rayleigh. There seems to be no systematic change in the character of the fading with either distance or direction. The median values of the rms nighttime field strength exceeded 10 per cent and 90 per cent of the time were, respectively, 6.7 db greater than, and 9 db less than the 50 per cent, or median value. Thus, although the depth of fading is greater than that given by Ebert, the median values of the 180-kc observations are more nearly Rayleigh distributed than log-normal. Ebert's values are based on data which are by far the most extensive observations extending over a year.

Atmospheric Noise

In the engineering of a low-frequency circuit, atmospheric noise is one of the most important factors to be taken into consideration. To assess the atmospheric noise in terms of the interference effect on signal reception requires a fairly complete knowledge of the noise characteristics and, in addition, the effect these characteristics will have on the receiving system. The practical approach to atmospheric noise problems requires a statistical description of the noise fluctuations in amplitude as a function of frequency, time of day, season, and geographic location of receiving site. This information together with information concerning the characteristics

¹⁷ W. v. Ebert, "Statistic der weidstanz-nachtpropagation von langund mittlewellen," *Tech. Mitt. Schweiz. Telegr-Teleph. Verw.*, vol. 34, pp. 198-209; May 1, 1956.

¹⁸ J. W. King, "The fading of radio waves reflected at oblique incidence," *J. Atmos. Terr. Phys.*, vol. 12, pp. 26-33; January, 1958.

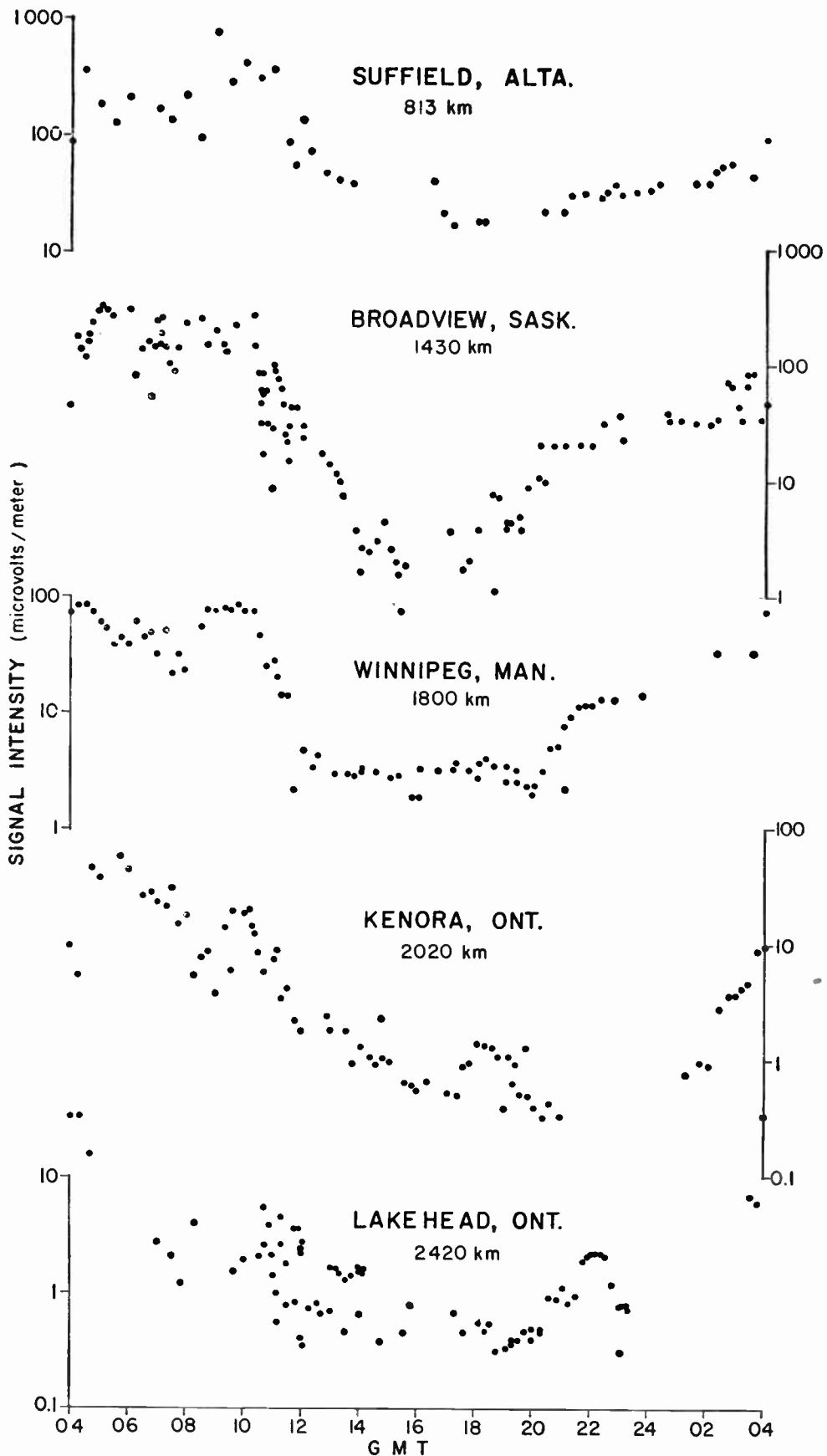


Fig. 9—Diurnal variation of field intensity of 133.15-kc radio waves at very oblique angles of incidence. The values refer to an inverse-distance attenuated field intensity at 1 km of 1010mv/m.

of the noise signal itself can then be utilized to determine optimum engineering parameters to facilitate signal detection. Considerable statistical work has been carried out by many workers, with little agreement concerning which parameters should be measured. Most commercial noise measurement equipment is of the quasi-peak variety, where the detector charge and discharge time constants are designed to approximate various types of subjective interference. The quasi-peak measurements have not been related to basic units, and it is necessary to define the measurements in terms of the time constant characteristics. As a result, the measurements made by the various equipments have little agreement and little interrelationship. It was recommended by the National Bureau of Standards, in a recent report to Commission IV of USRI, that steps be taken toward standardization of a mathematically significant measure of terrestrial radio noise and that the single parameter most closely representative of the interference effect is the noise power.

In recent years, the most extensive measurements, of the statistical characteristics of atmospheric radio noise, are those of Watt and Maxwell.¹⁹ The statistical parameters describing the noise depend on both time and geographic location.

Atmospheric noise is due primarily to electromagnetic radiation from lightning discharges. The majority of atmospheric noise originates from several large storm centers located over land masses near the equator. Those that affect the reception of noise in North America are located principally in Central America and the northern part of South America. There is a seasonal shift in these main storm centers, and there is considerable thunderstorm activity in summer over the major part of the North American continent. The precise seasonal and diurnal variation of the atmospheric noise intensity is dependent upon the receiving location. In general, the intensity is greatest at nighttime, greater during a summer day than a winter day, and greater in the late afternoon than in the early morning. The latter effect results from the diurnal variation of ionization at low levels in the ionosphere, and from the natural tendency for thunderstorm activity to be of greatest intensity during the midafternoon. Because major storm centers are to the south of Canada, the intensity of atmospheric noise decreases with increase of latitude. There is some indication that certain types of auroral activity may generate noise in the low-frequency band at high geomagnetic latitudes.

A good detailed study of the diurnal variation of the intensity of VLF noise, and the change of the diurnal variation with season, has been given by Lauter.²⁰

¹⁹ A. D. Watt and E. L. Maxwell, "Measured statistical characteristics of VLF atmospheric radio noise," *Proc. IRE*, vol. 45, pp. 55-62; January, 1957. Also, "Characteristics of atmospheric noise from 1 to 100 kc," *Proc. IRE*, vol. 45, pp. 787-794; June, 1957.

²⁰ E. A. Lauter, "Der atmosphärische störpegel in langwellenbereich und seine tages- und jahres zeit lichen variation," *Z. Met.*, vol. 10, pp. 110-121; April, 1956.

Graphs and charts have been made available by various organizations for the prediction of world-wide atmospheric noise levels.²¹ Median values of rms noise in a 1-kc bandwidth are predicted for 4-hour time blocks for each season of the year. However, the predictions are seriously limited by the small number of data-recording stations. To obtain information related specifically to atmospheric noise conditions in Canada, the Radio Physics Laboratory has developed and is now operating equipment at two locations. These measure the rms noise voltage in a 300-cycle bandwidth at frequencies of 11, 30, and 135 kc. Preliminary data concerning the mean logarithmic power of atmospheric noise at 10 and 107 kc have already been described by McKerrow.²²

The median values of the rms noise voltages obtained at Ottawa for the month of November, 1957, at a frequency of 11 kc are shown in Fig. 10. The spread in hourly values is indicated by the upper and lower decile plots. It will be noted that the median values vary from 38 to 46 db above 1 $\mu\text{v}/\text{m}$ in a 1-kc bandwidth. The rms values predicted for 1957 at Ottawa at 11 kc are shown in Fig. 11 for 1000 and 1800 hours as a function of season. The graph shows a seasonal variation of 15 db in rms noise amplitude.

LOW-FREQUENCY SYSTEMS ENGINEERING

The techniques required to exploit fully the various parts of the radio spectrum are determined customarily by the characteristics of the propagating medium. At low radio frequencies, the propagation medium is relatively stable and fading occurs at a slow rate relative to normal signaling speeds. Conventional receiving techniques can accommodate the fluctuations in signal amplitude that occur during the nighttime hours. Thus, at low radio frequencies, the propagating medium does not impose any unusual restrictions on the modulation systems to be employed. However, important limitations are imposed by other considerations such as system costs, available spectrum, physical limitations of the transmitting system, etc.

In the design of a low-frequency communication system, the cost is almost entirely determined by the design of the transmitting installation and by the transmitter power requirement. To effect economy of cost, the engineering design efforts must be directed toward minimizing the power requirement by the use of efficient radiating and modulation systems. The various factors involved in the system design are discussed separately.

Low-Frequency Radiators

The amount of power radiated by the antenna depends on the output power of the transmitter, the losses in the transmission line, and the radiation efficiency of the antenna. Practical antennas for use at low radio fre-

²¹ See, for example, NBS Circular 557.

²² C. A. McKerrow, "Some recent measurements of atmospheric noise in Canada," *Proc. IRE*, vol. 45, pp. 782-786; June, 1957.

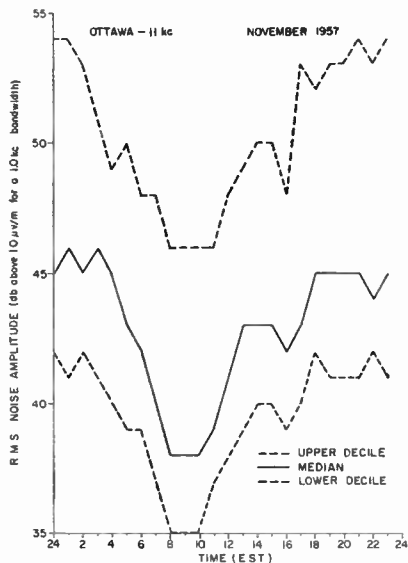


Fig. 10—Hourly median rms values of atmospheric noise recorded at Ottawa at 11 kc for the month of November, 1957.

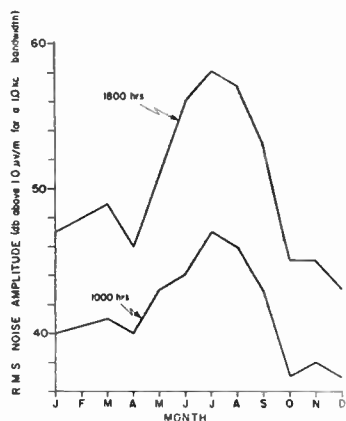


Fig. 11—Predicted hourly median rms values of atmospheric noise for Ottawa for a frequency of 11 kc during 1957.

frequencies are physically short compared with the operating wavelength; they also have low radiation resistance and relatively high capacitive reactance. To resonate the antenna at the operating frequency, it is necessary to insert an inductance in series with the antenna. Because the resistance of this inductance is usually of the same order as the radiation resistance, it is customary to consider this tuning inductance as a part of the antenna system.

The total antenna resistance, R_a , can be represented as the sum of four resistances, as follows:

$$R_a = R_r + R_g + R_i + R_w \tag{11}$$

where R_r is the radiation resistance, R_g is the ground terminal resistance, R_i is the resistance equivalent to the insulation losses, and R_w is the resistance equivalent to the conductor losses. The radiation efficiency of the antenna system, η , can then be expressed as

$$\eta = \frac{R_r}{R_a + R_c} \times 100 \text{ per cent} \tag{12}$$

$$\eta = \frac{R_r}{R_t} \times 100 \text{ per cent} \tag{13}$$

where R_c is the resistance of the antenna tuning inductance and R_t is the total resistance of the antenna system.

With proper antenna design the equivalent series loss resistance due to insulation leakage, R_i , is small compared with the ground terminal loss resistance, R_g . However, small fluctuations in insulator leakage produced by wetting by fog, rain, or sleet, may cause variation in the base impedance of the antenna which results in serious detuning. This problem is of particular importance in the case of FSK telegraphy because detuning of the antenna will result in signal distortion. Adequate insulation must be provided, and extreme precautions must be taken to restrict severely both the magnitude and the variability of the insulator losses.

The ground terminal resistance, R_g , can be reduced effectively by the provision of an extensive ground screen. The design of an efficient screen is a major problem in the over-all design of the antenna system. The dimensions of the ground screen for optimum performance depend upon the characteristics of the ground at the antenna installation. The various design factors must be considered on a comparative basis of effectiveness vs cost. A detailed theoretical discussion of the dependence of the antenna input resistance on the number, diameter, and length of the radial conductors in the screen, as a function of operating frequency and ground constants, is given by Wait and Surtees²³ and by Wait and Pope.²⁴ Applicable, also, is the paper by Wait²⁵ concerning the case of a ground screen located on the surface of a layer of ice or snow, representative of conditions encountered with a portable antenna operating over ice or snow-covered terrain.

Since the actual antenna height of low-frequency antennas is necessarily restricted to a small fraction of a wavelength, some form of capacitive top-loading should be provided to increase the equivalent electrical height. The top-loading may be in the form of a caged-T configuration requiring two supporting towers, or in triangular or diamond configurations requiring three or more supporting towers. The characteristics of such antenna configurations have been investigated in considerable detail (see Laport²⁶). Because of the high construction costs of multitower antenna systems, the Radio Physics Laboratory undertook an investigation of single-tower, top-loaded systems, directed toward the design of a radiating system that would provide per-

²³ W. J. Surtees and J. R. Wait, "Impedance of a top-loaded antenna of arbitrary length over a circular grounded screen," *J. Appl. Phys.*, vol. 25, pp. 553-555; May, 1954.

²⁴ J. R. Wait and W. A. Pope, "Input resistance of LF unipole aerials," *Wireless Eng.*, vol. 32, pp. 131-138; May, 1955.

²⁵ J. R. Wait, "Note on LF portable antennas operating over ice and snow," *Proc. Symp. on the Propagation of VLF Radio Waves, Boulder, Colo., January 23-25, 1957*, vol. 4, appendix C, pp. 62-66.

²⁶ E. A. Laport, "Radio Antenna Engineering," McGraw-Hill Book Co., Inc., New York, N. Y., 1952.

formance comparable to multitower systems but at less cost. One form of antenna system that appeared promising was the one investigated earlier by Smith and Johnson²⁷ for use at broadcast frequencies, and referred to by them as the umbrella-type top-loaded antenna. The top-loading consists of a number of wires strung obliquely to ground from the top of the radiator, electrically bonded to the top of the radiator, and insulated from the ground. The important physical parameters in such a system are shown in Fig. 12(a) and are the following: the height, h , of the radiator; the horizontal distance, D , from the base of the radiator to the extremities of the umbrella wires; and the vertical distance, d , from the top of the radiator to the height at which the umbrella wires are broken by an insulator. The experimental work by Smith and Johnson concerned a radiator with a height, h , of 300 feet, and with a distance, D , of 350 feet. Later work by Smeby²⁸ concerned a theoretical study of similar types of antenna systems but restricted to the case where $D = h$.

At the Radio Physics Laboratory, the umbrella-type top-loaded system was investigated to determine both its suitability for application at low frequencies, and the dependence of the electrical characteristics on the various physical parameters. A comparison of the two papers^{27,28} had indicated that some advantage might be gained in radiation resistance by increasing D with respect to h . At the time of preparation of our final report,²⁹ a paper was published by Smith, Hall, and Weldon³⁰ that described further experimental work using 8 and 12 umbrella wires and various values of D greater than the radiator height.

The experimental investigations of Belrose and Thain²⁹ were made using a 70-foot vertical radiator and 8 umbrella wires. The results of the series of measurements are shown in Figs. 12 and 13. Values of radiation resistance were deduced by field intensity measurements. From inspection of Fig. 12(b) and 12(c) it can be seen that the radiation resistance, R_r , increases with the parameter D . It is apparent that D should be made as large as is structurally possible. From inspection of Fig. 12(a) and 12(b), it can also be seen that for a fixed value of D , the radiation resistance varies with parameter d , and is at a maximum when d is approximately equal to $3h/7$. By increasing d to greater values the capacitive reactance of the system is decreased. Thus, by increasing d beyond the value for optimum radiation, increased bandwidth can be obtained at the expense of radiation efficiency. This effect permits the de-

signer to trade efficiency for bandwidth if bandwidth is an important requirement.

The experimental measurements were scaled for various antenna heights up to 300 feet for values of d equal to $3h/7$ and $5h/7$. The calculated radiation efficiencies of the systems were found to be comparable with T-type antennas of equivalent height. In order to check the extrapolation, and to meet a requirement for a radiator for propagation experiments, an umbrella-type top-loaded antenna system with 8 umbrella wires was installed at Ottawa. The physical parameters of this antenna were $h = 250$ feet, $d = 3h/7$, and $D = 350$ feet, the latter length being limited by structural considerations. The measured and predicted values of the input impedance as a function of frequency are shown in Fig. 14. Subsequently, a similar type of antenna system was installed by the Royal Canadian Air Force at St. Jacques, Que., but with top-loading parameters such that $d = 5h/7$. Measurements of this latter antenna system are shown in Fig. 15. For comparison, measurements of antenna reactance and resistance for a caged-T type of antenna system 300 feet high are shown in Fig. 16. The values given in Fig. 16 are the mean values obtained for three such systems installed by the Royal Canadian Navy.

Little effort has been devoted to the experimental determination of optimum ground screen parameters, or of the variation of ground losses as a function of frequency. The difficulties of such work at the low radio frequencies are evident from considerations of the magnitudes of the parameters involved. Values deduced from measurements of the 250-foot umbrella-type top-loaded antenna at St. Jacques and the 300-foot caged-T antenna of the Royal Canadian Navy are listed in Table 1.³¹ The ground-loss resistances for the 250-foot antenna at Ottawa were very much lower than those of all other antenna systems of similar size which have been measured by the authors. Using the values of Table 1, and assuming antenna tuning inductance Q factors of 1000, the radiation efficiencies and the operating bandwidths are shown in Fig. 17, as a function of operating frequency for the 250-foot and the 300-foot systems. The variation of radiation efficiency as a function of height is shown in Fig. 18 for a frequency of 150 kc.

The use of narrow-band radiating systems on a continuous basis throughout extreme changes of weather has indicated the desirability of providing a facility at the transmitter for the remote tuning of the antenna inductance. An automatic tuning system has been described recently by Wolff.³²

²⁷ C. E. Smith and E. M. Johnson, "Performance of short antennas," *Proc. IRE*, vol. 35, pp. 1026-1038; October, 1947.

²⁸ L. C. Smeby, "Short antenna characteristics—theoretical," *Proc. IRE*, vol. 37, pp. 1185-1194; October, 1959.

²⁹ J. S. Belrose and R. S. Thain, "The Characteristics of Short Vertical Radiators for Use at Low Radio Frequencies," *Defence Res. Telecommun. Establ., Ottawa, Ont., Can., DRTE/RPL Rep. No. 9-0-1*; September 2, 1954.

³⁰ C. E. Smith, J. R. Hall, and J. O. Weldon, "Very high power long wave broadcasting station," *Proc. IRE*, vol. 42, pp. 1222-1235; August, 1954.

³¹ Ground systems:

250-foot umbrella-type; 120 buried radials 1000 feet long bonded to a horizontal 8-foot copper plate buried at the base of the tower (center cut out to pass base pedestal).

300-foot T-type; 120 buried radials 500 feet long and a 4-foot by 4-foot copper plate buried on edge at the center of the radial system.

³² H. G. Wolff, "High-speed frequency-shift-keying of LF and VLF radio circuits," *IRE Trans. on Communications Systems*, vol. CS-5, pp. 29-42; December, 1957.

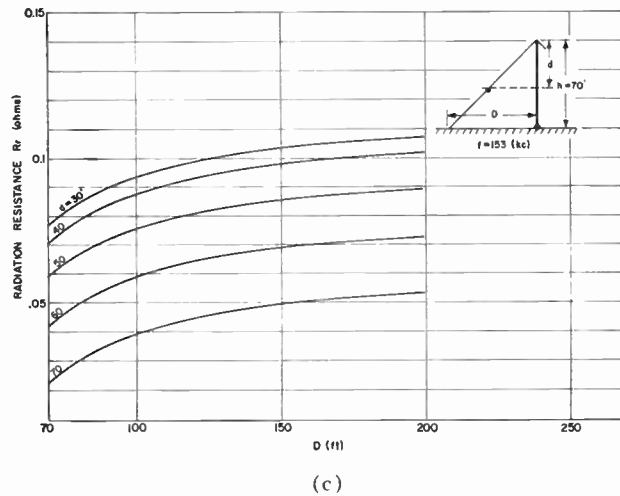
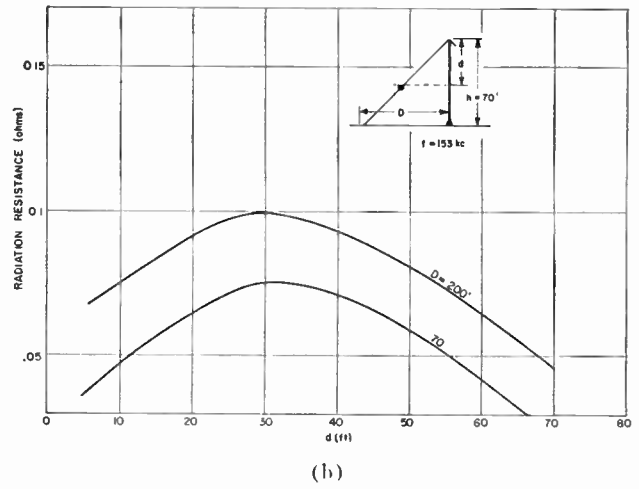
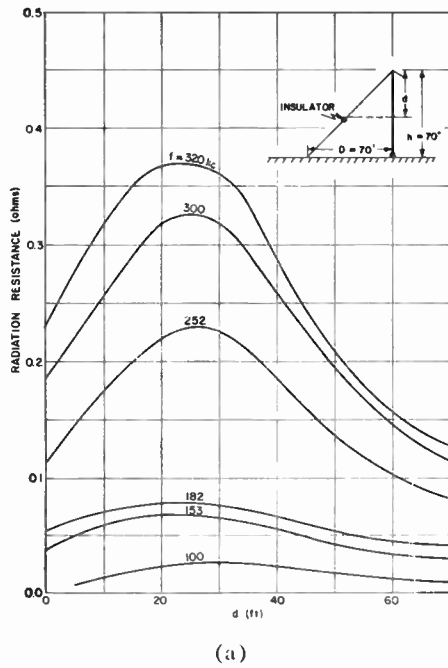


Fig. 12—Radiation resistance as a function of frequency and antenna parameters for a base-fed 70-foot umbrella-type top-loaded antenna.

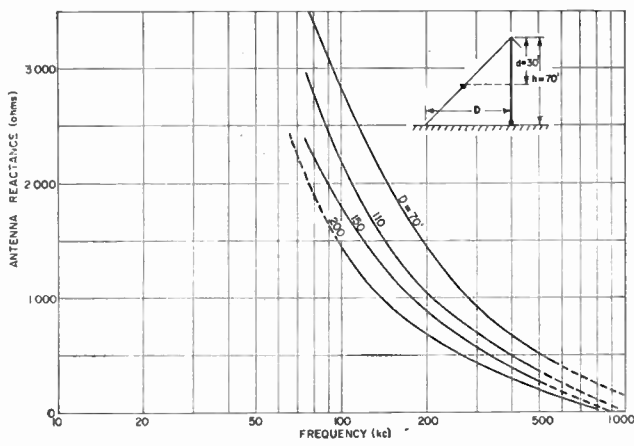
The Loop Receiving Antenna

The use of a screened loop antenna for the reception of low-frequency radio waves has been studied in some detail by Belrose.³³ The studies resulted in the conclusion that a small loop antenna is the most desirable form of antenna for the reception of low-frequency waves in either fixed or fixed-mobile operation in Canada. The main advantages accrue from the directional properties and the electrostatic shielding. The sharp directional null, characteristic of the loop antenna, permits discrimination against interference from a par-

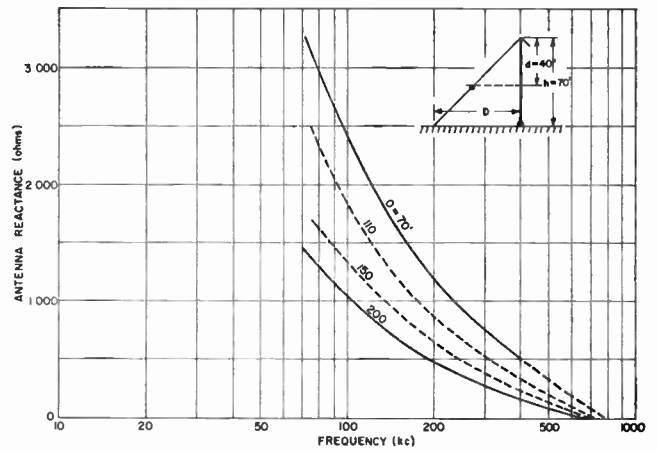
ticular direction. The electrostatic shielding reduces the effect of precipitation static and noise of an electrostatic nature such as local interference from motors, etc.

The basic factors that affect the design of a loop antenna have been discussed in detail by Belrose,³³ and only the conclusion will be given in this paper. Of the various configurations studied, the one found to provide the best signal-to-noise ratio (SNR) is illustrated in Fig. 19. The loop is tuned directly to the operating frequency by a capacitor, C , connected across the loop winding. The center of the loop winding is grounded to the center of the electrostatic shield, and one side of the capacitor, C , is connected to the grid of the first amplifier tube in a preamplifier, thus utilizing only one-half of the signal voltage across the capacitor, but maintaining

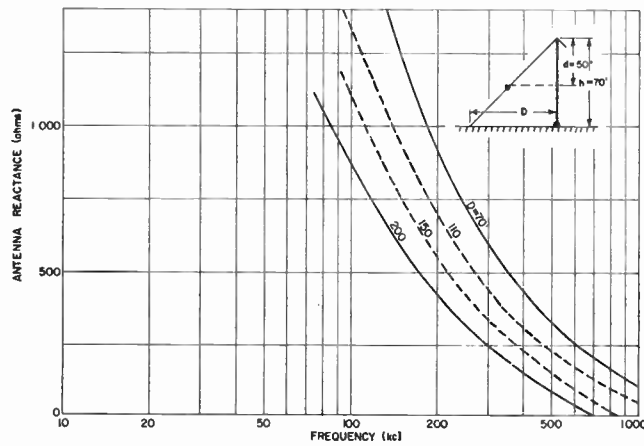
³³ J. S. Belrose, "The Use of a Screened Loop Antenna for Reception of Low-Frequency Radio Waves," Defence Res. Telecommun. Establ., Ottawa, Ont., Can., DRTE/RPL Rep. No. 10-0-1; December 11, 1952.



(a)



(b)



(c)

Fig. 13—Capacitive reactance as a function of frequency and antenna parameters for a base-fed 70-foot umbrella-type top-loaded antenna.

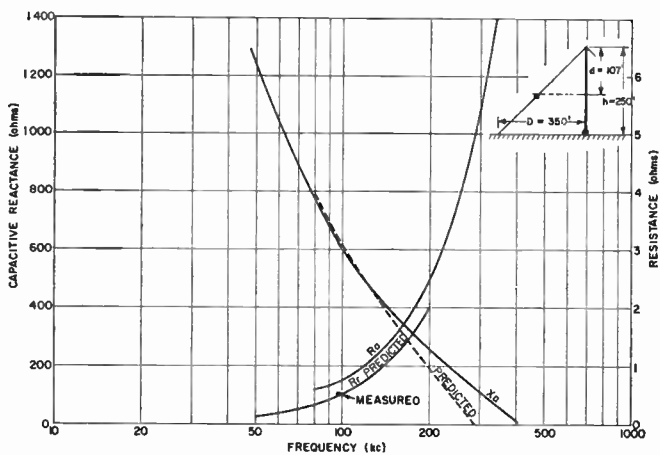


Fig. 14—Antenna resistance, R_a , radiation resistance, R_r , and capacitive reactance, X_a , as a function of frequency for a 250-foot umbrella-type top-loaded antenna, located at the Radio Physics Laboratory, Ottawa, with top-loading, d/h , designed for maximum radiation efficiency.

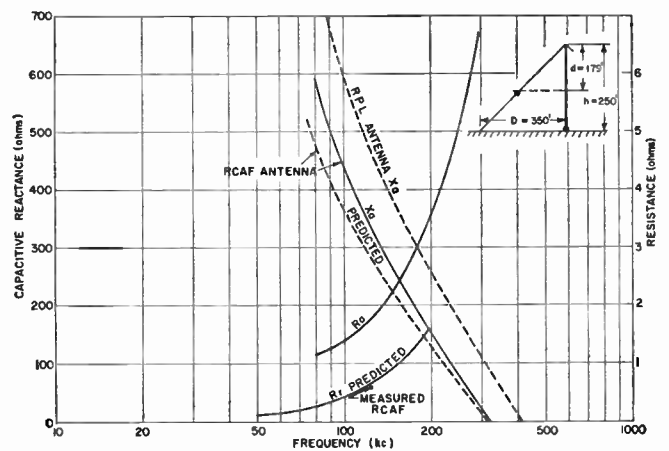


Fig. 15—Antenna resistance, R_a , radiation resistance, R_r , and capacitive reactance, X_a , as a function of frequency for a 250-foot umbrella-type top-loaded antenna, located at RCAF, St. Jacques, P. Q., with top-loading, d/h , providing greater than the optimum loading for maximum radiation efficiency, in order to provide a wider bandwidth of operation (RPL antenna reactance curve superimposed for comparison).

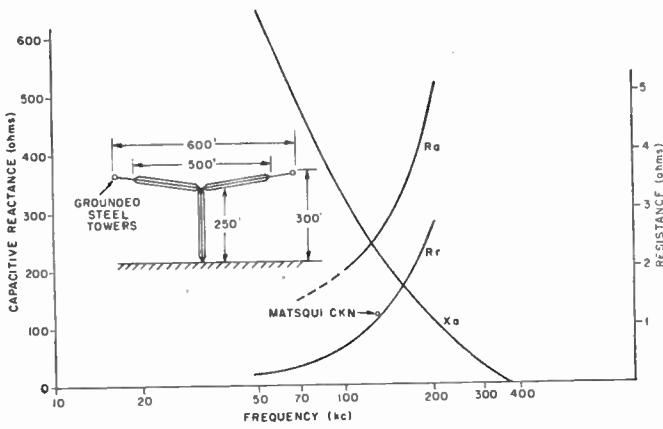


Fig. 16—Antenna resistance, R_a , radiation resistance, R_r , and capacitive reactance, X_a , for a 300-foot caged T-type top-loaded antenna. The values provide the average characteristics of 3 RCN installations.

TABLE I
GROUND-LOSS RESISTANCE, R_g (OHMS)

f (KC)	250-Foot Umbrella-Type	300-Foot T-Type
80	1.0	—
100	1.0	1.3
120	1.1	1.4
140	1.3	1.4
160	1.4	1.6
180	1.7	1.7
200	1.9	2.1

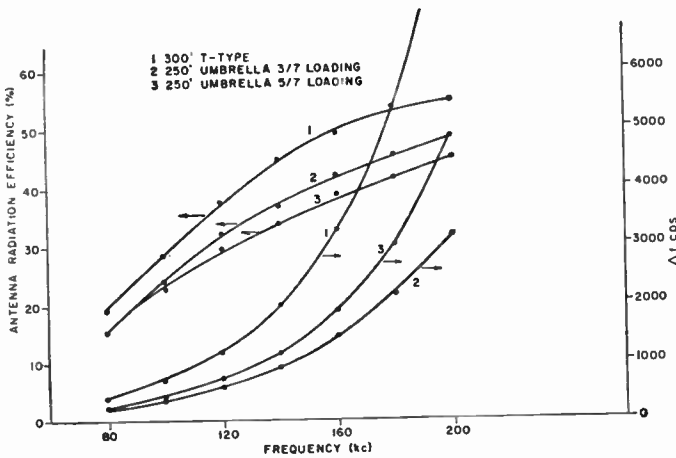


Fig. 17—Comparison of the radiation efficiencies and antenna bandwidths of the 3 antenna systems of Fig. 14-16, as a function of operating frequency.

a balance to ground. It is usually convenient to locate the loop and preamplifier in a small wooden hut several hundred feet from the receiver building. The signal output of the preamplifier should be balanced to ground by a transformer with an electrostatic screen. The transmission line connecting the preamplifier to the receiver should be twin-conductor screened cable. If the transmission line is extremely long, and hum is introduced, it may be necessary, in some cases, to break the shield of the cable at the preamplifier, and to ground the center of the input coil at the receiver.

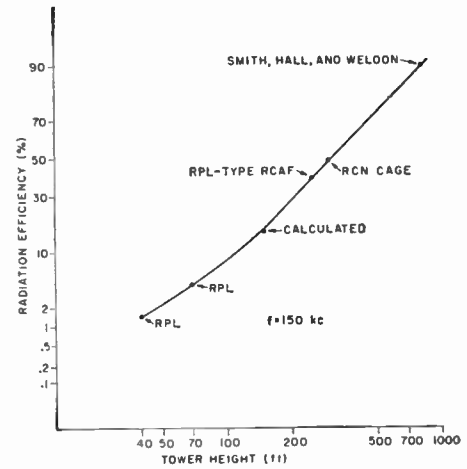


Fig. 18—Variation of the radiation efficiency at 150 kc with antenna height for an umbrella-type top-loaded antenna with top-loading, d/h , for maximum radiation efficiency.

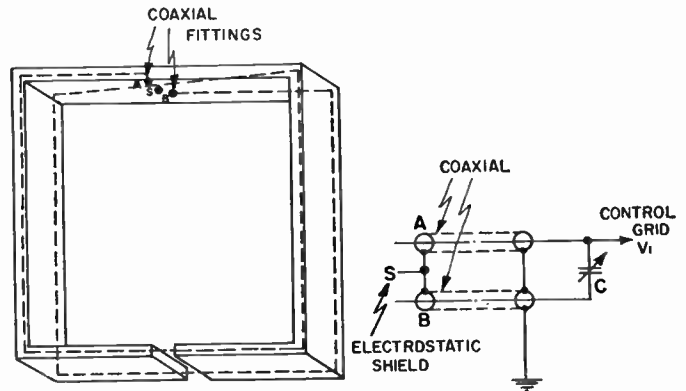


Fig. 19—Simplified two-turn loop antenna to illustrate the method of shielding, and balancing to ground.

With a 4-foot square loop antenna of 36 turns, it was found that a signal field intensity $0.4 \mu\text{v/m}$ at 150 kc produced a signal 20 db above the fluctuation noise of the loop preamplifier in a 100-cps bandwidth. This order of sensitivity is comparable to that required to detect the lowest atmospheric noise levels in Canada. Furthermore, it was determined that the sensitivity of the loop antenna has little dependence on the loop inductance but varies with the area and frequency.

Additional studies by Belrose³⁴ concerned the use of ferromagnetic materials in loop antenna design. The work was of an exploratory nature only, but the results were quite promising. For example, a 10-db SNR was obtained in a 100-cps bandwidth for a field intensity of $1 \mu\text{v/m}$ at 150 kc utilizing a rod-type antenna unit approximately 12 inches in length and $1\frac{1}{2}$ inches in diameter.

Modulation Systems

It has been indicated that the limitations of practical transmitting antennas rather than limitations of the

³⁴ J. S. Belrose, "Ferromagnetic loop aeriels for kilometric waves," *Wireless Eng.*, vol. 32, pp. 41-46; February, 1955.

propagating medium are the principal factors affecting the choice of modulation systems at low radio frequencies. It has been shown that in order to economize on transmitting installation costs it is necessary to restrict antenna height, but to design for high radiation efficiency, necessarily restricting transmission bandwidth. Other considerations lead also to restriction of transmission bandwidth, and influence further the choice of acceptable forms of transmission modulation. Some of these considerations are as follows:

- 1) In the 120 kc of radio spectrum between 80 and 200 kc, there are a large number of radio systems in operation for communication and navigation purposes. The use of minimum bandwidths is essential to reduce adjacent channel interference at the receiving terminals, and to facilitate economical use of the radio spectrum.
- 2) The effects of atmospheric noise have been discussed in an earlier section. Narrow-band operation permits the use of minimum signal power. During local thunderstorms the noise becomes quite spikey in nature and, even under normal noise conditions, is more spikey than noise that has a Rayleigh distribution. Strong impulsive interference effects can be reduced by amplitude limiting techniques in the receiver, preferably before bandwidth limiting circuits.

The preference for the employment of narrow-band systems at low radio frequencies implies the requirement for equipment that has a high order of frequency stability, and restricts information transmission to low signal rates. This restriction, in turn, permits only the use of hand-keyed or machine-keyed Morse at conventional signaling speeds, or else the use of single-channel telegraphy at 60 wpm.

Practical narrow-band teletype modulation systems suitable for low-frequency radio communications include shaped amplitude modulation³⁵ narrow-band frequency shift³⁵⁻³⁸ and narrow-band phase shift.³⁹ A special form of frequency shift, utilizing matched filters in the receiver at the mark and space frequencies, might be employed also, providing that the extent of the frequency shift is minimized. Because of the many large spikes of noise that occur randomly within the low-fre-

quency spectrum, some form of frequency modulation is to be preferred.

A number of theoretical and experimental studies have indicated that the optimum receiver bandwidth for frequency shift teletype in the presence of thermal noise is equal to about 5/3 times the frequency shift. It has also been indicated that the optimum frequency shift is equal to about twice the signaling speed expressed in cps. (A signaling speed of 23 cps corresponds to 60-wpm, 5-unit, stop-start teletype.) At the Radio Physics Laboratory, the dependence of SNR on the extent of the frequency shift was measured in the optimum receiving bandwidth. The signal was mixed with the constant output from a noise diode generator, and for each value of frequency shift the signal required for error-free FSK operation was measured in a detector bandwidth equal to 5/3 times the frequency shift. The results are plotted in Fig. 20 and indicate an optimum

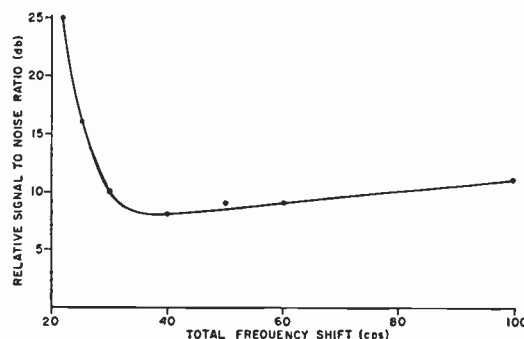


Fig. 20—Dependence of SNR, in relative value of db, on total frequency shift in an optimum receiver bandwidth as determined by measurements at the Radio Physics Laboratory.

quency shift of about 40 cps. The bandwidth corresponding to the latter measurement was 66 cps. However, neither the value for the optimum frequency shift nor the bandwidth are particularly critical, and both values may be increased considerably without significant deterioration in the presence of thermal noise, provided the ratio of frequency shift to bandwidth is maintained. Because of the other necessary considerations, the smallest permissible shift is to be preferred. With shaped keying, the sidebands at 120 cps from the center frequency are some 40 db below the unmodulated carrier. Without shaping the output spectrum the sidebands at 120 cps from the center frequency are 32 db below the unmodulated carrier. Since an efficient low-frequency radiator may have half-power bandwidth as small as 100 cps, the sideband power at 120 cps from the carrier may be reduced by a further 8 db. The bandwidths and frequency shifts suggested would permit 250-cps spacing of channel assignments, thereby making available 480 channels between 80 and 200 kc, and resulting in a very large increase over the number of channels now available.

After deciding on the optimum parameters of the frequency shift modulation system, as determined by the limiting and compromising factors, it is then necessary

³⁵ A. D. Watt, R. M. Coon, and V. J. Zurick, "Reduction of Adjacent-Channel Interference from On-Off and Frequency Shift Keyed Carriers," Natl. Bur. of Standards, Boulder, Colo, Rep. No. 5543; December 17, 1957.

³⁶ J. R. Davey and A. L. Matte, "Frequency shift telegraphy-radio and wire application," *Trans. AIEE*, vol. 66, pp. 479-493; 1947.

³⁷ J. S. Belrose, "Equipments for Use in a Narrow-Shift FSK Teleprinter System at Low Radio Frequencies," Defence Res. Telecommun. Establ., Ottawa, Ont., Can., DRTE/RPL Rep. No. 10-0-2; September 6, 1953.

³⁸ A. D. Watt, R. M. Coon, E. L. Maxwell, and R. W. Plush, "Performance of some radio systems on the presence of thermal and atmospheric noise," *PROC. IRE*, vol. 45, pp. 1914-1923; December, 1958.

³⁹ J. P. Costas, "Phase shift radio teletype," *PROC. IRE*, vol. 45, pp. 16-20; January, 1957.

to consider the required SNR necessary at the receiver to obtain the standard of operation specified for the system. The required SNR is determined by the characteristics of the atmospheric noise, the permissible error rate, and the characteristics of the receiving equipment. Watt, Coon, Maxwell, and Plush³⁸ have determined the SNR required for operation of a 60-wpm FSK system with a total shift of 50 cps, an IF bandwidth of 120 cps, and a post discriminator bandwidth of 70 cps, in the presence of atmospheric noise. Their results are shown in Table II.

TABLE II
SIGNAL-TO-NOISE IN DB AS A FUNCTION OF PERCENTAGE CHARACTER ERRORS FOR A NARROW-BAND FSK SYSTEM

Percentage Character Errors	RMS Carrier-to-Noise in a 1-KC Bandwidth (to 3-DB Points) in DB
10 per cent	-2
1 per cent	8
0.1 per cent	15

The numbers of Table II were obtained from tests made at Boulder, Colo., at a frequency of 22 kc using nonsynchronous stop-start teletype. The results are applicable only to the system tested under the given conditions but serve as a reasonable guide for determining the required SNR for narrow-band FSK at low radio frequencies. The dependence of the noise characteristics on geographical location, frequency, and time requires that the characteristics be determined for the actual location if an accurate estimate of the required SNR is to be determined. The fluctuation of the noise and of the signal about their median values, discussed in earlier sections of this paper, will cause a resultant variation of the percentage errors, as a function of time, about a median value. If the character error rate is specified not to exceed a given value, then the signal power must be increased by many db over the nonfading case.⁴⁰ In order to calculate the required power increase, allowances must be made for the statistical variation of both the signal and the noise, and the possible correlation between the two fluctuating amplitudes. Complete statistics are not available, and consequently it is difficult to effect realistic allowances. In practice, an estimate will have to be made based on median signal and noise levels.

Considerations of the characteristics of atmospheric noise affect the receiver design. The impulsive characteristic of the noise warrants the use of an RF bandwidth somewhat greater than optimum, with a form of limiter to suppress the amplitude of short duration pulses before the bandwidth is narrowed. Ideally, the receiver should consist of a comparatively broad-band amplifier with a limiter, a narrow-band amplifier, a discriminator, and a low-pass filter. From consideration

of operation in the presence of atmospheric noise, and requirements for adequate adjacent channel protection, an RF bandwidth of 200 cps is found suitable for the comparatively broad-band amplifier. The narrow-band amplifier should have a bandwidth of about 70 cps. The low-pass filter should have a cut off at 80 cps for non-synchronous teletype, and 40 cps for synchronous teletype. In practice, the benefit derived from reducing the prediscriminator bandwidth from 200 cps is small if the bandwidth is limited after the discriminator by the use of a low-pass filter.

COMMUNICATION EQUIPMENT AND FIELD TRIALS

After studying the propagation of low-frequency waves, the characteristics of radio noise, the design of transmitting and receiving antennas, and the relevant modulation systems, it was considered necessary to carry out field trials of experimental communication systems designed to test the validity of the various conclusions reached. The main obstacle to this work was the complete lack of suitable commercial communication equipment. With few exceptions, the receivers currently available for use in the low-frequency band were designed primarily for operation at high frequencies, and have neither the frequency stability nor the bandwidth characteristics essential for efficient use at low frequencies. Furthermore, the frequency shift keyers available were not sufficiently stable in frequency, nor was adequate provision made to permit the desired narrow shift operation. Some difficulty was also experienced with available types of frequency shift code converters.

The first phase of the field trials was designed primarily to investigate the possible use of semiportable low-power equipment for communication at low radio frequencies in the Canadian Arctic. The equipment was designed and built at the Radio Physics Laboratory. The transmitting installation consisted of a 50-watt CW transmitter and a sectionalized umbrella-type top-loaded antenna of 70-foot height. The receiving installation consisted of a loop antenna, loop preamplifier, and receiver of 100-cps bandwidth. The system was tested over several transmission distances to investigate the range dependence on transmitter power. In practice, the transmitter power was reduced to a level below which signal reception became unintelligible. Of some interest are the low orders of transmitter power found sufficient, on occasion, to communicate over ranges up to 1000 km. As an example, during November, 1952, with a 153-kc transmitter located at The Pas, Man., 685 km from the receiver at Churchill, the low atmospheric noise levels at Churchill permitted the reception at Churchill of signal intensities of 0.5 $\mu\text{v}/\text{m}$ during daytime and 1-2 $\mu\text{v}/\text{m}$ during nighttime. These values of signal intensity corresponded to radiated powers of 0.021 and 0.45 watt, respectively, or of transmitter output powers of 2.1 and 45 watts (1 per cent radiation efficiency obtained from the 70-foot antenna system).

The second phase of field trials was conducted

⁴⁰ F. E. Bond and H. F. Mayer, "The effect of fading on communication circuits subject to interference," *Proc. IRE*, vol. 45, pp. 636-642; May, 1957.

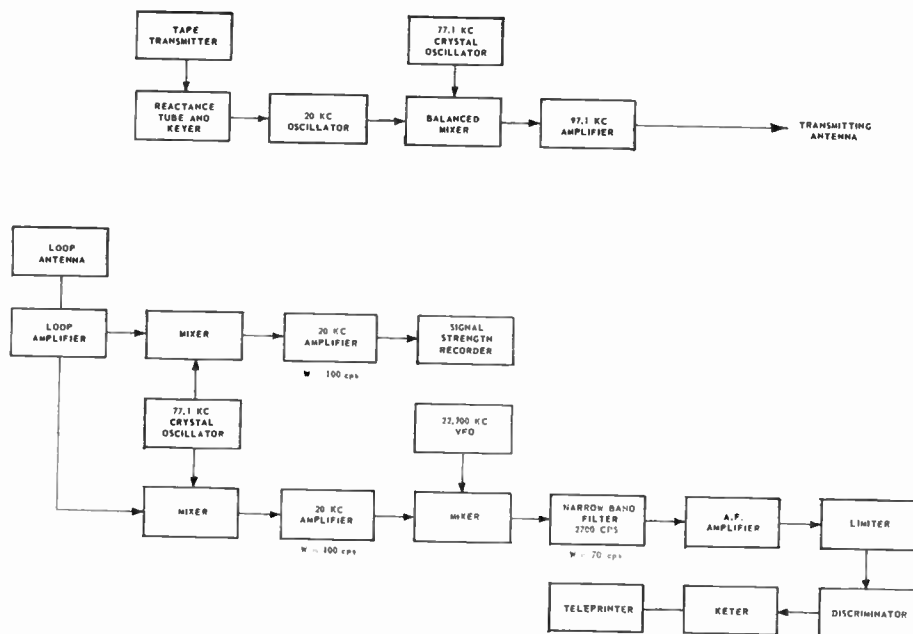


Fig. 21—Block diagram of the narrow-shift FSK teleprinter equipment used by the Radio Physics Laboratory in communications and propagation studies at low radio frequencies. (a) Transmitting terminal, (b) receiving terminal.

throughout 1954, and was designed to test the operation of narrow-band telegraphy communication using a total frequency shift of 40 cps, over distances up to about 2000 km. A block diagram of the equipment is shown in Fig. 21. The transmitting facilities at Ottawa consisted of a 250-foot umbrella-type top-loaded antenna system (previously described in this paper), a 5-kw transmitter, conventional teleprinter, and a frequency shift keyer designed and constructed at the Radio Physics Laboratory.³⁷ Receiving installations were located at Halifax, N. S., Goose Bay, Labrador, and Churchill, Man., at distances from Ottawa of approximately 965, 1415, and 1920 km, respectively. The receiving installations consisted of a loop antenna, loop preamplifier, receiver, frequency shift code converter, and teleprinter. Signal intensities were recorded continually with the transmitter being silenced at regular periods to obtain measurements of noise levels. The operating frequency for the trials was 97.1 kc,⁴¹ and the radiated power was 1000 watts. The trials were mainly concerned with tests of the modulation system under actual atmospheric noise conditions. The variability of the noise encountered at the three receiving sites, and the absence of precise information concerning the characteristics of the atmospheric noise, made it impractical to attempt any conclusive long-term assessment of the operational performance. However, the trials confirmed the usefulness of loop antennas at the low frequencies, verified the choice of various equipment techniques, and confirmed the practicability of communicating over long distances with FSK teletype at 60 wpm, using a frequency shift of the order of 40 cps.

Some of the conclusions of the Radio Physics Lab-

oratory studies have recently been applied by a group of commercial airlines operators in a system designed to transmit to aircraft in flight across the North Atlantic. From considerations of the increasing demands on existing HF radio telephone circuits for North Atlantic air routes, and the trend toward pilot-operated communications, it was decided to investigate alternative means of providing meteorological and other advisory information in a form that could be reproduced automatically in the aircraft. A system is now in test operation by the commercial airlines using transmitters located at Galde- noch, Scotland, Chatham, N. B., and Halifax, N. S., at operating frequencies of 121.6, 118.8, and 115.3 kc, respectively. A standard 60-wpm, 5-unit stop-start teleprinter signaling is transmitted, using a total frequency shift of 40 cps, to narrow-band receivers and teleprinters located in the aircraft. In the preliminary tests of the system, equipment was fitted in two aircraft of BOAC and one of TCA, and a total of over 50 flights were carried out prior to October 15, 1956. In the absence of suitable commercial equipment the airlines found it necessary to develop the frequency shift keyers, aircraft receivers, and various other items of the terminal equipment. It is reported⁴² that trials to date have been very promising, nighttime skywave coverage being fairly consistent to ranges of approximately 1000 nautical miles, and daytime ground-wave coverage between 700 and 900 nautical miles. A receiving equipment under development for such applications has been described by Bickers.⁴³

⁴² H. A. Ferris and B. G. Doutré, Trans-Canada Airline, private communication.

⁴³ A. Bickers, "An experimental airborne teleprinter service for North Atlantic airlines," *Marconi Rev.*, vol. 20, pp. 104-11; October, 1957.

⁴¹ The choice of this frequency is discussed later.

SUMMARY OF SYSTEMS DESIGN

The various factors that influence the design of a radio communication system at low frequencies have been discussed separately in the foregoing sections. An attempt has also been made to indicate optimum design parameters where applicable. The design of a reliable and efficient system can be achieved in practice only by consideration of all the various factors and their interrelationship.

The design of a system may proceed as follows:

1) Determine the frequency for optimum SNR at the receiver. This is dependent on the variation with frequency of the radiated power from the antenna system, of the propagation attenuations, and of the atmospheric noise level at the receiver. From considerations of the geographical location of the terminals, determine the various propagation modes and their amplitudes as a function of frequency. The choice of the best operating frequency can then be decided. This part of the design is usually the most difficult because it is often necessary to effect compromises. Consider, for example, the following cases:

a) $E_0 > E_s$ (daytime): The frequency for optimum SNR during the daytime is estimated after taking into account the variation with frequency of the ground-wave field intensity, the atmospheric noise level, and the radiation efficiency of a practical antenna system. The choice of frequencies at which there is a possibility that the nighttime skywave may have an amplitude comparable to that of the ground wave should be avoided because of possible phase interference effects. However, at frequencies below about 100 kc it is sometimes possible to choose a frequency such that the mean phase of the skywave at night is the same as the phase of the ground wave. The variability of the nighttime reflection height discourages the latter design approach for frequencies between 100 and 200 kc. Under the circumstances, it is then necessary to choose a frequency that will provide the best single mode propagation.

b) $E_s > E_0$: In this case, the frequency to provide optimum SNR at the receiver is estimated by taking into account the variation with frequency of the skywave, the atmospheric noise, and the radiation efficiency of a practical antenna system. In general, these considerations will suggest the use of a frequency below 100 kc for propagation during a summer day, because of the rapid decrease of skywave field intensity with increasing frequency. Considerations of the higher nighttime amplitudes of the skywave, and the higher amplitudes during a winter day, will suggest the choice of a higher operating frequency to take advantage of the improved radiation efficiencies and the lower noise levels at the high frequencies. In order to select a single operating frequency for daytime and nighttime use, a compromise is required.

2) Select suitable sites for the transmitting and receiving antennas. Because the conductivity of the ground in the vicinity of the transmitting and receiving antenna is of importance, the antenna sites must be chosen carefully. Care also must be taken to locate the receiving antenna in an area free from man-made noise.

3) Determine the radiated power necessary to provide the specified performance.

4) Determine transmitter power and transmitting antenna from considerations of cost, antenna efficiencies, and available transmitters.

The lack of information on the general behavior of atmospheric noise, and the requirement for information pertaining specifically to geographical locations of the circuit terminals, makes the design task very difficult and, in fact, makes it always advisable to carry out test measurements over the proposed circuit as a preliminary stage of the design. Although such a preliminary test program adds to the design costs and time delays, the information derived may effect a considerable saving in the eventual system costs.

Design Example

As a sample design calculation, consider the possible use of a low-frequency communication circuit between Ottawa, Ont., and Goose Bay, Lab., a distance of 1420 km.

The ground conductivity in the vicinity of Goose Bay is about 5×10^{-4} mhos/m, and at Ottawa, about 2×10^{-3} mhos/m. The average effective value over the path is expected to be less than 1×10^{-3} mhos/m. The ground-wave field intensity estimated from the CCIR graphs,⁴ and the skywave field intensity calculated from (9) are shown as a function of frequency in Fig. 22. From inspection of this figure, it is seen that the received signal will be due predominantly to the skywave. With the assumption of a variation of atmospheric noise levels proportional to $1/f^2$, and from consideration of the variation of transmitting antenna efficiency as a function of frequency, the frequency for optimum SNR is estimated to be about 120 kc for a March day and night, and below 100 kc for a summer day.

From the above conclusions and from considerations of available frequencies for allocation, a frequency of 97.1 kc was chosen for the test purposes discussed previously in this paper. The estimated values of field intensity for a radiated power of 1 kw are $4 \mu\text{v/m}$ for a summer day, $11 \mu\text{v/m}$ for a March day, and $50 \mu\text{v/m}$ for nighttime. Measured values were 3, 11, and $35 \mu\text{v/m}$, respectively. The diurnal variation of the field intensity in summer suggests a ground-wave field intensity smaller than that predicted in Fig. 22. This probably results from an overestimation of the effective ground conductivity along the path.

Consideration of the seasonal variation of atmospheric noise and of ionospheric reflection coefficients indicates that the transmitter power requirements will be a maximum during the summer. In Table III, the estimated

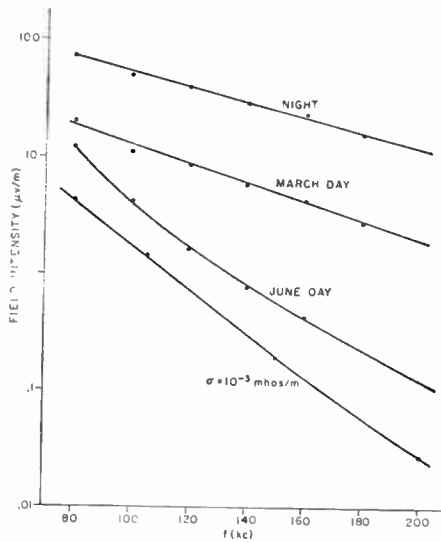


Fig. 22—Variation with frequency of the calculated field intensity for 1 kw radiated over the Ottawa-Goose Bay path. The ground-wave calculation assumes an effective conductivity of 10^{-3} mhos/m, and skywave values are for nighttime, and equinox and summer daytime propagation conditions.

values of the rms noise levels are shown for the months of June, July, and August, at Ottawa and Goose Bay, for a frequency of 100 kc.⁴⁴ The increase in the noise levels from day to night is compensated by the corresponding increase of about 20 db in signal field intensity.

Assuming the use of a 250-foot umbrella-type top-loaded radiator and summer daytime propagation as indicated by the measurements, the power required for 60-wpm teletype transmission using a total frequency shift of 40 cps and optimum receiver bandwidths is shown in Table IV, for the two transmitting terminals, and as a function of percentage error.

It is evident from Table IV that communication from Goose Bay to Ottawa at low radio frequencies can be maintained during summer daytime only by the use of a high-powered transmitter at Goose Bay, unless repetition of traffic can be tolerated in the interests of improved reliability. For example, a good arrangement would be to use some form of closed loop error correcting system. With a time division duplex transmission system at both ends of the circuit, one channel could be used for error correcting. Reduced rates during times

⁴⁴ These values have been deduced from NBS Circular 557, and represent the rms field strength for a 1-kc bandwidth above $1\mu\text{V}/\text{m}$ where the bandwidth is the effective noise bandwidth. The effective noise bandwidth is the bandwidth of a sharp cutoff system that would give rise to the same amount of noise power as the actual system. In Table II we quoted values for the rms carrier to rms noise ratio, in a 1-kc 3-db bandwidth, to yield a specific grade of performance. The relation between the effective noise bandwidth and the 3-db bandwidth depend on the shape of the receiver IF response. For a receiver which consists of a single-tuned circuit the effective noise bandwidth is approximately the same as the 6-db bandwidth, and the ratio between the 6-db bandwidth and the 3-db bandwidth is 1.6. A practical receiver for use at low frequencies would consist of several cascaded double-tuned band-pass coupled circuits. The effective noise bandwidth in this case would more closely approximate the 3-db bandwidth than the 6-db bandwidth. In the calculations to follow, the noise intensity values of Table III are taken to represent those that would be measured in a 3-db bandwidth.

TABLE III
MEDIAN RMS NOISE LEVELS AT 100 KC IN DB ABOVE $1\mu\text{V}/\text{M}$ FOR A 1-KC BANDWIDTH

	Day	Night
Ottawa	13	33
Goose Bay	4	16

TABLE IV
POWERS REQUIRED FOR COMMUNICATION OVER THE OTTAWA-GOOSE BAY PATH AT 100 KC

Per Cent Error	Radiated Power (KW)		Transmitter Power (KW)	
	Ottawa to Goose Bay	Goose Bay to Ottawa	Ottawa to Goose Bay	Goose Bay to Ottawa
10	0.2	1.4	0.7	5.5
1	1.8	14	7.1	56
0.1	8.6	69.5	34	280

of poor signal-to-noise conditions are not unreasonable since the main application of low-frequency circuits is to provide support to other radio systems. Normally, assessment would involve an examination of the operational requirement to determine the minimum permissible information rate, or alternative locations for the Ottawa receiving terminal.

The calculations for this particular circuit clearly illustrate two important points made in this paper:

1) The relatively smaller transmitter powers required for satisfactory signal reception in the Arctic compared with power requirements for reliable signal reception at lower latitudes;

2) The great reduction in the radiated and received skywave field intensities due to the finite conductivity near the terminal antennas. For the circuit discussed, the field intensities are some 14–20 db less than those which would be obtained if the antennas were located on a perfectly conducting earth, or if propagation for the first few wavelengths in the direction of the receiving (or transmitting) terminal were over sea water.⁴⁵

ACKNOWLEDGMENT

The authors are indebted to a number of their colleagues at the Radio Physics Laboratory for assistance in field trials. Specific mention should be made of T. W. Straker, who was in charge of the low-frequency group when it was first formed, and who always showed great enthusiasm for this field of endeavor. The work could not have been done without the complete support and encouragement of F. T. Davies, former Superintendent of the Defence Research Telecommunications Establishment, and of his successor, J. C. W. Scott.

⁴⁵ Since the preparation of the original manuscript, a report has been published which shows very clearly the effects of very poor ground conductivities on low-frequency propagation in Arctic areas. See A. D. Watt, E. L. Maxwell, and E. H. Whelan, "Observations on Some Low Frequency Propagation Paths in Arctic Areas," Natl. Bur. of Standards, Boulder, Colo., Rep. No. 5574; May, 1958. To be published in *J. Res. NBS*, pt. D; July, 1959.

The Development of Radio Traffic Frequency Prediction Techniques for Use at High Latitudes*

O. A. SANDOZ†, SENIOR MEMBER, IRE, E. E. STEVENS†, AND E. S. WARREN†

Summary—A brief historical review is presented of ionospheric work leading to the development of radio prediction techniques in order to provide a background and to emphasize the fact that all prediction methods are relatively quite "young." A bibliography of Canadian work is given. The merits and defects of some prediction techniques are discussed with particular attention to the Canadian problem in the far north. The results of recent research are shown to hold promise in the field of predictions. The paper concludes with some speculation on possible trends in both long- and short-term prediction methods.

RADIO WAVES have been used for more than fifty years for long distance communication. It is natural, perhaps, to assume that the methods used for predicting the most appropriate frequencies for the purpose are sound as a result of tests and modification over many years. Therefore, it is with some surprise that newcomers to the field of radio communication discover that some of the ideas basic to the U. S. Bureau of Standards method were conceived as late as 1939, by Newbern Smith, and that the corresponding concepts for the British methods were developed in 1940 by Appleton and Beynon. There is value, therefore, in reviewing the structure of the prediction systems in order to make clear that each of them contains techniques that have definite need of improvement. A secondary aim is to provide a reminder of the largely empirical nature of so many of the working assumptions that are often accepted as established fact, even by experienced communication engineers. Further, by reference to recent work, a promise is indicated that Arctic predictions may yet be made sufficiently simple and accurate to gain general acceptance.

This paper deals primarily with the prediction problems of northern latitudes. These problems are a specialized outgrowth from the task of predicting radio traffic conditions on a world-wide scale and, therefore, many of the comments are relevant for all latitudes.

HISTORICAL NOTES

The first approach to the modern concept of our upper atmosphere dates back to 1882 when Stewart¹ proposed that there is a conducting region at the outer fringe of our atmosphere. When Marconi achieved successful radio communication across the Atlantic in 1901, con-

siderable interest was stimulated and discussion arose throughout the scientific world as to the means by which electromagnetic waves could be propagated for so great a distance around a curved surface.

Kennelly in the United States, and Heavyside in England, in 1902, predicted independently the presence of a conducting layer in the upper atmosphere. They proposed that radio signals were reflected back to earth by this layer and, for this reason, followed the curvature of the earth. Eccles,² in 1912, presented his theory of wave propagation through an ionized medium. This theory was supplemented by Larmour³ in 1924, and today the Eccles-Larmour theory is considered the basic theory for wave propagation in the ionosphere.

The use of radio communications expanded greatly. Radio amateurs, forced out of the wavelengths above 200 meters, pioneered the way into the use of higher frequencies shortly after the end of World War I. Observations of certain characteristics in the propagation of both long and short waves gave additional support to the theory that a reflecting layer must exist. However, it was not until 1925 that Appleton's experiments⁴ confirmed the presence of not only the Kennelly-Heavyside layer, but of a still higher layer above. This was called the Appleton Layer. Today we know the Kennelly-Heavyside layer as the E Layer and the Appleton layer as the F layer.

In 1926, Breit and Tuve,⁵ employing a new technique, were able to show direct evidence of the ionospheric layers. They transmitted pulses of extremely short duration, and measured the delay time between the reception of the ground wave or "direct wave" and the sky wave or "indirect wave." This technique was soon adopted by other workers in the field and today is still the most useful method for investigating the properties of the ionospheric regions. As a matter of interest, observations and equipment techniques used in early ionospheric soundings contributed directly to the discovery and development of radar.

² W. H. Eccles, "On the diurnal variations of the electric waves occurring in nature and on the propagation of electric waves round the bend of the earth," *Proc. Royal Soc.*, part A, pp. 79-99; August, 1912.

³ J. Larmour, "Why wireless electric rays can bend round the earth," *Phil. Mag.*, vol. 48, pp. 1025-1036; December, 1924.

⁴ E. V. Appleton and M. A. F. Barnett, "Local reflection of wireless waves from the upper atmosphere," *Nature*, vol. 115, pp. 333-334; March, 1925.

⁵ G. Breit, and M. A. Tuve, "A test of the existence of the conducting layer," *Phys. Rev.*, vol. 28, pp. 554-573; September, 1926.

* Original manuscript received by the IRE, January 26, 1959.

† Defence Res. Telecommun. Est., Ottawa, Ont., Can.

¹ B. Stewart, "Terrestrial Magnetism," *Encyclopaedia Britannica*, vol. 21, pp. 965-966; 1957.

Terman has said,⁶ "The literature giving experimentally determined characteristics of the ionosphere is almost limitless." The same could be said with almost equal weight of theoretical studies. In this paper, therefore, no attempt has been made to provide more than representative references to indicate the course of developments.⁷

By 1930, a considerable amount of ionospheric data was available from scientific measurements and from performance figures on different communication circuits. Attempts were made to provide communication organizations with predictions of optimum working frequencies over different transmission paths based on the time of day and season. In that year Prescott⁸ provided 1) a source of HF propagation data for treating long distance circuits, and 2) data that would show the diurnal and seasonal performance of HF transmissions.

The advent of sweep frequency vertical incidence ionospheric recording together with the photographic reproduction of records meant that the fine structure of the ionosphere could be studied in detail. Gilliland⁹ summarized the results of measurements taken by the National Bureau of Standards, at Washington, D. C., during 1933-1934. He pointed out the diurnal and seasonal variations, their practical application to communication problems, and the possibility of a long-term trend connected with the sunspot cycle.

Hulburt,¹⁰ in 1935, drew a relationship between the sunspot cycle and skip distances. He indicated in the same paper that the ionization of the ionospheric regions might increase by fifty to one hundred per cent from the minimum to maximum of the sunspot cycle.

There were still insufficient ionospheric data to permit accurate frequency predictions to be made on a world-wide basis. However, significant trends in HF propagation conditions were being noted. This was apparent from an article published by Young and Hulburt¹¹ in 1936. They pointed out in this paper that useful long distance communications appeared to agree more closely with the yearly average sunspot number than with the magnetic character in solar constant during the period from 1923 to 1936.

Additional information on the diurnal, seasonal, and long-term trends of HF propagation appeared the following year when an article on the subject was published

by Gilliland, Kirby, Smith, and Reymer,¹² of the National Bureau of Standards.

Publication of ionospheric information from the National Bureau of Standards, Washington, D. C., in the form of monthly bulletins, began in September, 1937, in the PROCEEDINGS OF THE IRE. These contained the hourly mean critical frequencies for the month, maximum usable frequencies for different path lengths, and ionospheric storm and disturbance observation charts. These bulletins continued until January, 1942.

The development of the transmission slider by Smith¹³ of NBS was an important step forward in the practical application of ionospheric measurements to HF communications. This slider permitted a simple method to be used in relating critical frequencies to maximum usable frequencies for different path lengths.

It was soon apparent, during the early years of World War II, that sufficient radio propagation information was just not available to meet the requirements of the Armed Forces fighting a war on a world-wide basis. Data in all cases were incomplete and, in some instances, nonexistent. The recognized need for efficient radio communications led to the establishment, in 1941, of the British Inter-Services Ionospheric Bureau (ISIB) and the Australian Radio Propagation Committee (ARPC). In the United States the Inter-Service Radio Propagation Laboratory (IRPL) was established in 1942 in the National Bureau of Standards by order of the Joint Chiefs of Staff. The main function of these organizations was to supply radio propagation information to the respective Armed Forces, but the scientists who staffed them retained a research interest.

Ionospheric observations on a regular basis were only available at this time from Slough, England, Huancayo, Peru, Washington, D. C., and Watheroo, Australia. More comprehensive data had to be obtained. An effort was made to obtain world-wide coverage of ionospheric information and by the end of the war regular data were being obtained from forty-four stations as a result of cooperation between Australia, Canada, the United Kingdom, the USA, and the USSR.

These organizations were soon able to produce world-wide frequency prediction charts. Bennington,¹⁴ of the Marconi Wireless Company, appears to have been responsible for the first of these. The first IRPL charts appeared in January, 1942, for the month of March of that year. As additional information became available, former anomalies were explained and a consistent world-wide pattern became apparent.

The "longitudinal" effect of the F2 Layer, although known for some time, was not corrected in prediction systems until 1943, when a division of the world into

⁶ F. E. Terman, "Radio Engineers' Handbook," McGraw-Hill Book Co., Inc., New York, N. Y., p. 720; 1943.

⁷ For a comprehensive bibliography refer to L. A. Manning's "A Survey of the Literature of the Ionosphere," Radio Propagation Lab., Stanford Univ., Stanford, Calif., final report prepared under contract AF19(604)-686 for AFCRC/ARDC; July 31, 1955.

⁸ M. L. Prescott, "Diurnal and seasonal performance of HF radio transmission over various paths," *Proc. IRE*, vol. 18, pp. 1797-1920; November, 1930.

⁹ T. R. Gilliland, "Multifrequency ionosphere recording and its significance," *Proc. IRE*, vol. 23, pp. 1076-1101; September, 1935.

¹⁰ E. O. Hulburt, "The ionosphere, skip distances of radio waves, and the propagation of microwaves," *Proc. IRE*, vol. 23, pp. 1492-1506; December, 1935.

¹¹ L. C. Young, and E. O. Hulburt, "Radio and the sunspot cycle," *Phys. Rev.*, vol. 50, pp. 45-47; July, 1936.

¹² T. R. Gilliland, S. S. Kirby, N. Smith, and S. E. Reymer, "Characteristics of the ionosphere and their application to radio transmission," *Proc. IRE*, vol. 25, pp. 823-840; June, 1937.

¹³ N. Smith, "The relation of radio skywave transmission to ionosphere measurements," *Proc. IRE*, vol. 27, pp. 332-347; May, 1939.

¹⁴ T. W. Bennington, "Wireless World," vol. 52, pp. 246-250 and pp. 292-295; August, 1946.

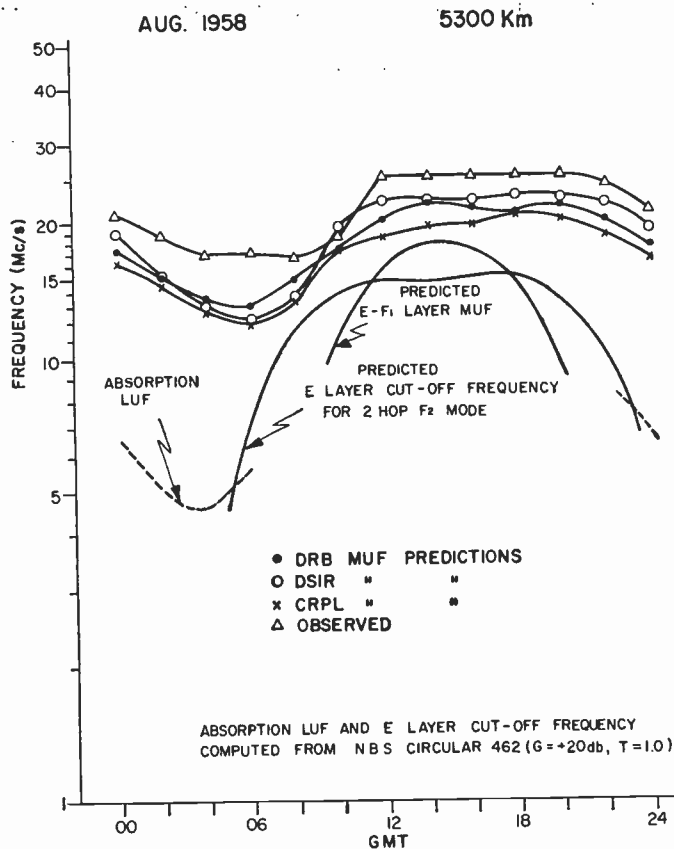


Fig. 1—A comparison of predicted and observed MUF on the Ottawa, Can., to Slough, England, circuit.

In particular, the one-hop F₂ high angle mode of propagation to great distances predicted by Dieminger¹⁹ and observed by DRTE deserves more study. It should be of particular interest in equatorial regions.

The methods used at present for calculating the obscuring effect of the lower layers are inadequate. The basis on which such calculations are made is that the LUF for the higher-layer mode is the MUF for the lower-layer mode at that distance for which the elevation angle is the same as for the higher-layer mode. This concept is unsound in that it ignores the role of the Pedersen ray in the obscuring of the higher layer. Such Pedersen rays are fairly common for E Layer propagation to distances of 1000 km or less. The assigning of a constant height of 320 km to the F region for the purpose of E Layer cutoff calculations is an even more doubtful approximation. If, in the future, a constant height must be used, a somewhat lower one, having a value in the vicinity of 250 km, should be chosen.

It has been assumed implicitly in prediction work that if a frequency is below the MUF and above the LUF, it will be usable. This assumption has proved incorrect for propagation on the North-Atlantic circuit, where a gap in the received spectrum commonly occurs between the lowest frequency propagated by the one hop F₂

¹⁹ W. Dieminger, H. G. Moller, "Echo sounding experiments with variable frequency at oblique incidence," *Nuovo Cimento Suppl.*, ser. 10, vol. 4, pp. 1532-45; 1956.

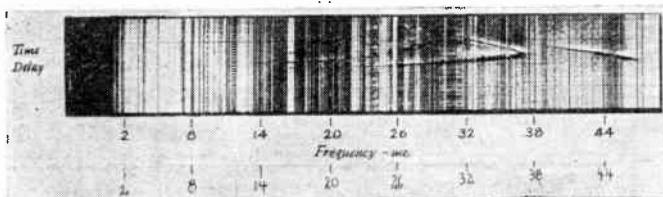


Fig. 2—Sweep-frequency oblique incidence sounding record on trans-Atlantic path, showing gap in usable frequencies around 38 mc. October 27, 1957, 1220 GMT.

mode and the highest frequency propagated by the two-hop F₂ mode. This is shown in Fig. 2, where the gap occurs in the vicinity of 38 mc.

SOME SPECULATION

Basic to any conventional prediction technique is some method for deducing the frequencies that will be propagated at oblique incidence by the ionosphere, from a knowledge of those returned to earth at vertical incidence. The relationship for present systems is calculated on the assumption that the radio wave is refracted in an ionosphere that has no horizontal gradient of electron density. The vertical incidence sounding records in Canadian latitudes give abundant evidence, however, that the layers are irregular, especially at night and particularly in the auroral zone. It is plausible that at oblique incidence, since the wave must travel through greater numbers of irregularities, the relative amount of power scattered from the incident wave would be greater. Oblique incidence sounders on circuits that pass close to the auroral zone rarely record a high angle ray between midnight and sunrise, the interval during which the ionosphere is most irregular as judged from the spread echoes²⁰ observed at vertical incidence. A typical spread echo record is shown in Fig. 3. The energy that would have revealed the high angle ray under more favorable circumstances has apparently become so diffused through scatter that it arrives at the receiver at a level below that of the noise. These sounders operate at a peak output power of 10 kw and a bandwidth of 30 kc. It is of interest to speculate whether, if a higher power-to-bandwidth ratio were used, a diffuse high angle one-hop F₂ mode might be observed on the North-Atlantic circuit at night. If so, a worthwhile extension of the spectrum might be achieved with but a modest increase in the power level. For the rough ionosphere of northern latitudes, the approximation that radio waves are propagated by refraction alone yields maximum frequencies that are on the average ten per cent or more below the observed values.

In addition to the regular changes correlated with time of day, season, and running average sunspot number, the ionospheric parameters undergo short-term fluctuations. These disturbances have a duration of

²⁰ E. L. Hagg, G. H. Hanson and D. Fowle, "The Interpretation of Ionospheric Records," DRTE/RPL Rep. No. R-2; May, 1953.

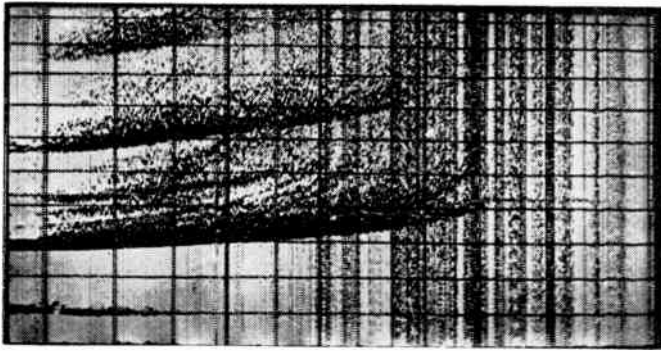


Fig. 3—Vertical incidence sounding record (h/f ionogram) showing spread echoes. Eureka; 80.0° N, 85.9° W; 0713 EST, November 8, 1957.

from a few minutes to several days and, on a long-term basis²¹ are unpredictable at present. Fig. 4 demonstrates the variability of the penetration frequency of the F Layer during a single month at Alert. The highest and lowest values for a given hour of the day are plotted, and the solid lines indicate the range of values occupied by fifty per cent of the total number of occurrences. Present techniques of prediction make use of the median value. It might be better to deal separately with quiet days and disturbed days. Conceivably the MUF for quiet days would be more readily predictable than the median MUF for all days of the month. Such a prediction would have to be supplemented by long-term average reliability figures for the purposes of equipment design and frequency allocation. The necessity for supplementing the long-term prediction service with day-to-day information on ionospheric conditions is generally recognized. A short-term forecast for disturbances on the North-Atlantic and North-Pacific circuits is now operated by CRPL.²² The prediction of disturbances several days in advance still has a large amount of uncertainty connected with it, although promising work is being done on this problem.²³

For the next few years at least, it appears that the best way to deal with disturbances is to make the most of what little HF spectrum may be usable when disturbed working conditions occur unexpectedly. Oblique incidence sounders could be installed on the longer and more important domestic and foreign circuits so that the operators can be kept informed of the changing frequencies and conditions on the circuit as they occur. It is on long circuits, where the band of suitable frequencies is often small, and for Arctic circuits, which are particularly vulnerable to ionospheric fluctuations, that the oblique sounder would have most value.

Backscatter sounding in azimuth and range at a few frequencies has been proposed by some users to provide guidance for the operation of their circuits. Such a

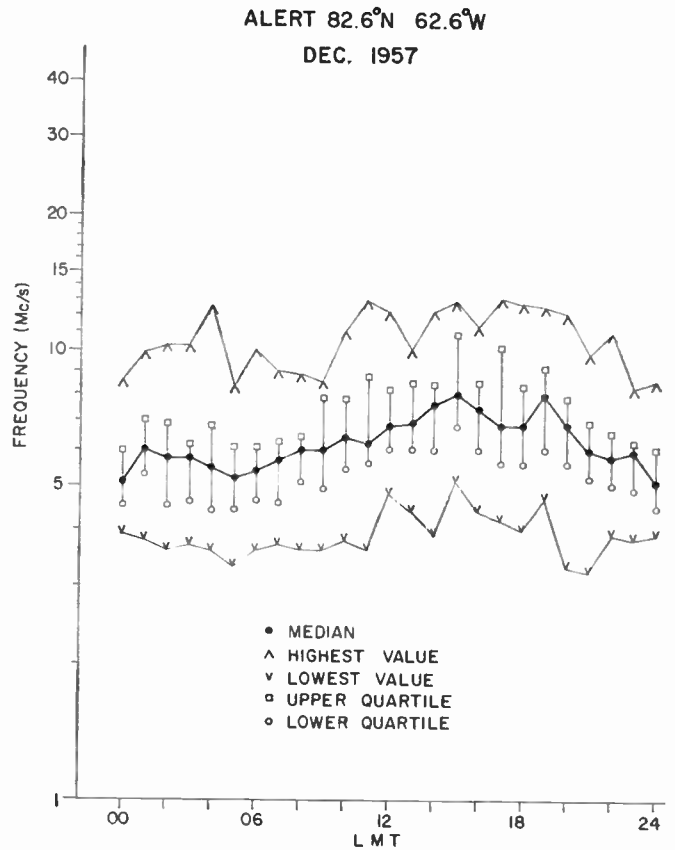


Fig. 4—Variability of the penetration frequency of the F Layer; Alert, NWT.

Communication Zone Indicator, generally called COZI equipment, gives directly only a small part of the complex propagational information required, yet it could be employed usefully to show the azimuth over which data from the two terminal oblique sounder might be extrapolated safely.

It might not be too speculative to anticipate that the pressure of congestion in the HF band, combined with the requirement to make rapid frequency changes due to the instability of ionospheric conditions, will lead to the establishment of prediction centers from which data from oblique sounders would be supplied, as is weather information now. This congestion may well increase to the point where, despite the obvious coordination difficulties, some form of frequency-time sharing will be forced upon us.

The limitations of prediction techniques are recognized generally by workers in the field, and several revisions are under way. Japan has already issued a set of F2 maps applicable to most of the world.²⁴ In Canada, the Defence Research Board has an active research program for the study of propagation in the high frequency band. A part of this program has consisted in increasing the number of vertical incidence stations, as shown in Fig. 5, operating in Canada from five to

²¹ J. H. Meek, "Polar disturbances," *J. Atmos. Terr. Phys. (Suppl.)* pp. 120-128; 1957.

²² "Sun-Earth Relationship and Radio Disturbance Forecasts," U.S. Nat. Bur. Stand., Boulder, Colo.

²³ T. R. Hartz, "The Prediction of geomagnetic and ionospheric disturbances," *Electronics and Commun.*, vol. 6, pp. 26-29; August, 1958.

²⁴ "World Maps of F2 Critical Frequencies and Maximum Usable Frequencies for 4000 KM," Radio Research Labs., Ministry of Post- and Telecommun., Tokyo, Japan; August, 1958.

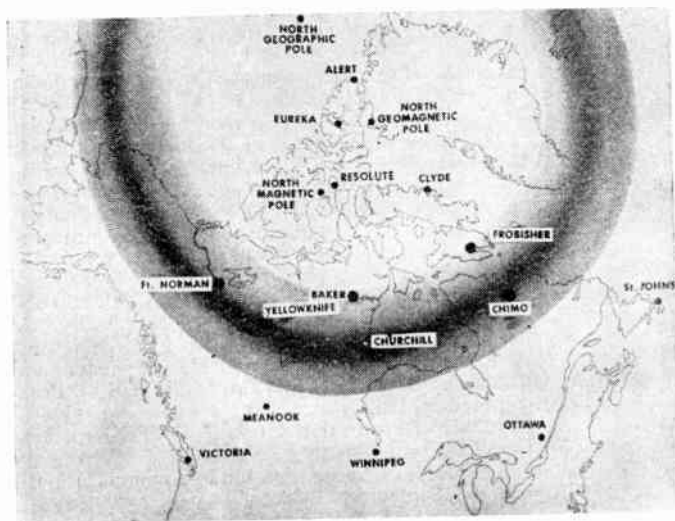


Fig. 5—Canadian vertical incidence sounding stations during the IGY.

fifteen during the IGY. The data from these stations²⁵ are expected to have considerable value for prediction purposes, and to yield a knowledge of the statistics, not only of penetration frequencies, but also of disturbances and blackouts. An oblique incidence sounder has been operating since September, 1957, on the Ottawa, Ont.—Slough, U.K., circuit, recording the propagation modes in the frequency range 3–49 mc. More recently, other oblique incidence circuits have commenced operation inside Canada. The absorption of 2 mc radio waves reflected from the E region has been measured hourly for several years at five Canadian stations. These many measurements provide not only the necessary basic information for predictions, but also appropriate tests of the principles involved.

A prediction system for northern latitudes could probably be constructed using new concepts that would yield medians of maximum and minimum frequencies with less than five per cent error. It is recognized that this accuracy is required only for certain circuits such as the North-Atlantic circuit in summer, when a ten per cent error would yield a prediction outside the range of usable frequencies that frequently lie in a band hardly a megacycle wide. A probable development is the devising of prediction systems of differing levels of complexity suitable to the length of the circuit and to the use to be made of the prediction.

A gradual improvement in the prediction of traffic frequencies at high latitudes is expected. Long-term “median” prediction will benefit from the inclusion of new data in systems now based on an inadequate survey of ionospheric conditions. The methods may also be expected to improve as greater attention is paid to the geometry of the propagation modes. The radio operators’ short-term problem will likely be reduced greatly by the installation of simplified oblique incidence sounders on important circuits.

²⁵ Published in the series, “Canadian Ionospheric Data,” with IGY Supplements.

V. ACKNOWLEDGMENT

The authors wish to acknowledge the contributions of their coworkers through discussion and suggestion, and to thank the Defence Research Board for permission to publish.

APPENDIX I

REPRESENTATIVE CRWPC PUBLICATIONS

- [1] Meek, J. H., “Instructions for Observers: Canadian Ionospheric Stations,” CRWPC/IP/1; July, 1944.
- [2] Meek, J. H., “The Application of Ionospheric Measurements,” CRWPC/IP/2; October, 1944.
- [3] Davies, F. T., “First Report on the Solar Eclipse Observations of July, 1945,” CRWPC/IP/6; September, 1945.
- [4] Davies, F. T., “Aurora at Churchill, Manitoba, 1943–44,” CRWPC/IP/3; February, 1945.
- [5] Davies, F. T., “Long Distance Transmission of 40–50 mc/s Band in Canada,” CRWPC/IP/5; August, 1945.
- [6] Scott, J. C. W., “Measurement of Signal Strength of WWV at Goose Bay,” RCAF Rep. No. ERD4-1; August, 1945.
- [7] Davies, F. T. and Scott, J. C. W., “Radio Frequency Prediction for Canada 1946–55,” CRWPC/IP/11; March, 1946.
- [8] Meek, J. H., and Scott, J. C. W., “Graphical Method for Obtaining Times of Sunrise and Sunset,” CRWPC Information Bulletin, No. 2; May 29, 1946.
- [9] Meek, J. H., “Graphical Method for Obtaining the Elevation of the Sun,” CRWPC Information Bulletin No. 8; November 15, 1946.
- [10] Meek, J. H., “Sporadic E. Ionization at Churchill, August 1943–July 1946,” CRWPC/IP/14; November 26, 1946.
- [11] Scott, J. C. W., “Radio Noise Measurements in Canada,” CRWPC Information Bulletin No. 3; June 25, 1946.
- [12] Scott, J. C. W., “Seasonal Variation in WWV Reception at St. John’s,” JCC/WP/IP/13; July, 1946.
- [13] Scott, J. C. W., “Practical Radio Frequency Prediction Charts,” CRWPC Information Bulletin, No. 4; July 5, 1946.
- [14] Davies, F. T., “Sunspots and Radio Disturbance,” CRWPC Information Bulletin, No. 11; June 5, 1947.

APPENDIX II

REPRESENTATIVE DRTE PUBLICATIONS

- [1]† Scott, J. C. W., “Magneto-ionic effects at Clyde River,” URSI-IRE, Washington, D. C.; October, 1947.
- [2]† Meek, J. H., “The variation of radio wave attenuation across the auroral zone,” presented at DRB Symposium, Ottawa, Ontario, Can.; December, 1948.
- [3] Scott, J. C. W., “The Auroral Zone and Frequency Assignment,” ITU Provisional Frequency Board Document No. 130; March 1, 1948.
- [4]† Scott, J. C. W., “Critical frequency difference variations and the Poynting vector in the ionosphere,” *Nature*, vol. 163, p. 993; June, 1949.
- [5] Scott, J. C. W., “Nomograms for ionosphere control points,” *PROC. IRE*, vol. 37, pp. 821–824; July, 1949.
- [6]† Meek, J. H., “Occurrence of E Region Sporadic Ionization Through the Auroral Zone,” DRTE/RPL Rept. No. 3; January 24, 1950.
- [7]† Meek, J. H., “Attenuation of Radio Waves in the Auroral Zone,” DRTE/RPL Rep. No. 5; August 16, 1950.
- [8] Scott, J. C. W., “Longitudinal and transverse propagation in Canada,” *J. Geophys. Res.*, vol. 55, pp. 267–269; September, 1950.
- [9] Scott, J. C. W., “The Poynting vector in the ionosphere,” *PROC. IRE*, vol. 38, pp. 1057–1068; September, 1950.
- [10] McKerrow, C. A., and Meek, J. H., “Ionospheric Observer’s Instruction Manual,” DRTE/RPL; February 15, 1951.
- [11] Meek, J. H., “Reception of 2 mc/s Loran in Central Canada,” DRTE/RPL Rep. No. 9; July, 1951.
- [12]† Hagg, E. L., and Hanson, G. H., “Recommendations to the Department of Mines and Technical Surveys Regarding Radio Communication with Survey Parties in 1952,” DRTE/RPL Rep. 1–2; February 25, 1952.
- [13] Meek, J. H., “Oblique reflection of radio waves by way of triangular path,” *Nature*, vol. 169, p. 327; February 23, 1952.
- [14] Meek, J. H., “Ionospheric disturbances in Canada,” *J. Geophys. Res.*, vol. 57, pp. 177–190; June, 1952.
- [15] Meek, J. H., “Correlation of Magnetic, Auroral and Ionospheric Variations at Saskatoon,” a) M. A. Thesis, University of Saskatchewan, Saskatchewan, Can.; November, 1952. b) *J. Geophys. Res.*, vol. 58, pp. 445–456; December, 1953. c) Part II, *J. Geophys. Res.*, vol. 59, pp. 87–92; March, 1954.

- [16] Scott, J. C. W., "The Solar control of the E and F layers at high latitudes," *J. Geophys. Res.*, vol. 57, pp. 369-386; September, 1952.
- [17]‡ Hagg, E. L., Hanson, G. H., and Scott, J. C. W., "An Arctic Prediction Service," DRTE/RPL Proj. Rep. No. 1-1-1; March 17, 1953.
- [18] Campbell, W., Hanson, G. H., and Serson, H. V., "Maximum usable frequencies and lowest usable frequencies for the path Washington to Resolute Bay," *J. Geophys. Res.*, vol. 58, pp. 487-491; December, 1953.
- [19] Davies, K., and Hagg, E. L., "Optimum Working Frequencies for Use on the Circuit Ottawa-Halifax," DRTE/RPL Project Rep. 1-3-5; November 23, 1953.
- [20] Farmer, J. C., and Hanson, G. H., "Seasonal and Solar Cycle Variation in the E Layer," DRTE/RPL Project Rep. 1-2-1; September 10, 1953.
- [21] Scott, J. C. W., "The distribution of F2 region ionization at high latitudes," *J. Atmos. and Terr. Phys.*, vol. 3, no. 6, pp. 289-291; 1953.
- [22] Hagg, E. L., Hanson, G. H., and Fowle, D., "The Interpretation of Ionospheric Records," DRTE/RPL Rep. No. R-2; May, 1953.
- [23] Hurtubise, J. C., "A Portable Ionosphere Recorder," DRTE/RPL Proj. Rep. No. 3-4-1; February, 1954.
- [24] Cox, J. W., and Davies, K., "Statistical study of polar radio blackouts," *Can. J. Phys.*, vol. 32, pp. 743-756; December, 1954.
- [25] Davies, K., and Hagg, E. L., "Reception of WWW Aboard H.M.C.S. Labrador," DRTE/RPL Project Rep. No. 3-6-2; January, 1955.
- [26] Davies, K., and Jelly, Doris, "An Ionospheric Index for Prediction Purposes," DRTE/RPL Proj. Rep. No. 1-1-4; March, 1955.
- [27] Davies, K., "A Study of the Reception of WWW at Baker Lake and Churchill," DRTE/RPL Proj. Rep. No. 1-3-7; March, 1955.
- [28] "Predictions of Optimum Traffic Frequencies for Northern Latitudes," Parts I and II, DRTE/RPL Rep. No. 1-1-3; 1954.
- [29] Cox, J. W., and Davies, K., "Oblique-incidence pulse transmission over a 2360 km path via the ionosphere," *Wireless Engineer*, vol. 32, pp. 35-41; February, 1955.
- [30] Davies, K., and Hagg, E. L., "Ionospheric absorption measurements at Prince Rupert," *J. Atmos. and Terr. Phys.*, vol. 6, pp. 18-32; January, 1955.
- [31] Chapman, J. H., Davies, K., and Littlewood, C. A., "Radio observations of the ionosphere at oblique incidence," *Can. J. Phys.*, vol. 33, pp. 713-722; December, 1955.
- [32] Peebles, P. J. E., "Ionospheric Absorption Measurements on a Frequency of 2 mc at Baker Lake and Churchill," DRTE Proj. Rep. 3-6-3; September, 1956.
- [33] Hartz, T. R., "Solar-Terrestrial Relations," DRTE Rep. No. 23-2-3; February, 1958.
- [34] Hatton, W. L., "Radio Propagation and Communication," DRTE Rep. 40-0-8; September, 1958.
- [35] Hagg, E. L., and Warren, E. S., "Single-hop propagation of radio waves to a distance of 5300 Km." *Nature*, (letter), vol. 181, pp. 34-35; January 4, 1958.
- [36] Hartz, T. R., "The Prediction of geomagnetic and ionospheric disturbances," *Electronics and Commun.*, vol. 6, pp. 26-29; August, 1958.

Tropospheric Scatter System Using Angle Diversity*

J. H. VOGELMAN†, J. L. RYERSON†, AND M. H. BICKELHAUPT†

Summary—This paper proposes an extension of the angle diversity technique as a means towards solution of several problems restricting use of over-the-horizon microwave communications. The paper suggests application of a microwave multibeam system as an attack on the problems of transmitter tube limitations, "medium to aperture" coupling loss, and UHF-band interference. A simple experiment is described and results evaluated. The experiment provides a measure of verification of previous theoretical work on which the proposed technique is based. Results of numerous path calculations are included to provide the reader with an estimate of improvement to be gained by the use of this expansion of the angle diversity technique over a conventional system.

INTRODUCTION

IN an attempt to get long-range tropospheric communications, the United States Air Force has been forced to go to lower and lower frequencies and larger and larger antennas to achieve the necessary power and signal strength to permit reliable communications. The resulting equipments have stretched the state of the transmitter tube art to the maximum, and require antennas as large as 120 feet in diameter. The use of the lower UHF region has resulted in extreme crowding of this frequency band because of the many

different applications of this region in communication, both military and civil. In order to obtain new frequencies for use in tropospheric scatter, it is essential to move into the microwave-frequency region. Various investigators¹ have shown that the use of large aperture antennas at super-high frequencies, resulting in antenna beamwidths narrow compared to 1°, have produced effective gains considerably less than the free space gain of the antennas. This apparent loss in gain, by currently accepted theory, is the result of a failure to illuminate the entire scatter region. For a fixed aperture antenna at both the transmitting and receiving end, the total path loss for any specific distance increases as the frequency increases. A part of this additional path loss results from a phenomenon described by some as "the medium to aperture coupling loss" and by others as "effective beam broadening." General curves relating total loss between transmitter and receiver to effective range are provided in Fig. 1. Total loss includes free space loss, scatter loss, system feeder loss, and medium to aperture coupling loss minus the transmitting and receiving antenna gains.

¹ K. Bullington, "Characteristics of beyond-the-horizon radio transmission," *Proc. IRE*, vol. 43, pp. 1175-1180; October, 1955. See also H. G. Booker and J. T. De Bettencourt, "Theory of radio transmission by tropospheric scattering using very narrow beams," *Proc. IRE*, vol. 43, pp. 281-290; March, 1955.

* Original manuscript received by the IRE, December 1, 1958; revised manuscript received February 24, 1959.

† Rome Air Dev. Center, Griffiss Air Force Base, N. Y.

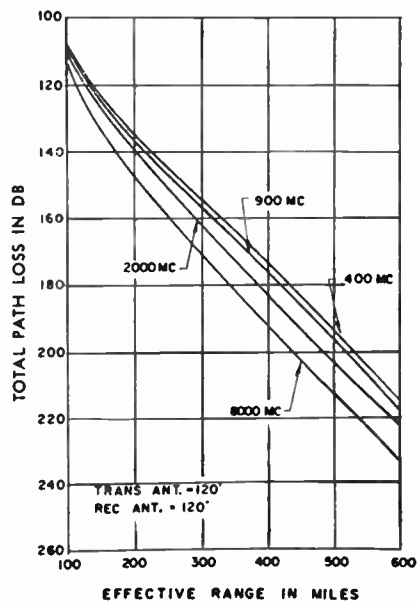


Fig. 1—Total path loss in db vs effective range in miles. The values for these curves were calculated for a general situation using parabolic antennas 120 feet in diameter.

Data for the curves were calculated using the method applied in Appendix II. Calculations are made for antennas 120 feet in diameter. Fig. 2 is a series of curves relating the total loss between transmitter and receiver minus the medium to aperture coupling loss to effective range. Data for Fig. 2 was calculated by the same method as the data shown in Fig. 1. Comparison of these figures shows the relative importance of medium to aperture coupling loss when very large antennas and microwave frequencies are considered. The system to be described, which uses angle diversity, is intended to overcome the limitations cited, and to permit the use of the higher microwave-frequency region where much less interference may be expected. This idea was jointly proposed by Harry Davis and Dr. Joseph H. Vogelmann, of the Rome Air Development Center.

PROPOSED ANGLE DIVERSITY SYSTEM

The system that is proposed utilizes an antenna capable of producing a beamwidth in the order of a fraction of a degree and illuminated by the use of many feed-horns. This parabolic antenna can be used to produce a multiplicity of beams, each at a different angle, to form an envelope which illuminates the entire scatter volume. Identical antennas are used at the receiving and transmitting ends of the path for each of the antenna feeds. Fig. 3 shows a typical antenna utilizing cassegrainian optics. Fig. 4 shows the antenna pattern which results from this type of antenna. Only the main lobes are shown. Each of the beams, one through five, is fed from its own transmitter at the transmitting terminal, and each beam is fed to its own receiver at the receiving terminal. The frequencies of the transmitters and receivers are adjusted to form beam pairs. The use of different frequencies for the various beam pairs is dictated

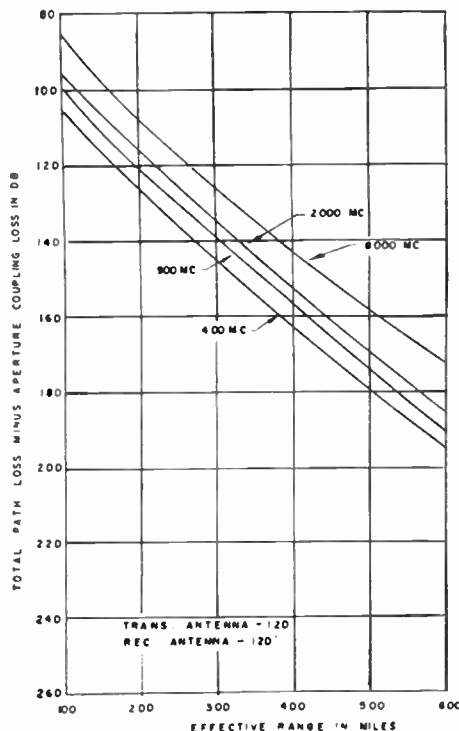


Fig. 2—Total path loss minus aperture coupling loss in db vs effective range in miles. Comparison of the curves in this figure with those shown in Fig. 1 provides an indication of the relative effect of medium to aperture coupling loss in a conventional circuit at various transmission frequencies and effective ranges. Values were calculated for the same parabola diameter, 120 feet.

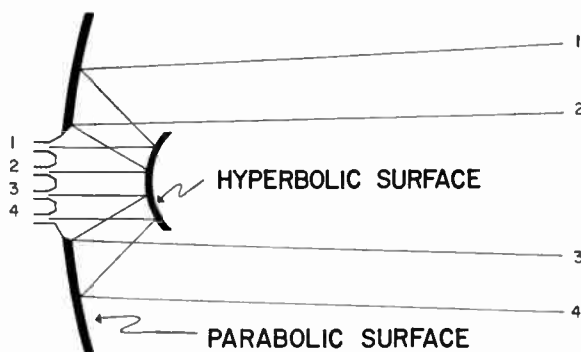


Fig. 3—Proposed antenna configuration.

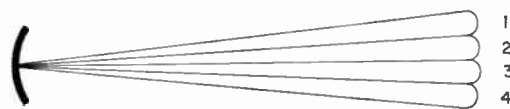


Fig. 4—Typical beam pattern. Adjacent beams cross at half power points.

by the need in this case for precise beam forming. Any frequency-diversity advantage gained is incidental, but welcome. Figs. 5 and 6 show the intersection of the beams, as seen from one side and from above, respectively, within the scattering region. For a particular frequency and effective range, an antenna size and beamwidth may be chosen at the point where realized an-

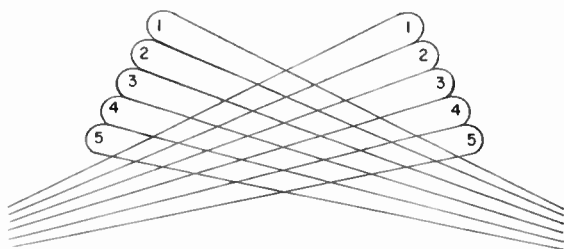


Fig. 5—Beam pattern—side view.

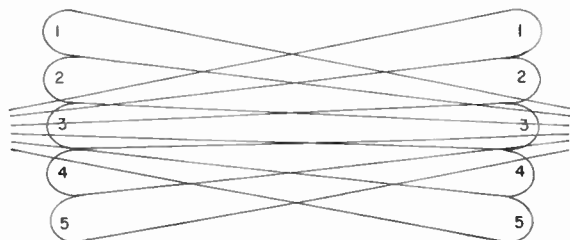


Fig. 6—Beam pattern—top view.

tenna gain begins to appreciably depart from the antenna free space gain. Illumination of the same region by a multiplicity of beams assures that the effective scattering region is illuminated in space to about the same degree. If one will accept the idea referred to above that medium to aperture coupling loss is a result of failure to fully illuminate the effective scatter region, and if the received signals from the various beams in this arrangement can be efficiently combined, it appears reasonable to presume that antenna size can be increased without experiencing the over-all effect on the system of medium to aperture coupling loss. If the transmitting antenna is of large size, the effective illumination of each portion of the scattering region is very much larger than would be the case if a smaller antenna is used to obtain the large beam necessary for the reduction of the medium to aperture coupling loss. For very long paths, use of the largest practical antenna and the highest practical transmitter power will certainly be required. For this reason, an approach to this problem on the basis of single-frequency operation using a single broad antenna beam and multiple receiving beams is not considered practical.

An additional advantage which arises from the multiplicity of antenna feeds is the ability to feed in a simple manner a separate transmitter on each beam if the transmitter frequencies are different. The total effective power will almost certainly be increased by the total number of beams, if we assume a low correlation between the beams. The combination of the received signals in a diversity combining system will greatly reduce the Rayleigh fading normally encountered in tropospheric scatter. To achieve the desired results, the received signals in each of the receivers must be translated in frequency to a common IF with an appropriate instantaneous phase-correction mechanism to permit their combination in an IF diversity combiner. Alter-

nately, they could be reduced to baseband and combined at audio frequencies with no requirement on the accuracy of phase correction.

A high order of space diversity can, of course, be achieved by use of a large number of receiving antennas spaced over a considerable area. Use of high transmitter power, a relatively broad transmitting beam, and an array of such receiving antennas all having relatively broad beams achieves the same result as that described above. This is certainly inefficient when compared on a dollar basis with the single transmitting and receiving antenna arrangement given above. Indeed, the real estate requirement for receiving antennas alone makes it impractical in many cases.

UNKNOWNNS

The effectiveness of this system depends on the degree of correlation between the signals on individual beam pairs in the angle diversity system. The computations which have been made in connection with this program have ignored the possibility that any improvement may be associated with the frequency diversity inherently possible in the system. This course of action appears to be desirable until sufficiently accurate data on the characteristics of frequency diversity in tropospheric scatter have been accumulated and published. Since the statistical combination of the reliabilities of the individual beams assumed the existence of a low order of correlation between the Rayleigh fading on each beam, it was considered essential to conduct an experiment to determine the degree of correlation which would be encountered in such a system. It has been shown by various investigators^{2,3} that for cross-correlations of less than 0.6, the degree of improvement due to diversity can be considered as essentially the same as that from completely uncorrelated signals. Correlation between beams was determined by the following experiments.

EXPERIMENT IN ANGLE DIVERSITY

The Directorate of Communications of Rome Air Development Center has a 7630- and 8450-mc experimental tropospheric test link between Mount Rose, N. J., and Forestport, N. Y., a distance of slightly more than 200 miles. This link uses two 28-foot parabolic reflectors at each end with space diversity. The transmitters are 2 kw, frequency modulated. This system is already operating with the antennas having a realized gain considerably smaller than the free space plane wave gain. The transmitting and receiving antennas have gains of 51.6 db each, as measured by conventional methods.⁴ In tropospheric scatter service, loss measurements indicated that these antennas provided approxi-

² R. Bolgiano, Jr., N. H. Bryant, and W. E. Gordon, "Diversity Reception in Scatter Communications with Emphasis on Angle Diversity," Contract AF30(602) 1717, Final Rep., pt. I; January, 1958.

³ H. Staras, "Diversity reception with correlated signals," *J. Appl. Phys.*, vol. 27, pp. 93-94; January, 1956.

⁴ Rome Air Dev. Center Contract AF30(635)-2856.

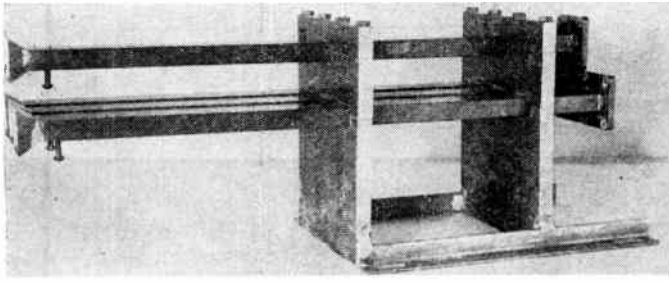


Fig. 7—Experimental feed configuration.

mately 89-db gain. The system, therefore, lent itself readily to an investigation of the usefulness of angle diversity for the improvement of tropospheric scatter communications. To eliminate the need for individual transmitters and for a multiplicity of feed horns, a 6-foot-diameter dish with a 2-kw transmitter was erected at the Mount Rose site. This combination provided a beam with a realized gain of 38 db compared to the free space plane wave gain of $41\frac{1}{2}$ db. This dish, transmitting to the 28-foot receiving parabola, would show a reduction in median signal to noise of $3\frac{1}{2}$ db arising from medium to aperture coupling loss. Its beamwidth would be sufficiently wide to permit the use of four beams on the receiving antenna with four individual receivers.

The existing feeds on both the transmitting and receiving antennas had been carefully positioned to provide maximum signal for the receiver on this particular path. An assembly of four feeds was made as shown in Fig. 7 using mounting plates dimensioned so that the middle feed of the row of three would be in identically the same position as the original feed. The outer feeds of the row of three were positioned from the centerline of the center feed, a distance calculated to skew their associated beams about three beamwidths. The feed above the row of three was positioned to provide a beam with axis skewed three beamwidths in the vertical plane from the axis of the middle beam in the horizontal row of three. Unfortunately, the patterns provided by this configuration could not be measured because of terrain difficulties. However, theoretical considerations indicate that the beamwidths were 0.3° wide and the angle between adjacent beams was 0.9° .

Fig. 8 and Fig. 9 are block diagrams of the equipment used at Forestport and Mount Rose, respectively.

Data were recorded on two beams at a time for each of the six possible combinations. The second part of the experiment consisted in transmitting 60 wpm teletype and reception of the teletype signal on each of two beams for the six possible combinations of antenna feeds. Both of the individual channels were recorded, as well as the combined signal using post-detection ratio combining.

Data were collected in two groups. First of all, signal-level information was obtained directly from the IF strips on the two receivers. Second, teletype signals were received and recorded at the video outputs of each receiver and the combined output. Reducing these data

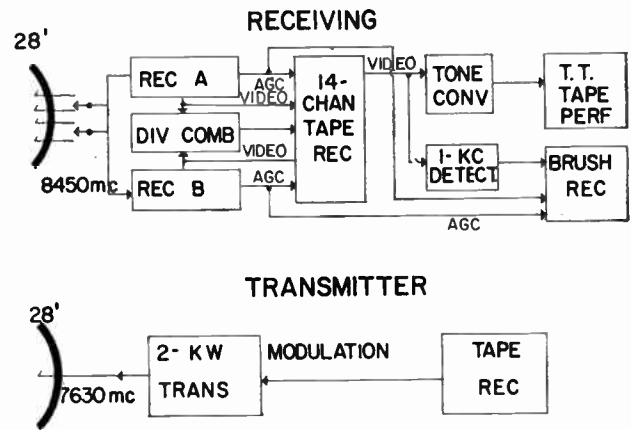


Fig. 8—Equipment diagram—Forestport terminal.

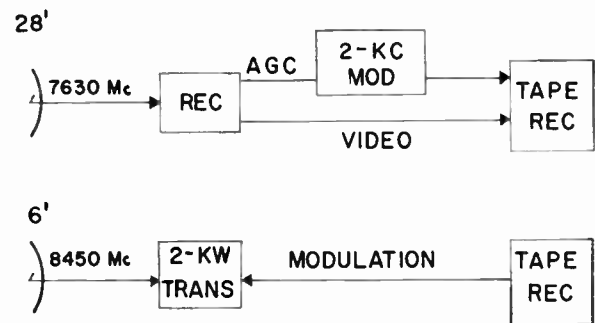


Fig. 9—Equipment diagram—Mt. Rose terminal.

allowed the determination of: a) signal level during the period of the test giving some idea of degradation of signal level with variation of beam direction from the great circle and from the horizon; b) cross-correlation of the signals between pairs of horns; c) the relative reliabilities on an individual circuit basis of the four horn configuration in comparison with calculated results. Reduced data are shown in Table I and in Fig. 11.

This test was conducted during the latter part of September and the first week of October, and so the path loss was about average for the year.

REDUCTION OF DATA

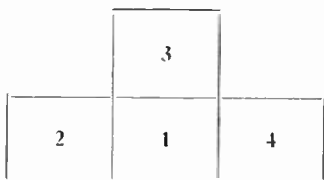
There were eleven runs of teletype data. Two channels at a time were used as inputs to the diversity combiner. The channels were individually monitored and teletype tapes were taken. The combiner output was also monitored and a tape was taken. These teletype tapes were analyzed by methodically comparing the received signal tapes with standard tape and noting the number of losses of synchronization as well as the numbers of correct and incorrect characters. Thus, the number of correctly-received characters was obtained by direct count.

Table I shows the cross-correlation figures obtained between the particular pairs of feeds. These were obtained by taking a very large number of points on the individual IF signal-level records and transposing these into digital data by means of a curve-translation device. The resulting numbers were then fed into a digital com-

TABLE I
REDUCED TELETYPE DATA

Combined Channels	Cross Correlation	
	Between Combined Channels	
1 & 4	0.075	
1 & 4	0.073	
2 & 4	0.037	
2 & 4	0.017	
3 & 4	0.120	
3 & 4	-0.001	
1 & 2	0.112	
1 & 3	-0.052	
1 & 3	-0.010	
2 & 3	0.184	
2 & 3	0.029	

Feed Diagram



puter programmed to calculate the standard coefficient of cross-correlation.⁵

It should be noted that the greatest cross-correlation obtained was 18.4 per cent, whereas in most cases, the advantages of diversity reception are lost if the signal correlation is above 60 per cent.³ The individual median signal to rms noise ratio was sufficiently high (at least 9 db) during the intervals in which the cross-correlation was measured, that the effects of noise in lowering the correlation may be considered small. Signal-strength data were not taken over periods long enough to allow their direct use in calculating long-term median estimations, but long-term median data is not necessary for the evaluation of this multibeam system.

These measured values of correlation were in close agreement with extrapolation of a curve originally prepared by Dr. W. E. Gordon and his associates.² This curve relates the correlation coefficient to the beam separation, and is shown in Fig. 10, with the spread of correlation values obtained also indicated.

"Signal Reliability" is defined as the percentage of time the signal is at least 10-db above the receiver threshold. This reliability on each feed horn is obtained from the theoretical calculations by taking path length, azimuth, elevation, etc., into account. Assuming no correlation between the signals and switch combining, the reliability of the combined signal is found by: $1 - P_c = (1 - P_a)(1 - P_b)$, where P_c is the reliability of the combined signal, P_a the calculated reliability of the signal on horn A, and P_b the calculated signal reliability on horn B.

The measured reliability of the combined signal was obtained from the teletype data. In order to relate the measurement of character reliability (percentage of cor-

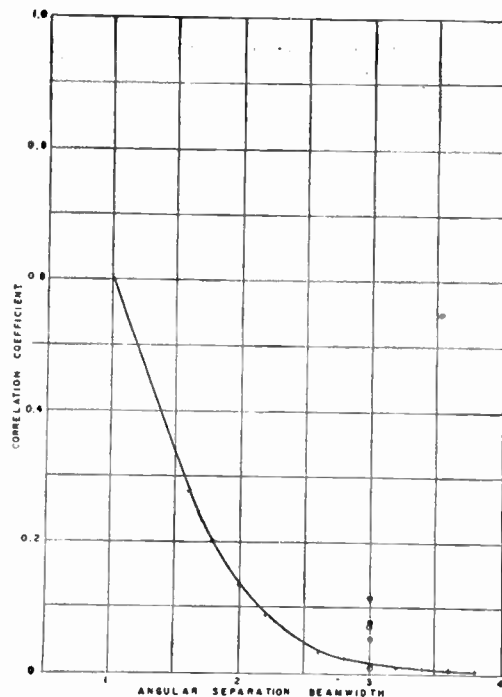


Fig. 10—Correlation coefficient vs angular separation in beamwidths. The curve is an extension of a curve calculated by Bolgiano, Bryant, and Gordon.² Five experimental points are plotted at the appropriate angular separation.

rect characters received) to a measured reliability of the combined signal, the average number of wrong bits per wrong character had to be estimated. This was done by assuming a simple linear relationship: $X = 5(1 - P)$, where X is the average number of wrong bits per character and P is the measured statistical signal reliability of the two channels under consideration. In general, a direct knowledge of the fade rate would be necessary to provide a better curve for the error rate instead of a straight line. This was not considered necessary for this evaluation since precise values were not needed.

To obtain P , it was necessary to compute the character reliability from the tape count as a function of the average number of bits lost per character by using $D = 1 - (XW/5R + 5W)$ where R is the number of correct characters, W the number of wrong characters, and D the character reliability. The values thus obtained for X enable us to compute P , the measured reliability of the two combined channels. The points thus obtained are compared to the line $P = P_c$ in Fig. 11. In this case, a linear relationship between signal reliability and error rate seemed to be a reasonable assumption.

In all cases, the measured reliability was higher than the calculated reliability for each pair. This may indicate that the method used to calculate reliability described in Appendix II results in slightly lower values than can be obtained in an actual system.

EVALUATION

The cross-correlations obtained were low enough to allow us to consider each beam as being statistically in-

⁵ H. Cramer, "The Elements of Probability Theory," John Wiley and Sons, Inc., New York, N.Y., p. 173; 1955.

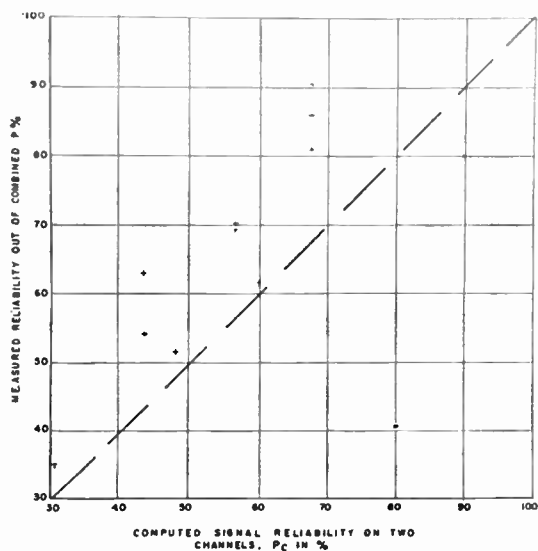


Fig. 11— P vs P_c . The abscissa P_c is the computed reliability of the signal out of the combiner, obtained by multiplication of the individually computed signal reliabilities. The ordinate P is the quantity resulting from the error count of the actual teletype output of the combiner. They are assumed to be related linearly by the equation given in the text. The 10 resultant points are from the 6 possible channel combinations used during the runs of teletype data. Since P_c was computed on the basis of switching diversity, the use of linear combining diversity provides measured values of reliability all higher than the dashed line $P_c = P$.

dependent and to expect angle diversity performance identical with that achieved by space diversity. The theoretical calculations were an adequate approximation for the experiment.

BASIS OF COMPUTATION

All of the computations made in the angle diversity system were computed for each beam separately. These computations were made using the system prescribed by Yeh.⁶ They were made with an antenna oriented so that the beam under consideration was directed along a great circle route and at the horizon. Path loss was calculated by the Yeh method and adjusted for a deviation in azimuth from the great circle route and in elevation from the horizon by means of published data.⁷ The corrections for azimuth and elevation angles are shown in Figs. 12 and 13, and are taken directly from Friis, *et al.*⁷ They were calculated for a particular path and verified by an experiment described in the reference. They are used for this general computation because the losses indicated are considered pessimistic. The reliability of each of the received beams was then computed independently. They were then combined on a statistical basis to obtain a combined reliability.

⁶ L. P. Yeh, "Tropospheric Scatter System Design," Westinghouse Electric Co., Tech. Rep. No. 5; October, 1957.

⁷ H. T. Friis, A. B. Crawford, and D. C. Hogg, "A reflection theory for propagation behind the horizon," *Bell Sys. Tech. J.*, vol. 36, pp. 627-644; May, 1957.

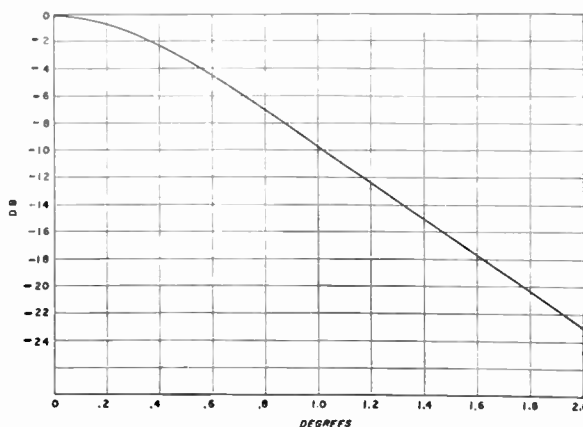


Fig. 12—Relative received power vs azimuth angle. This curve was calculated and published⁷ with regard to a particular path. It is required for general path loss calculations utilized in the preparation of Figs. 16-18 and Table II. For design purposes, a similar curve appropriate to the path in question may be easily prepared by means of the method given in the literature.⁷

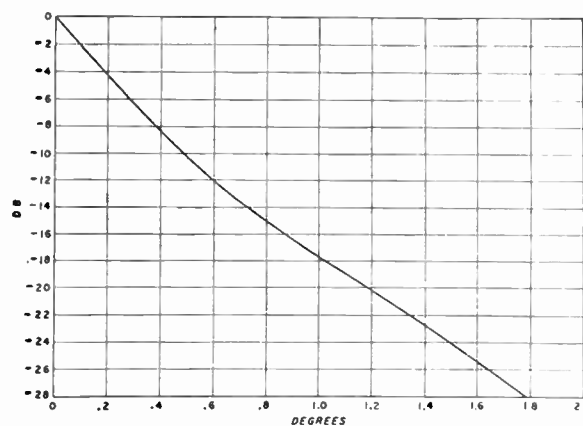


Fig. 13—Relative received power vs elevation angle. This curve was calculated and published⁷ with regard to a particular path. It is required for general path loss calculations utilized in the preparation of Figs. 16-18 and Table II. For design purposes, a similar curve appropriate to the path in question may be easily prepared by means of the method given in the literature.⁷

THEORETICAL COMPARISON OF THE SYSTEM

In order to compare the usefulness of this approach, theoretical calculations have been made on the following configurations.

Four systems were computed to operate at 400, 900, 2000, and 8000 mc. Each would employ transmitting and receiving antennas 120 ft in diameter. In addition to angle diversity, space diversity is utilized in the unconventional system. Four antennas would be used, two at the transmitting terminal and two at the receiving terminal, with appropriate polarization to achieve a total of quadruple space diversity in the conventional circuit. Angle diversity improvement is added to this in the unconventional system. The transmitter power considered was 10 kw per feed in every case. In the unconventional system, one feed (*i.e.*, 10 kw) is used at 400 mc, six feeds (*i.e.*, 60 kw) at 900 mc, and twelve feeds

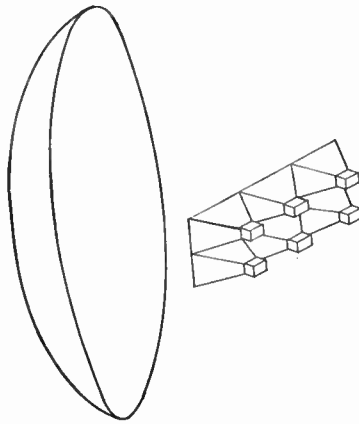


Fig. 14—Six feed system.

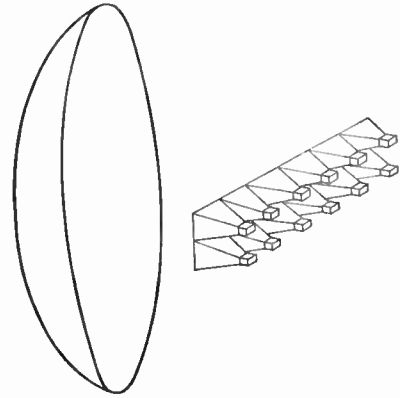


Fig. 15—Twelve feed system.

(i.e., 120 kw) at 2000 and 8000 mc, respectively. A combiner gain of 3 db was assumed in all cases. A receiver thermal noise for the 500-kc HF bandwidth of minus 146 dbw was used in each case. The use of parametric amplifiers with a noise figure of 1 db was assumed. Figs. 14 and 15 show the arrangements for six and twelve feeds. Feed size is not in scale in the interest of clarity. Table II shows the ranges that can be expected under the conditions described at each of the frequencies.

The computed values indicate the benefit we might expect to enjoy after utilizing this application of angle diversity. With all other parameters the same, it provides the extra order of reliability required for good operations over an unusually long circuit without exceeding the state-of-the-art of the transmitter field. This is so because an increase in order of diversity is primarily effective only when a high order of reliability is required. Further, application of the higher microwave frequencies, heretofore rejected because of high total loss, some of which is caused by medium to aperture coupling loss, becomes attractive.

CONCLUSIONS

Table II demonstrates the advantages to be achieved by using angle diversity at super-high frequencies, as compared to a single-feed 120-foot antenna used at 400 mc. The cost of achieving this performance is the increase in transmitters and receivers that are required. To compensate for this cost, however, the increased equipment reliability that results cannot be overlooked. The failure of a single transmitter or receiver in a six or twelve feed system would reduce the over-all reliability only slightly. This would be particularly true if switching provisions were incorporated to permit utilizing those horns which produce the highest median signals. A critical shortage of frequencies in the lower UHF band can now be alleviated by use of the angle diversity systems described in this paper.

Over and above this technique, the addition of multiple receivers on each of the feeds would increase the order of diversity without increasing the transmitter requirement. In a practical system, the summation of

Frequency in mc	400	900	2000	8000
Number of feeds	1	6	12	12
Path loss (db)	167	171	178	183
Effective range in miles	360	370	380	385
Total power required in kw	10	60	120	120

Reliability—99.9 per cent

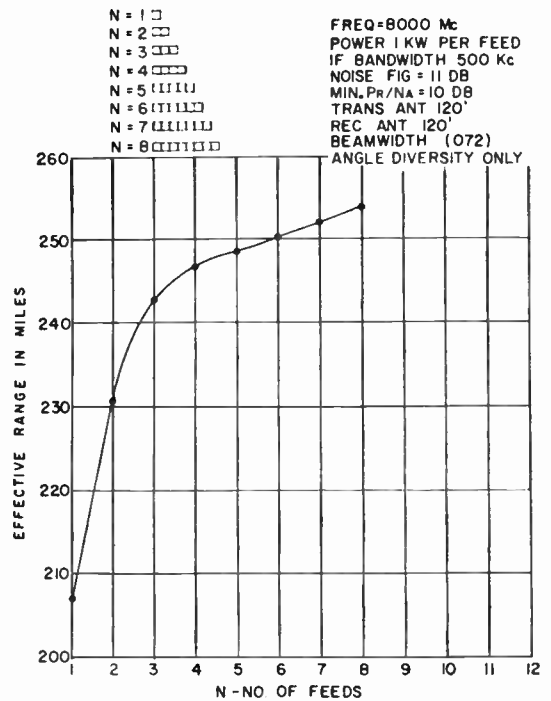


Fig. 16—Effective range in miles vs number of feeds of angular diversity system at 99 per cent reliability. This calculated curve shows the effect of adding additional beams in the horizontal plane in a very narrow beam system.

both frequency and angle diversity is probably optimum. More data on frequency diversity is required before adequate calculations can be made.

APPENDIX I

Fig. 16 shows the improvement which can be obtained by increasing the number of feeds from one to eight for

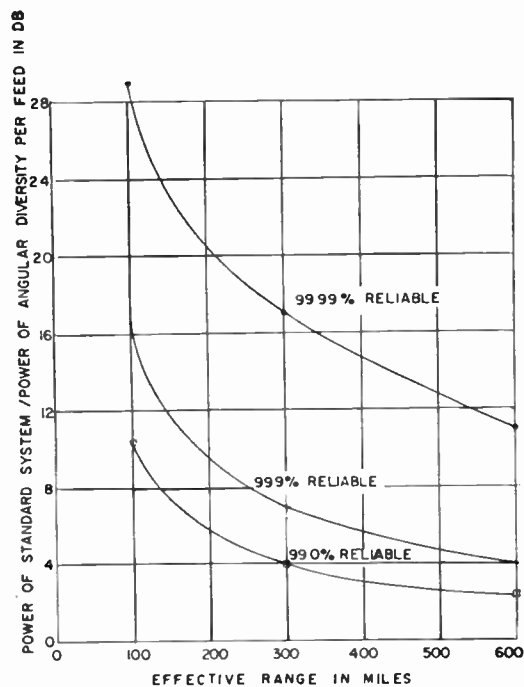


Fig. 17—Ratio of power of standard system over power per feed of angular diversity system in db vs effective range in miles. These curves indicate the reduction in design problems involving the transmitting terminal expected with the use of a multiple angular diversity system at 900 mc.

an 8000-mc system. It is a plot of the effective range achievable in miles as a function of the number of feeds for 99 per cent reliability. It assumes the following characteristics:

1 kw of power per feed.

An IF bandwidth of 500 kc.

A noise figure at the receiver of 11 db.

Minimum signal to noise ratio for acceptable performance—10 db.

This curve experiences a saturation effect with a number of feeds higher than shown because of the rapid increase of loss with increase of angle of beam skew as well as the reduced effectiveness of very high orders of diversity. Obviously, the gain in range is not worthwhile in terms of equipment complexity.

Fig. 17 shows the ratio in db of the power required in a standard single-feed antenna system to that required per feed in an angle diversity system for 99 per cent, 99.9 per cent, 99.99 per cent reliability as a function of effective range in miles. These data have been plotted for 900 mc. An antenna size of 120 feet has been assumed.

Fig. 18 is a plot of the ratio in db of the power of a conventional system to the total power of an angle diversity system for various reliabilities and path lengths at 900 mc and utilizing 120-foot antennas. Calculations for Figs. 17 and 18 were based on a conventional system utilizing standard space diversity and a system utilizing this diversity plus angle diversity. All other parameters

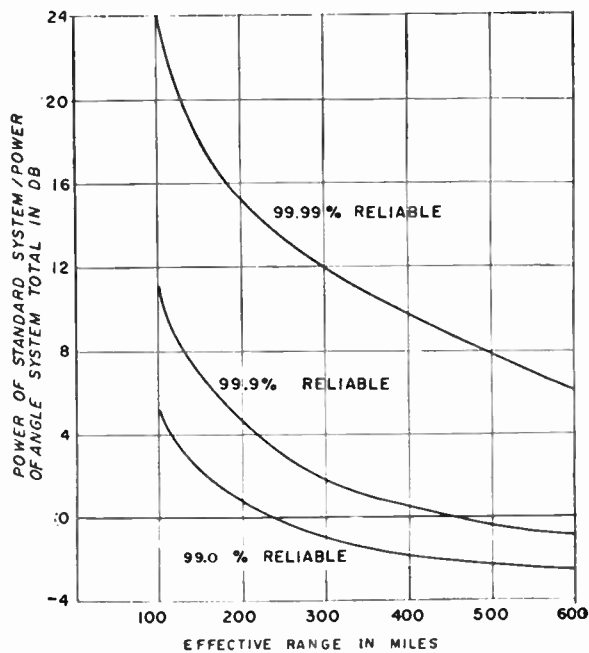


Fig. 18—Ratio of power of standard system over total power of angular diversity system in db vs effective range in miles. These curves indicate that a reduction in total power requirement could be expected at the transmitter terminal for circuits which must operate with 99.99 per cent reliability.

are the same. These curves are offered as an illustrative case only. Obviously, with change of parameters, the improvement would be somewhat different. In general, the improvement made possible by use of this technique is indicated here, as well as the fact that the technique may have economic value when applied to relatively short distance circuits.

APPENDIX II SAMPLE CALCULATIONS

These calculations for reliability for the following tropospheric scatter circuit are based on a procedure outlined by Yeh.⁶

Path length in statute miles = 400

Frequency in mc = 900

Modulation = FM

Power/feed in kw = 2

Diversity (space) = 4

Number of feeds (angle diversity) = 6

Antenna diameters in feet = 120

Antenna heights above sea level in feet = 1000

IF bandwidth of receiver in kc = 500

Receiver noise figure in db = 1.

Free Space Loss (L_{fs}):

$$L_{fs} = 36.6 + 20 \log d + 20 \log f$$

d = path length in statute miles

f = frequency in mc

$$L_{fs} = 36.6 + 52 + 59 = 147.6 \text{ db.}$$

Transmitting Antenna Gain (G_t):

$$G_t = 8 + 20 \log D/\lambda$$

D/λ = ratio of antenna diameter to wavelength

$$G_t = 8 + 20(2.042) = 48.8 \text{ db.}$$

Receiving Antenna Gain (G_r):

$$G_r = 8 + 20 \log D/\lambda$$

D/λ = ratio of antenna diameter to wavelength

$$G_r = 48.8 \text{ db.}$$

Aperture Coupling Loss (L_c): from Fig. 4 of Yeh:⁶

$$L_c = 20.6 \text{ db.}$$

System Feeder Loss (L_a):

$$L_a = 5 \log f - 10$$

f = frequency in mc

$$L_a = 14.8 - 10 = 4.8 \text{ db.}$$

Scatter Loss (L_s): (for the distance horizon to horizon only as per the *Effective Distance* method of calculation)

Distance to horizon from 1000-foot antenna is 45 miles.

Distance between horizons is $400 - 2(45) = 310$ miles.

Scatter loss for 310 miles

$$L_s = 57 + .15 (d - 100)$$

d = distance from horizon to horizon in statute miles

$$L_s = 57 + 31.5 = 88.5 \text{ db.}$$

Total path loss (L_t):

$$L_t = L_{f_s} - G_t - G_r + L_c + L_a + L_s$$

$$L_t = 147.6 - 48.8 - 48.8 + 20.6 + 4.8 + 88.5$$

$$L_t = 163.9 \text{ db.}$$

Transmitted power (P_t): 33 dbw.

Received power (P_r):

$$P_r = P_t - L_{f_s}$$

$$= 33 - 163.9 = -130.9 \text{ dbw.}$$

Thermal Noise (N_a):

$$N_a = 10 \log FKT B_{IF}$$

F = Noise figure of receiver (1 db or ratio 1.26)

K = Boltzmann's constant (1.38×10^{-23} joules/°K)

T = 290°K

B_{IF} = Receiver IF bandwidth in kc

$$N_a = 10 \log (1.26) (1.38 \times 10^{-23}) (290) (500 \times 10^3)$$

$$N_a = -146 \text{ dbw.}$$

$$\text{Median } \frac{P_r}{N_a} = -130.9 + 146 = 15.1 \text{ dbw.}$$

$$\text{Median } \frac{P_r}{N_a} - \text{Minimum } \frac{P_r}{N_a} = 15.1 - 10 = 5.1 \text{ dbw.}$$

Combiner gain (G_c): 3 db.

Fading Margin: $5.1 + 3 = 8.1$ dbw.

Reliability for one feed on center: from Fig. 8 of Yeh:⁶

99.1 per cent.

A separate reliability is calculated for each of the six feeds which are arranged in this manner.

The beamwidth (BW) of the individual beams using a 120-foot dish is:

$$BW = 70\lambda/D$$

λ/D = ratio of wavelength to antenna diameter

$$BW = 70 (1.09)/120 = .635.$$

From curves by Friis,⁷ the attenuation due to the beams deviated from the optimum path are calculated. These curves are shown in Figs. 12 and 13. The individual fading margin for each received signal is entered in the box corresponding to its receiving horn.

-3.7	1.3	-3.7
-3.7	1.3	-3.7

The individual reliabilities are entered in the appropriate box (from Fig. 8 of Yeh⁶).

55	85	55
55	85	55

The individual reliabilities are combined on a statistical basis to give the combined reliability of 99.91.

ACKNOWLEDGMENT

The contribution made by the Philco Corporation in equipment operation and by the Scientific Computation Branch of the Rome Air Development Center in data reduction is acknowledged.

Special acknowledgment is made to Harry Davis, scientific director of the Rome Air Development Center, for the basic idea which led to this investigation.

A Critical Analysis of Some Communications Systems Derived from Amplitude Modulation*

WARREN D. NUPP†

Summary—For radio operation in the frequency range of 30 mc downward, ordinary amplitude modulation (AM) has been most widely used. In order to overcome certain known deficiencies of AM, variations of AM have been evolved. One of these, single sideband operation with the carrier suppressed at the transmitter (SSB), has proved its superiority over AM in transatlantic operations for nearly twenty-five years. A notice of proposed rule-making by the FCC, in 1955, to make mandatory the use of SSB below 25 mc in certain services, quickly caused two other systems to be put forward for consideration. These were 1) a system of transmitting both sidebands with the carrier suppressed at the transmitter (DSB), and 2) a form of single sideband with the carrier transmitted, called Compatible Single Sideband (CSSB).

This paper deals mostly with a comparison of SSB and DSB, with emphasis on the applications to the aeronautical mobile service. There is discussed 1) spectrum requirements as influenced by transmission of unwanted frequencies outside the nominal modulation band, 2) the capability of DSB to reject an interference occurring on one sideband, 3) the problem of whistle heterodynes in a radio environment where signals with and without the carrier are both present, 4) the problem of AGC with SSB and DSB, and the effect on phase lock of the reinserted carrier, 5) obtaining a phase lock with synchronous detection, 6) the factors affecting the required accuracy in frequency of the reinserted carrier, and the accuracy obtainable at the present state of the art, 7) the effect of a carrier in an AM signal on reinsertion of a carrier in a DSB receiver receiving this signal, 8) the effect of type of modulation on peak RF envelope power, and a discussion of multiplexing binary signals, 9) the advantages of DSB for binary transmission and a discussion of diversity operation, 10) the effects of impulse type interference, and, 11) a brief discussion of CSSB.

INTRODUCTION

SINGLE sideband operation with the carrier suppressed at the transmitter (SSB) has proved its superiority over AM in transatlantic operation. The action of the Federal Communications Commission in a notice of proposed rule-making, in 1955, to make mandatory the use of SSB below 25 mc in certain services was prompted also by the shortage of space in the high frequency radio spectrum.¹ This action caused two other systems to be put forward quickly for consideration. These were 1) a system of transmitting both sidebands with the carrier suppressed at the transmitter (DSB),² and 2) a form of single sideband with the carrier transmitted, called Compatible Single Sideband (CSSB).³ A lack of certain kinds of information on these systems may well result because they were not yet fully

ready for presentation by their sponsors. However, the potentialities of these systems were sufficient to cause an investigation to be launched in a government laboratory, an investigation of variations of AM for purposes of determining their relative value, principally with regard to use in the aeronautical mobile service. The description of the systems has been purposely omitted because of the general availability of this information in the literature.

In the course of the work, a scope of the subject was revealed that hardly could have been realized at the beginning. Each phase of the subject is material for one or more papers. Admittedly, then, the subject is imperfectly understood at this stage of the development; therefore, no apology is offered for omissions or inaccuracies that may be found in this paper. However, it is hoped and believed that it will provide the stimulus to further work of benefit in the field of high frequency communications.

SPECTRUM REQUIREMENTS

Normally there is assumed, in the design of a communications system, some band of modulation frequencies from a high limit down to a low limit. For voice transmission stressing high intelligibility rather than high fidelity, these limits might well be 3300 cps and 300 cps. The band of radio frequencies required for AM transmission in this case would be 6.6 kc, whereas the band of frequencies required for SSB transmission would be only 3.0 kc. This comparison is typical, in that double sideband transmission requires a little over twice the bandwidth required for SSB. The situation found in practice, however, may be much different due to the transmissions containing spurious frequencies outside the desired RF band.

In generating the SSB signal, the entire unused sideband and all other frequencies outside the desired sideband may be suppressed to practically any degree desired by proper filtering. The problem of spurious transmission lies principally in the matter of amplifying the weak generated signal to the final power to be transmitted. Nonlinearity in these amplifiers produces spurious frequencies in a band within and adjacent to each side of the desired band. At the present state of the art, these spurious frequencies can be expected to be suppressed about 40 db (as measured by a two-tone modulation test) with respect to the desired transmission. It appears that this is not sufficient to prevent serious interference in the aeronautical mobile service with a signal in an adjacent band. A determination of the width of the guard band, required to reduce this inter-

* Original manuscript received by the IRE, December 1, 1958; revised manuscript received, February 17, 1959.

† U. S. Naval Air Dev. Center, Johnsville, Pa.

¹ D. G. Fink, "Danger! the radio spectrum is bursting at the seams," *J. Frank. Inst.*, vol. 261, pp. 477-493; May, 1956.

² J. P. Costas, "Synchronous communications," *PROC. IRE*, vol. 44, pp. 1713-1718; December, 1956.

³ L. R. Kahn, "A compatible single-sideband system," presented at Second Annual Symp. on Aeronautical Communications, Utica, N. Y.; October, 1956.

ference to a satisfactory level, would require an investigation of greater scope than may be undertaken here. Nevertheless, this writer will present a form of reasoning in the matter that may be of value. The close-by interference spectrumwise under discussion occurs from odd-order terms, starting with the third order, in the transfer characteristic of the RF amplifiers. At least two modulation frequencies are required to be present simultaneously. The minimum requirements, then, to produce the interference, are two modulation frequencies and third order distortion. Assuming these minimum requirements, and assuming further that one modulation frequency may be near the low end of the band of modulation frequencies when the other is near the high end, the guard band required is equal to the nominal band of transmission. The departure from this condition that occurs in practice is that more than two modulation frequencies are present and that higher than third order terms are present. The effect of more than two modulation frequencies is to cause the interference to be concentrated closer to the desired band, and the effect of the other is to spread out the interference. These effects, being opposite, would tend in some degree to cancel. A thorough investigation increases greatly in scope from this point, involving such things as expected orders of nonlinearity of the RF amplifiers, expected types of modulation, for many equipments in many geographical locations, under expected conditions of service, etc. Short of such an investigation, the material that has been given indicates that the guard band should be equal to the nominal band.

The matter of spectrum requirements as influenced by linearity of the RF amplifiers appears the most difficult to deal with, although there are other sources of close-by interference spectrum-wise, such as results from imperfect stabilization of a mixer oscillator, or noise amplified within the transmitter. The SSB signal may be constructed at a high power level as in the systems of Villard⁴ and Kahn,⁵ thus eliminating the need for linear amplifiers. However, the close-by spurious frequencies resulting with these systems tend to be of no less amplitude than with linear amplifiers.

A comparison of the bandwidth required for double sideband operation as compared with single sideband operation, therefore, depends on the spurious frequencies to be found in the bands adjacent to the band of desired frequencies, for double sideband operation. The means to generate the DSB signal is somewhat different from that to generate the SSB signal or the AM signal. It is shown⁶ that the performance of ordinary AM is generally inferior to SSB and DSB operation in the presence of selective fading, so that the matter of spurious fre-

quencies incidental to generation of the ordinary AM signals need not be considered. Information on the spurious frequencies accompanying generation of the DSB signal by the General Electric Company equipment² has not been published.

REJECTING INTERFERENCE ON ONE SIDEBAND OF THE DSB SIGNAL

The DSB receiver of the General Electric Company^{2,7} has provision for rejecting an interference appearing on one sideband. This rejection of interference, however, does not take place automatically but must be done by manual switching at the receiver, the operator trying each sideband to learn which sideband has the interference. The claim that both sidebands of the signal continue to be utilized when the interfering signal is removed is only partly true; both sidebands continue to be utilized for the phase lock of the reinserted carrier, but one sideband of the desired signal is removed from the receiver output. Therefore, the reception in this case is essentially SSB. This brings to attention the fact that reception of a DSB signal by a SSB receiver provides 3 db less signal-to-noise than reception of the same signal by a DSB receiver; this is because the signal in the SSB receiver is reduced 6 db but the noise is reduced only 3 db. The noise referred to is the normal noise at the receiver. With respect to the interference as a noise, a net improvement of signal-to-noise could, of course, result from removal of interference on one sideband.

WHISTLE HETERODYNES WHEN SIGNALS WITH AND WITHOUT THE CARRIER ARE BOTH PRESENT

The whistle interference produced by heterodynes between carriers of AM signals is well known, especially to listeners of the amateur bands. It would be a welcome relief from whistle heterodynes for the carriers of all radio signals to be suppressed. In ordinary AM reception it takes two overlapping signals to produce the whistle, and the whistle is strong only if the strength of the undesired signal is of comparable strength to the desired signal. However, while searching the frequency spectrum with a SSB or DSB receiver, *all* AM signals individually produce a very strong whistle heterodyne. The trend in the aeronautical mobile service has been towards the use of receivers automatically tunable to the desired frequency, which eliminates the whistle heterodynes resulting while searching the spectrum. This writer, however, rebels at the idea of being constrained to use such a receiver. Certainly, the environment would be more favorable for obtaining the kind of information to be learned in on-the-air listening tests if portions of the spectrum were reserved exclusively for SSB and DSB transmission. Some degree of compatibility in reception of SSB and DSB signals is evidenced by the fact that the General Electric synchronous de-

⁴ O. G. Villard, Jr., "A high-level single-sideband transmitter," *Proc. IRE*, vol. 36, pp. 1419-1425; November, 1948.

⁵ L. R. Kahn, "Single-sideband transmission by envelope elimination and restoration," *Proc. IRE*, vol. 40, pp. 803-806; July, 1952.

⁶ W. D. Nupp, "Effects of selective fading on AM signals," *Proc. IRE*; to be published.

⁷ J. K. Webb, "A synchronous detection adapter for communications receivers," *CQ*, vol. 13, pp. 30-33, 117; June, 1957.

tection receiver, designed primarily for DSB reception, receives equally well a SSB or a DSB signal.

THE PROBLEM OF AGC AND NOISE LIMITERS WHEN THE CARRIER IS SUPPRESSED

The severe fading and distortion of the AM signal commonly encountered in the presence of multipath propagation is discussed,⁶ and it is shown that most of the time the sidebands and carrier essentially fade together. For ordinary AM transmission, therefore, AGC is a very necessary adjunct at the receiver, but one which serves very well. Fading rates encountered are sufficiently slow that the time constant of the AGC in most AM receivers is a goodly part of a second for both response and decay. Two advantages of the long time constant are as follows: 1) The long time constant acts essentially as a filter to prevent the modulation actuating the AGC. The advantage of this is that the receiver remains at the proper gain in the complete absence of modulation and, therefore, is at the proper gain the instant modulation returns. 2) Impulse type noise, the most common type of man-made interference, does not affect operation of the AGC. These advantages, which one tends to take for granted in AM operation, show their true worth by being absent in AGC systems used with SSB and DSB operation.

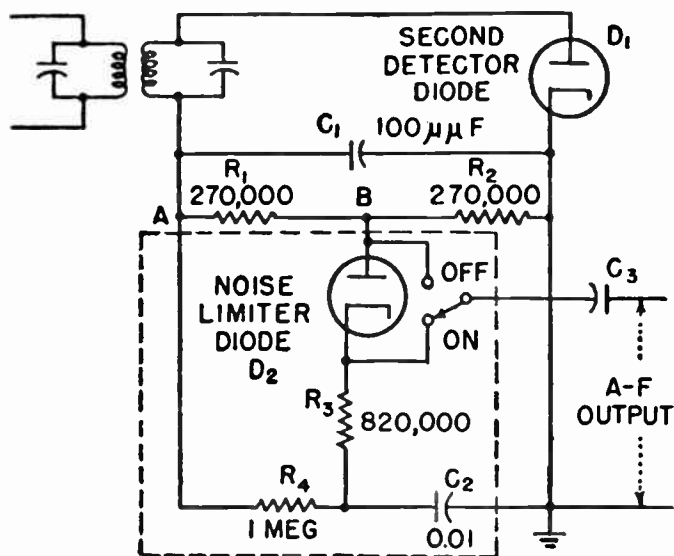


Fig. 1—Circuit for a series limiter.

The ordinary series limiter, as shown in Fig. 1, and described by Toth,⁸ is a remarkably simple and effective device for reducing the effect of impulse-type radio interference at the receiver. Its effectiveness is in part due to the fact that the proper level of limiting is automatically set by the carrier strength. Therefore, it is not surprising that it is almost universally incorporated in good communications receivers for reception of AM signals.

⁸ E. Toth, "Noise and output limiters, part 1," *Electronics*, vol. 19, pp. 114-119; November, 1946.

It is essential for proper location and reception of weak signals that a communications receiver be at full gain in the absence of a signal. In the reception of the SSB signal, it is necessary that the AGC be actuated by the modulation and be very quickly responsive to prevent receiver overload at the start of transmission. The response time for the AGC of a typical receiver for reception of the SSB signal in the aeronautical mobile service is only a fraction of a millisecond, and the decay time is a goodly part of a second to maintain AGC action during and between syllables of voice transmission. This AGC characteristic is, however, adversely affected by impulse-type interference.

Quantitative information on the adverse effect of impulse interference on SSB reception is not available, but it is undertaken here to present information on AM reception that appears suitable to show results to be expected for SSB reception. The type of subjective tests generally employed to evaluate the effect of interference on voice reception requires the use of special word lists and many selected speakers and listeners. By comparison, the test results to be presented are oversimplified in that only one listener, namely the writer, was employed, but the results are believed of value as a first look at the problem. When an impulse signal from an impulse generator, at the rate of 100 per second, is superimposed on AM voice transmission from a U. S. broadcast station as received on a standard AM communications receiver, over fifty per cent of the transmission is intelligible when the peak amplitude of the interference, in volts, is ten times the peak amplitude of the voice signal, and overloading is not evident, as viewed on an oscilloscope at the receiver output. This repetition rate of the impulse signal is roughly that to be expected when the interference from an aircraft engine ignition system comes from all cylinders. The rate would be much lower, of course, if the interference were from one cylinder, in which case the intelligibility of the signal would be considerably higher. For a response time of only a fraction of a millisecond, as described above, the AGC operating level of the SSB receiver would be essentially set by the peak value of the interference. The resulting reduction in amplitude of the voice signal by about 20 db might well completely remove a weak voice signal, and this as a result of an interference which by AM standards is only moderately effective in reducing intelligibility of the voice signal.

The test results just given are oversimplified in another respect, in that the effects of receiver overloading by the impulse signal, which may well occur in practice, are not included. This indicates that the test results given should be taken only as an indication of the potential seriousness of the problem and to show the need for further investigation.

The tendency for the SSB signal level to be affected by impulse noise, and the lack of a carrier with the SSB signal, obviate use with SSB and DSB a series limiter of the type used with AM. This situation might be

partly overcome by a compromise; that is, the AGC response time could be lengthened at the expense of an explosive sound at the start of modulation, thus reducing the effect of interference on receiver gain. The bias supplied in the noise limiter circuitry by the carrier could be replaced by a fixed dc voltage. However, possibly another problem exists. The series limiter of the AM receiver works with a signal out of the envelope detector from an impulse signal that is of one polarity only. This is not true of the product detector used for SSB reception.

The impulse signal, after traversing the radio frequency portion of a superheterodyne receiver, has a length roughly the reciprocal of the IF bandwidth. It might be called a damped sinewave (it is essentially a damped sinewave in the early stages of the receiver, but the IF stages shape it so that the time of rise and fall are about equal at the detector) at the center frequency of the IF pass band. The reinjected carrier of a SSB receiver, however, has a frequency approximately at the edge of the passband. This superimposes on the detector output from an impulse signal a "beat" at a rate close to half the IF bandwidth. Therefore, the shape of the detector output from the impulse signal depends on the phase relation that happens to exist between the reinjected carrier and the RF impulse signal. It is not difficult to see that the detector output from the impulse signal will be randomly of either polarity and may reverse polarity during the pulse. Because of the many problems involved, the noise limiters attempted by manufacturers of SSB receivers have been of low effectiveness.

OBTAINING A PHASE LOCK WITH SYNCHRONOUS DETECTION

It might appear that the carrier reinserted at the receiver of a DSB signal must be at exactly the frequency of the suppressed carrier, and, in addition, in the proper phase. This is true on the average, but it will become clear that the instantaneous frequency of the reinserted carrier must make deviations from the frequency of the suppressed carrier. It is shown⁶ that the phase of the reinserted carrier with the sidebands must vary in a prescribed manner as relative phase of the signals arriving by multipaths changes.

As a simplification for the first case to be considered, suppose that the DSB signal arrives undistorted by fading, in which case the reinserted carrier is required to be in the phase of the suppressed carrier. Suppose further, that the reinserted carrier happens to be at the exact frequency of the suppressed carrier but not in phase with it. In order to get in phase, the reinserted carrier must momentarily deviate in frequency, and the more rapidly it is brought to proper phase, the greater the deviation in frequency. The changing phase relation required of the reinserted carrier, mentioned as being caused by selective fading, the effect of changing frequency content of the modulation in combination with

selective fading, and the fact that synchronization is lost during the gaps in modulation, all cause the frequency of the reinserted carrier, generally, to be in a state of considerable deviation in order to maintain proper phase. No analysis of the more complex dynamic condition will be made, but one more simple case will be considered, which lends clarification to the mechanism of synchronizing the reinserted carrier.

In the matter of reinserting the carrier in proper phase, it is to be noted that the phase position 180° removed is equally good. Therefore, the greatest correction in phase ever required of the reinserted carrier is 90° . Let us suppose that a gap in modulation causes synchronization to be lost. The frequency of the unsynchronized reinserted carrier might differ from the frequency of the suppressed carrier in a practical case by, say, 50 cycles. However, every 10 msec, the reinserted carrier would be of exactly the proper phase, at which times no synchronizing voltage would be generated by the modulation. Between these times, there would be momentary synchronizing voltages of alternating polarity; and if these were strong enough and the response time of the system sufficiently short, the system could "take hold" of one of the synchronizing voltages and fall into essentially proper frequency and phase in a time not much greater than 10 msec. The fact that the synchronization is dependent on signal strength explains at least to some degree some of the synchronization problems encountered by the writer with operation of a DSB receiver. Even with AGC applied, some manual adjustment of RF gain was required to maintain proper synchronization for wide variation in signal strength. With the receiver operated at fixed gain, the tuning of the receiver became correspondingly more critical with a weaker signal, and, in listening tests, the weaker parts of syllables sounded not quite synchronized when the stronger parts were synchronized. It is the writer's experience, based on operation of just one DSB receiver, that the frequency of the reinserted carrier is required to be initially not over 100 cycles from the frequency of the suppressed carrier for reliable synchronization to occur, and the difference in frequencies was required at times to be somewhat less than this figure. Based on this, the required accuracy in the frequency of the reinserted carrier must be of the same order for DSB reception as for SSB reception.

The very considerable reduction in effects of selective fading obtained by the SSB receiver and the DSB receiver when receiving the SSB and DSB signal is also to be obtained by using these receivers for the AM signal. This reception of the AM signal by the SSB or DSB receiver could be a logical type of interim operation in the process of a transition from AM operation to SSB or DSB operation. It is not feasible, however, in the aeronautical mobile service because the frequency stability of transmitters in the aeronautical mobile service is not sufficient for the signal to be properly tuned by the SSB or DSB receiver.

FACTORS AFFECTING ACCURACY IN FREQUENCY OF REINSERTED CARRIER

The quality of the demodulated SSB and DSB voice signal is very good when the frequency of the reinjected carrier is sufficiently close to that of the suppressed carrier. Natural questions that arise are: What is the effect on the signal of error in frequency of the reinserted carrier, and, what is the allowable deviation? The one effect of error in frequency of the reinserted carrier on SSB reception that is well known, is that all modulation frequencies are translated in the same direction, as evidenced by a change in pitch of the modulation. The permissible error as regards the effect on intelligibility of the voice signal has in the past been a matter of controversy. For HF operation, it does not now appear to be of consequence because of the strides made in stabilizing the frequency of oscillators used in airborne equipment.

When the reinserted carrier for the DSB signal is not phase locked and is not at the proper frequency, the output of the detector contains two sets of modulation frequencies where normally there is one set, and one set is raised in frequency and the other set is lowered an equal number of cycles, equal to the error in frequency of the reinserted carrier. The phase of the modulation frequencies is such that the entire modulation signal "beats" down to zero amplitude at a rate twice the error in frequency of the reinserted carrier. The net effect is to impair intelligibility of the voice signal to a somewhat greater extent than the same error in frequency of the reinserted carrier impairs intelligibility of the SSB signal. In the presence of an error in frequency of reinserted carrier of as much as 30 cycles, the intelligibility of voice modulation in DSB reception remains fairly good. However, since the DSB receiver is designed to provide a phase lock, in which case the signal quality is very good, and this is the only way to avoid a whistle heterodyne when the signal happens to be an AM signal, it is desirable that the DSB receiver be operated in a phase-lock condition.

The frequency stability of the airborne crystal-controlled oscillator is today equal to what was obtainable a few years ago only in a fixed ground installation. The accuracy in frequency of one part in 10^7 to one part in 10^8 is obtained only by very close temperature control of the crystal. This is accomplished in present equipment by an electrically heated oven capable of holding the temperature of the quartz crystal to within 0.01°C of some fixed value. Such equipment does not have a 2-minute, or 5-minute warm-up time, but the warm-up time can be expected to be more in the order of 20 minutes. The very best frequency stability of which the equipment is capable requires a somewhat longer time than this. Since the oven is very well insulated and requires only a very few watts for operation, it may well be desirable in the aeronautical mobile service that the oven heater remain on when the equipment is turned off.

EFFECT OF SIGNAL CARRIER ON REINSERTED CARRIER OF DSB RECEIVER

The carrier of the AM signal tends generally to be ignored by the reinserted carrier of the DSB receiver. An exception is when an AM signal is modulated by a single tone, as a 1000-cycle test tone of a signal generator. The reinserted carrier will more readily lock at the frequency halfway between the side frequency and carrier than it will at the carrier frequency. This is due to the fact that this position is encountered first in attempting to tune the receiver. Therefore, this is only an apparent false capture of the reinserted carrier, in that the receiver is improperly tuned to one side of the signal band. When the receiver is tuned as to be properly centered on the signal, the reinserted carrier readily locks at the carrier frequency of the signal.

EFFECT OF TYPE OF MODULATION ON PEAK RF ENVELOPE

It is shown⁶ that quadrature components of the modulation sidebands, which cancel out in the undistorted AM signal, make their appearance when the sidebands become unbalanced, as in selective fading. They produce harmonic distortion of modulation in the AM detector, and they are in greatest amplitude when one sideband is completely missing. These same quadrature components combine in a surprising manner, most in evidence at the transmitter, in the particular case of square wave modulation of the SSB signal.⁹ All frequencies of the signal then add up to produce peak amplitudes of the signal envelope many times the amplitude of the square wave. Because of the overloading of the transmitter, and accompanying undesirable side effects, square wave modulation is practically ruled out in SSB operation. Speech clipping, which is effective in improving the speech level of the AM signal and DSB signals, tends to produce square-wave modulation, and, therefore, may be applied only to a very limited extent in SSB operation.

DATA TRANSMISSION, MULTIPLEXING, AND DIVERSITY RECEPTION

Data transmission, particularly since the advent of electronic computers, is becoming more often a type of signal sent over a radio communications system. However, a form of data transmission, teletype, has been sent over radio links for many years. The coding of the teletype signal for radio transmission has generally been similar to that used for land line operation. Since the signal starts off as typing at the rate of about 60 words per minute, and each letter is formed by a series of five on-off type signals, the rate of information signals plus synchronizing signals is about 40 per second. The fundamental frequency and essential harmonic frequencies of the modulation may be considered as extending to only

⁹ L. R. Kahn, "The use of speech clipping in single-sideband communications systems," *Proc. IRE*, vol. 45, pp. 1148-1149; August, 1957.

200 cycles. Using AM modulation (simple on-off keying of an RF carrier), the sidebands are so narrow that fading could be expected to be almost never selective of the frequencies in the sidebands, and, therefore, to be so-called "flat" fading. In this case, there would be no distortion of the signal other than by noise. There is a problem of AGC level of the AM teletype signal when fading is very fast. Under this condition, or when fading is very deep, or the signal very weak, the signal may become lost in the background noise or possibly otherwise improperly detected. Space diversity reception is then helpful, triple diversity being somewhat better than dual diversity.¹⁰

The transmission of a teletype message by AM modulation is obsolete, having been displaced by frequency shift keying (FSK). FSK may be displaced shortly by newer methods, but it will be discussed to some extent, leading to a discussion of other systems.

The process of improving signal-to-noise by increasing the bandwidth, as implemented by FM operation, has not generally been successful in the presence of selective fading because of the resulting severe distortion of modulation. However, in the case of FSK, it has provided an improvement over AM transmission. The relative performance of AM and FM with the on-off type signal (for FM, more properly termed "mark-space" keying) depends on many details of operation, but certain basic differences stand out. For one, FM is not much affected by the rate of fading. For another, FM provides frequency diversity.

An analysis of FM performance for the general case in the presence of selective fading would be exceedingly difficult. However, FSK is amenable to a simplified form of treatment by viewing the process as one of sending simultaneously two AM pulsed signals, one channel transmitting the marks, the other the spaces. In transmitting the mark-space signal by AM, a mark is known by the presence of a signal, and a space by the absence of signal, or vice-versa. Thus, by transmitting only the marks or spaces, both the marks and spaces are known; however, the entire signal is lost by a fade-out. FM transmission, by transmitting twice the minimum information required, provides a redundancy that increases the reliability of reception in the presence of fading. FSK operation is reported to have provided superior performance to AM operation when triple space diversity was used for AM operation and dual space diversity was used for FM operation.¹¹ A frequency shift of 850 cycles (deviation of ± 425 cycles) has been commonly used. Viewed as two separate "channels," the centers of these channels are then spaced 850 cycles. Since the sidebands of each channel may be considered

confined to ± 200 cycles, some inefficiency in use of the band is apparent because the center 450 cycles are not used. However, the frequency shift appears to have been chosen to permit signaling at a rate much higher than 40 per second, as in multiplexing several messages by time division. Nevertheless, the system was not capable of multiplexing as many messages as efficiently as other systems to be described.

It is recognized in information theory that the receiver should be in possession of the invariant part of the information transmitted. For FSK, this includes shape and duration of the pulse and accurate knowledge of the RF shift frequencies. This permits the receiver to better integrate the signal for the duration of the mark or space as an aid to identifying it as a mark or space in the presence of noise. The essential element used as integrator in the system called "predicted wave"¹² is a magnetostrictive resonator. The signal is permitted to enter the resonator at the beginning of the mark or space, to build up an oscillation during the pulse, when the signal amplitude is sampled and the resonator quickly freed of energy so as to be ready to sample another mark or space. As applied to FSK, one resonator is tuned to a mark, another to a space.

The magnetostrictive resonator is basically no more than the electromechanical equivalent of the simple resonant electrical circuit. However, the application to on-off keying is so simple and useful that it is important to understand the operation involved. Suppose the resonator to be of high Q in its basic design, and to have added to it the small positive feedback required to give it infinite Q . A sine wave sent into the resonator will produce an oscillation with zero crossings at a rate corresponding exactly to the resonant frequency of the resonator, even though the frequency of the applied wave differs from the resonant frequency. At the start of the baud the applied wave and the resulting oscillation are in phase. If the applied wave is at the resonant frequency, the oscillation will build up linearly. When the applied wave is not at the resonant frequency, the waves will become out of phase such that the oscillation eventually becomes zero. It is easy to see that the oscillation is reduced to zero at multiples of the time $1/F$, where F is the discrepancy in frequency of the applied wave and resonant frequency of the resonator. Thus, it is seen that the resonator has the dual quality of not only reinforcing a wave at its resonant frequency but in canceling out waves at regular frequency intervals of the resonant frequency. This immediately suggests a means of frequency division multiplexing since waves so spaced in frequency will not interfere with each other.

Viewing the means of frequency division multiplexing as made up of a group of individual AM channels as was done above, the centers of the channels would be spaced 40 cycles for a signalling speed of 40 per second

¹⁰ J. R. Davey and A. L. Matte, "Frequency-shift telegraphy—radio and wire applications," *Bell Sys. Tech. J.*, vol. 27, pp. 265-304; April, 1948.

¹¹ H. O. Peterson, J. B. Atwood, H. E. Goldstine, G. E. Hansell, and R. E. Schock, "Observations and comparisons on radio telegraph signaling by frequency-shift and on-off keying," *RCA Rev.*, vol. 7, pp. 11-31; March, 1946.

¹² M. L. Doelz, "Predicted wave radio teleprinter," *Electronics*, vol. 27, pp. 166-169; December, 1954.

per channel. Since the bandwidth of a channel operated alone is to be considered as about 400 cycles, considerable interleaving of frequencies of each channel occurs with neighboring channels. Therefore, multiplexing provides the means to realize a narrow bandwidth per channel. Other rather obvious reasons for multiplexing are as follows: 1) There happens to be enough teletype traffic with a certain geographical location to warrant sending simultaneously many messages; 2) the synchronizing of a number of messages may be accomplished with one synchronizing signal; and 3) the frequency control required for the subcarriers in SSB operation may be obtained simultaneously on all channels.

There appear to be no restrictions on multiplex operation of the channels other than that they be placed at the spacing producing no interference. For example, suppose magnetostrictive resonators to be used with FSK in which signaling is at the rate of 40 pulses per second and the frequency shift is 800 cycles. The band required to transmit this message is 1200 cycles. Now, placing at 40-cycle intervals other pairs of channels, each an FSK signal with frequency shift of 800 cycles, twenty simultaneous messages each signaling at the rate of 40 pulses per second could be placed in a band of 1960 cycles, for a band of 98 cycles per message. Since power is being transmitted on each AM channel only 50 per cent of the time on the average, the signal power may be increased by filling in the gaps of transmission with a signal displaced 180° in phase (each channel is then phase modulated), provided a method of detection is used that can distinguish the phase. The system of data transmission known as Kineplex employs phase modulation, infinite Q resonators, and frequency division multiplexing in a manner resembling to some degree what has been described.

In the discussion of multiplexing mark-space information by frequency division multiplexing, it is not intended to suggest or recommend any particular method but rather to show the versatility inherent in multiplexing and to provide some insight into the possibilities for devising very many systems. It is to be noted that a certain minimum frequency spacing of a pair of channels is necessary for good frequency diversity. The best frequency diversity occurs if one channel is at a condition of maximum reinforcement of the signal when the other is in a condition of greatest cancellation. For example, when multipath propagation is by two paths differing 93 miles in length, the best frequency diversity occurs for channels whose centers are spaced 1000 cycles. Such a path difference is associated with a range of transmission that is at least several hundred miles and likely several thousand miles. Therefore, for even the very long range transmission, a channel spacing of several hundred cycles would be required for good frequency diversity. This brings to attention the fact that frequency diversity is not possible at all for flat fading, whereas space diversity is effective for both flat fading and selective fading.

The AM, FM, and phase modulation mark-space signal sent as a single radio message, comprises both an upper and a lower sideband. When removal of one sideband and the carrier is considered as a means to conserve frequency spectrum and transmitter power, the FM signal becomes an AM signal. The AM signal and phase modulation signal reduce to the single sideband square-wave modulated signal, which was found above to be unfeasible, because of the excessively high peak envelope power required. For the mark-space signals multiplexed by frequency division, the carrier of each channel is, of course, a subcarrier, and the main system of modulation may be selected independently. Single sideband operation is the type most conserving of the frequency spectrum and, in this case, is feasible because the composite modulation of many messages departs so far from square-wave modulation.

Irrespective of the main type of modulation used with multiplexing of the mark-space signal by frequency division, space diversity provides a great increase in reliability of the signal, just as it has in the past for transmission of only one message. In the absence of space diversity, and assuming selective fading, the fading effects are as follows:⁶ 1) Using AM for the main modulation, all channels tend to fade simultaneously. 2) For SSB operation, the channels tend to fade individually or in groups, the fading skipping momentarily from channel to channel. 3) For DSB operation, the fading tends to be confined mostly to certain channels or groups of channels. It appears, therefore, that DSB operation would be desirable when the required antenna spacing for space-diversity reception, or antenna polarization for polarization-diversity reception,¹³ is not available, as in reception aboard an aircraft. The strong channels could be expected to remain strong for a matter of hours, during which time the weak channels could be put out of operation.

Another advantage of DSB operation is that the close control of frequency of the sub-carriers required at the receiver, to be discussed next, would be obtained automatically.

A requirement generally accompanying efficient systems of transmission and detection is that of accurate control of frequency of the transmitted signal, and accurate tuning of the receiver. For the type of multiplexing that has been described, using channels only 40 cycles wide, the desired accuracy in frequency of the radio signal is about 2 parts in 10^8 . This apparently cannot be considered a requirement peculiar to the infinite Q resonator. Another attempt at optimum detection of the data signal is in the use of a delay line.¹⁴ It is claimed to be superior to the infinite Q resonator in resisting the

¹³ G. L. Gridale, J. G. Morris, and D. S. Palmer, "Fading of long-distance radio signals and a comparison of space-and polarization-diversity reception in the 6-18 mc/s range," *Proc. IEE*, vol. 104, pp. 39-51; January, 1957.

¹⁴ F. A. Losee, "A digital data transmission system using phase modulation and correlation detection," *Proc. 1958 Aeronautical Electronics Conf.*, Dayton, Ohio.

effects of selective fading. This device also requires very close control of the subcarrier frequency.

COMBINATIONS OF SYSTEMS

Because of the limitations of AM operation in HF communication, a change-over to some other system or systems appears inevitable in some classes of service. It has appeared that some special interim type of operation will be necessary during the period of change-over. There are many possible means of providing compatible operation of the new equipment with present AM equipment in the aeronautical mobile service. There is a choice of applying it to the equipment in use or to the equipment of the new system, applying it to the transmitter or the receiver, applying it to the airborne or the ground equipment, and a choice in selecting the means employed in each instance.¹⁵ Certainly the ultimate selection of method will be greatly influenced by the results obtained in service use of equipment. Of the systems proposed to replace AM operation, SSB operation has been used the most, but this has been to a comparatively limited extent in the aeronautical mobile service. If a trend can be said to have developed, it appears in the case of new airborne SSB transceiver combinations to be to provide for both AM or SSB operation selectable on a panel control. Present specifications for SSB equipment are likely to contain a provision for alternative operation with a carrier present, so that the signal may be received on an ordinary AM receiver. However, considerable distortion of modulation results at the receiver in this mode of operation from the quadrature component of the sideband frequencies. Kahn has devised a system called Compatible Single Sideband

(CSSB)³ in which he provides a full carrier but predistorts the sideband in a manner that *removes* distortion from the envelope and thus cancels out the distortion of modulation produced in the AM receiver. Because of the similarity of the equipment to produce this signal to the equipment used by Kahn to produce the SSB signal by the method of envelope elimination and restoration,⁵ it appears that not much more equipment is required to permit a choice of operation by either system, as compared to one system alone, so that provision for a change in operation by means of a panel control is feasible.

In the literature, the advantages and disadvantages of various types of carrier transmission with the SSB signal are discussed.¹⁵ Because of the problems of AGC and radio noise with SSB transmission, cited above, a short discussion of the effect of the carrier in this regard is included herein. In the presence of selective fading, the amplitude of the carrier transmitted with a sideband, at any given moment, bears no relation to strength of the sideband, as received. Since the relation of the carrier to sideband at the receiver is a statistical one only, the carrier appears to be of questionable value for AGC in SSB operation. It follows also that the signal consisting of a sideband and carrier will be very prone to overmodulation when received by the envelope detector in the presence of selective fading, unless the percent modulation at the transmitter is low. This is not a basic limitation in that the condition can be largely corrected by enhancement at the receiver of the transmitted carrier by what has been termed "exalted carrier."¹⁶ This, however, does not aid in the problem of AGC.

¹⁵ J. F. Honey, "The problems of transition to single-sideband techniques in aeronautical communications," *PROC. IRE*, vol. 44, pp. 1803-1809; December, 1956.

¹⁶ M. G. Crosby, "Exalted-carrier amplitude- and phase-modulation reception," *PROC. IRE*, vol. 33, pp. 581-591; September, 1945.

Radiation Measurements at Radio Frequencies: A Survey of Current Techniques*

W. A. CUMMING†, SENIOR MEMBER, IRE

Summary—A survey is given of techniques used in the measurement of radiation fields at radio frequencies. These measurements are divided into three broad classifications dealing with fundamental studies of electromagnetic waves, antenna design, and antenna performance. These classifications in turn include diffraction, scattering, transmission and reflection, current distribution, aperture fields, radiation patterns, and gain.

I. INTRODUCTION

THE detection and investigation of electromagnetic waves at radio frequencies had their beginnings before 1900, in the early experiments of Hertz, Marconi, Righi, and others, whose work has recently been reviewed in an excellent paper by Ramsay [1]. While Maxwell, in 1864, had shown theoretically the electromagnetic nature of light waves and had predicted similar electromagnetic phenomena at lower frequencies, it was not until Hertz, in 1888, was successful in generating, transmitting, and detecting such waves, that interest in what we now call "radio frequencies" became general. The early experimenters focused their attention on the investigation of the wave nature of the radiation, and on its similarity to radiation at optical wavelengths. Having devised means of generating and detecting electromagnetic signals at radio frequencies, their experiments were concerned chiefly with developing means of effectively launching waves into space in a concentrated beam, and determining the effects of various obstacles on their transmission. This emphasis on the external or radiated field was a logical one, not only because it paralleled optical experience, but also because the more interesting applications of radio frequencies depend on waves propagating in space. The evolution of present-day RF technology from these early beginnings some sixty years ago has depended to a large degree on a continuing emphasis on the study of both radiated and guided electromagnetic fields. Now, as then, the actual detection and measurement of the radiated fields play an important part in the understanding and application of electromagnetic waves, and it is to review this phase of the art that this paper is written. In particular, it deals with the measurement of the radiation fields associated with antennas, but excludes propagation experiments, as well as impedance and other measurements which are not peculiar to a radiating system.

Radiation measurements play an important part in fundamental investigations of the behavior of elec-

tromagnetic waves, in antenna design procedure, and in the testing and assessment of antennas and radiating systems. While the basic quantity measured in each case is usually the electric or magnetic field at different points in space, the measurement techniques vary widely, and depend on the extent of the region to be explored, the manner of presentation of the result, and whether or not the phase and state of polarization of the field vector are required, in addition to the amplitude. Furthermore, some investigations require a knowledge of absolute field strength, or at least a knowledge of the relation between the field strength and the power being radiated by the source; while others require only a knowledge of relative field strength. A rigid separation of the measurements to be discussed into the three classes mentioned is difficult, and has no particular significance; however, diffraction and scattering by various obstacles, and reflection and transmission at boundaries between materials have arbitrarily been termed "fundamental phenomena." Similarly, the measurement of antenna aperture and near-zone field distributions, and the measurement of current distributions, have been classed as pertaining to antenna design. Finally, the determination of radiation patterns, gain, polarization, efficiency, and the effects of radomes, have been included in those measurements which are made to assess the antenna as part of a system. In each case, the description of the measurement deals with the method of sampling the field and with the instrumentation required to produce the RF signal and to record the output and position of the sampling device.

II. DIFFRACTION AND SCATTERING

Consider a region in which an electromagnetic field has been set up by an arbitrary source. Let this field be characterized by the electric vector, E_i . If an obstacle is now introduced into the region, displacement currents and conduction currents will be induced in such a manner as to produce fields which, together with the original field, satisfy the boundary conditions on the surface of the obstacle. If the source and obstacle are sufficiently far apart that these secondary waves do not interact with the source, we may then write

$$E_t = E_i + E_s$$

where E_t is the total or diffracted field, E_i the initial or incident field in the absence of the scatterer, and E_s the secondary or scattered field. In cases where the scattered field interacts with the source, the equation must be modified so that E_i represents the initial or incident field in the presence of the obstacle.

* Original manuscript received by the IRE, February 13, 1959.

† Radio and Elec. Eng. Div., Natl. Res. Council of Canada, Ottawa, Ont., Can.

Since the calculation of either the scattered field or the diffracted field involves the solution of a boundary-value problem, it is of great interest to the theoretician. However, due to the complexity of the problem, only certain geometrically-shaped obstacles yield rigorous solutions. Thus measurement of the scattered and diffracted fields is necessary, not only to verify the theoretical calculations in the case of simple obstacles, but because it is the only means of obtaining the result in the case of complex obstacles.

Scattered and diffracted fields are of more than academic interest, since they are intimately connected with the operation of antennas and radiating systems. For example, consider a source field in the form of a plane wave, and an obstacle in the form of an infinite sheet placed normal to the wave, and in which is cut a narrow half-wavelength slot. The solution of this fundamental boundary value problem gives a diffracted field whose variation with angular position on the shadow side of the plane and at a great distance from the opening, represents the pattern of the well-known half-wave slot antenna. In fact, determination of the patterns of many microwave antennas involves the solution of a diffraction problem in one form or another.

In practice, the scattered field *per se* is not usually of interest so far as an antenna alone is concerned, but is often involved in the analysis of radiating systems. An example is the analysis of a radar system in which a knowledge of the field scattered from a target is of prime importance. Certain special antenna problems are also best solved by considering the scattered field separately. To illustrate, the pattern of a parabolic cylinder illuminated by a line source can be arrived at in two ways. In both methods the unperturbed aperture field which would exist if the feed system were not present to obstruct the emerging energy is first calculated. Then in one method, the total or diffracted field in the aperture plane is determined, from which the far-zone pattern is calculated. In the other and simpler method [2], the far-zone pattern of the unperturbed aperture distribution is first calculated, and on this is superimposed the scatter pattern of the feed system.

A great many means have been devised for measuring scattered and diffracted fields; a recent survey paper by King and Wu [3] outlines over a dozen methods which have been used at Cruft Laboratory alone. Since the measurement of the total or diffracted field is basic and involves techniques common to other measurements, it will be discussed first.

A. Diffraction Measurements

The diffracted wave resulting when an obstacle is placed in a source field is described by the electric or magnetic field strength, as defined by amplitude, phase, and polarization, at various points in space. Generally it is required to determine the field at distances from the obstacle varying from a small fraction of wavelength to no more than ten wavelengths. The problems in meas-

urement are: to produce a suitable source field, to devise a means of supporting the obstacle within this field, to sample the total field by a probe or similar device whose presence does not disturb the field, and finally, to measure and record the probe response as a function of its position. In general, the probe response is interpreted to include amplitude, phase, and polarization information. It is not feasible to construct a single instrument suitable for all types of diffraction measurements; in practice, the apparatus used depends very much on the type of obstacle to be investigated, on the polarization of the source field, and on the amount of information required.

First there is the parallel-plate type of instrument [4], one form of which is shown schematically in Fig. 1,

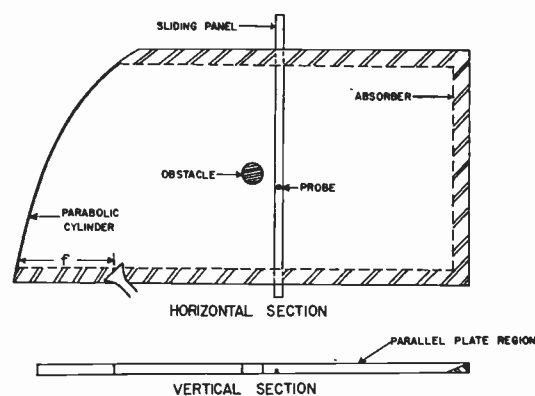


Fig. 1—A parallel-plate type of diffraction-measuring equipment.

and which is limited in its use to the study of cylindrical obstacles of infinite equivalent length, located in fields where the electric vector is parallel to the cylinder axis. The plate spacing is less than $\lambda/2$ to limit propagation to the TEM mode, which is launched from a small horn projecting into the parallel-plate region, as shown. Since most applied diffraction problems involve an obstacle placed in a uniform, plane field, the energy from the horn is collimated by a parabolic reflecting strip, so that an essentially plane wave is launched into the main parallel-plate region. The width of the region must satisfy two conditions. First, it must be wide enough so that over the transverse region of interest the variation in incident field lies within prescribed limits. For example, suppose the object to be studied has a transverse dimension of one wavelength, and that it is desired to examine the diffracted field across a strip 8 wavelengths wide. Assuming the transverse variation of the field produced by the reflector to be sinusoidal with a 10-db cutoff at the edges, the width of the parallel-plate region must be 80 wavelengths in order that the field variation be less than 0.25 db over the 8-wavelength-wide central section. The second consideration in the choice of width is that scattered energy from the obstacle must not be reflected from the boundaries of the parallel-plate region and returned to the obstacle with sufficient amplitude to distort the measurement. The

sides of the parallel-plate region are terminated in lossy material to absorb the scattered energy, but nevertheless a reflection of perhaps -20 db is likely to occur, and in addition a portion of the scattered energy will strike the parabolic strip and be reflected back into the parallel-plate region. For this reason, the width of the plates and the position of the obstacle forward of the focal plane must be chosen to provide reasonable space attenuation of these reflections. Returning to the example, the width of 80 wavelengths arrived at previously provides some 19 -db additional space attenuation between the scattered field in the vicinity of the obstacle and the field at the same point, caused by a specular reflection of the scattered field from the sides of the region.

The effect of a reflection from the parabolic strip is more difficult to assess. However, on a simple ray-theory basis it can be shown that if the obstacle is placed at least a distance f in front of the focal plane, where f is the focal length, the energy reflected from the strip will not reach the region of interest except by multiple reflections. This requires in the present case that the obstacle be located at least 40 wavelengths, and preferably 50 wavelengths, in front of the focal plane. The length of the parallel-plate region beyond the obstacle can be arbitrarily arrived at by again demanding space attenuation of the order of 20 db for a scattered signal specularly reflected from the end termination, which in this case results in a length of 40 wavelengths.

To summarize, a parallel-plate region of width 80λ by length 90λ , excited by a semiparabolic cylinder of focal length 40λ , is required in order to produce an incident field which is uniform in amplitude to within $\frac{1}{4}$ db, and sensibly plane in phase over a transverse dimension of 8 wavelengths; and in order to insure that scattered energy returning to the test area via reflections from the boundaries of the parallel-plate region is down some 40 db from the initial scattered wave.

It is evident that with this type of excitation of the parallel-plate region it is relatively simple to obtain an incident wave which is uniform in phase over the test area, but it is difficult to obtain a uniform amplitude distribution and to prevent the scattered field being reflected back into the parallel-plate region by the source.

Other forms of excitation are possible, each having advantages and disadvantages. For example, a waveguide array with slots in the broad face might be used, and would permit a considerable degree of control over the amplitude distribution. However, it would give rise to very large source reflections, and would therefore require large separation of source and obstacle. A second alternative is to use a small horn placed on the center line of the parallel-plate region, and to correct the resulting cylindrical wavefront by a dielectric lens. With this configuration, reflections are considerably reduced, but either a large lens or a large lens-to-obstacle spacing is required if phase and amplitude variations are to be held to satisfactory levels. Finally, a small

horn can be used without a corrective lens. This is the simplest feed to construct, and produces the smallest source reflection, but requires very large separation of the source and obstacle in order to obtain a uniform phase front. A common practice is that of using a separation sufficient to give a reasonable amplitude distribution, applying a correction to the measurements to take into account the cylindrical phase front. For example, a typical system might employ a ground plane of dimensions 100λ by 50λ , with a source-to-obstacle spacing of, say, 50λ . With these dimensions, a feed horn could be designed to produce a satisfactory amplitude distribution over a region some 8 wavelengths wide.

Since they are convenient to construct and mount, electric-field probes are used with parallel-plate diffraction-measuring instruments. Such a probe consists of the extended inner conductor of a coaxial line, which is connected to a coaxial or waveguide-tuning device, and thence to a receiver. Ideally, the probe should sample the electric field within the parallel-plate region without causing any disturbance to that field. This is theoretically and practically impossible, since the probe itself gives rise to scattered fields necessary to satisfy the boundary conditions on its surface. However, these scattered fields can be kept to an acceptable value by limiting the probe length to less than $\lambda/4$, by tuning the coaxial line leading to the receiver to a high input impedance, and by restricting the spacing between the probe and the obstacle—usually to one half or one wavelength.

In order to measure the diffracted field at all points within the parallel-plate region, provision would have to be made to move the probe at will throughout the entire region surrounding the obstacle. Because of the difficulty involved, a simpler system is usually employed in which the probe can be continuously moved along transverse lines at various distances from the obstacle. This is accomplished by mounting the probe and its associated coaxial line on a movable panel which is inserted in a slot in one of the parallel plates. Since the slot is transverse and the flow of surface currents on the plates longitudinal, care must be exercised to obtain good contact between the sliding panel and the plate, otherwise a further field perturbation will be introduced.

In the general case, the diffracted field is described by its magnitude, phase, and polarization, as a function of position. In the parallel-plate type of instrument, only one component of polarization is possible, so that one of these variables is eliminated. Since the probe-output voltage is proportional to the field, means must be provided to determine and record the amplitude and phase of this voltage as a function of probe position. In the last few years, considerable effort has gone into developing the necessary microwave and electronic circuitry necessary to accomplish this, either manually or automatically, and a recent report by Bekefi [5] contains an excellent bibliography on the subject. The art of phase and amplitude measurement at microwave fre-

quencies is highly developed; in fact, a limited variety of instruments is now available commercially.

Determination of the phase of the probe-output voltage is far more difficult than determination of the amplitude, and will be dealt with first. All methods employed at microwave frequencies are based on comparison of the phase of the probe signal with that of a known reference signal, and make use of a calibrated phase shifter of known characteristics. The simpler systems involve a direct comparison between the probe and reference signals, either at the signal frequency or at a lower frequency obtained by heterodyning. The relative simplicity of these systems is, however, offset by the fact that for accurate results, the probe and reference signals must be adjusted to have approximately the same level, which makes the systems undesirable for automatic operation. More elaborate techniques which lend themselves to automatic instruments involve modulating one or both of the signals and combining them in such a way that the resultant signal has a measurable property which is dependent on the relative phase of the individual signals, but independent of their relative amplitudes.

A fully automatic instrument using the simple comparison method was described by Barrett and Barnes [6] and involves the conversion of both reference and probe signals to an intermediate frequency of 30 mc. Limiting amplifiers at this frequency are used to maintain equal amplitudes of the two signals, which are then mixed in a phase detector whose output is used to position the reference phase-shifter.

One of the more sophisticated systems makes use of the homodyne principle [7], in which the probe signal is modulated at an audio frequency by a balanced modulator to produce a double-sideband output with suppressed carrier. The reference signal is passed through a variable phase-shifter, mixed with the probe signal and detected. The audio signal from the detector has two components: one at the fundamental modulating frequency, and one (which is filtered out) at twice the modulating frequency. The fundamental component can be shown to vary in amplitude as the cosine of the reference signal-phase angle, passing through zero when the reference signal is in quadrature with the original carrier. Since the null condition is independent of the relative amplitudes of the probe and reference signals, the detector output can be fed into a servo system and used to control the reference phase-shifter. However, since the amplitude of the detector fundamental output for unbalanced conditions is not independent of the relative amplitudes of the probe and reference signals, the servo-loop gain must be provided with automatic gain control.

A block diagram of such a system [5] is shown in Fig. 2. The microwave source is a CW klystron whose frequency stability need not be great, provided the reference and probe paths are made of approximately equal length. In this connection it should be pointed out that

for accurate measurement at microwave frequencies, the transmission line linking the probe to the mixing hybrid must not be coaxial line because of the changes in phase-shift which occur as this type of line flexes. Therefore, in order to permit motion of the probe, the connection must be made with waveguide-transmission line into which are incorporated two rotating joints whose phase-shift is independent of rotation. This calls for extreme care in the design of the rotating joints. The ferrite type of balanced modulator is preferred since the earlier crystal modulator is not stable, because of variations in the RF impedance of crystals. Many types of phase-shifter may be used in the reference branch, but the rotary phase-shifter is well suited because of its linear characteristic and unrestricted travel. The addition of the two signals in the phase-measuring circuit requires particular attention to insure adequate isolation between the channels. This is achieved by feeding the signals into the E and H arms of a magic-T and adding the outputs of matched crystal detectors placed on the other two arms. It will be noted that because of the nature of the signal coming from the probe, the output of the amplitude detector is an audio frequency twice that of the modulating frequency.

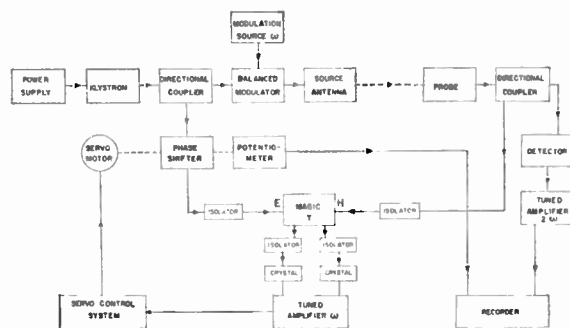


Fig. 2—Phase and amplitude measurement using the homodyne technique (after Bekefi).

Both the magnitude and phase of the probe voltage can be recorded automatically on graphic recorders whose detailed operation is discussed in Section V. The amplitude is most conveniently recorded on a logarithmic recorder, while the phase is plotted on a linear recorder whose pen servo is fed from a linear potentiometer linked to the variable phase-shifter. The charts of both recorders are positioned by servo systems whose inputs are a function of the probe position.

While the homodyne system has been described as suitable for automatic operation, other means are available. For example, if either the reference or probe signal is phase-modulated [8] at an audio rate, then the resultant signal obtained when they are combined will be amplitude-modulated. If this signal is detected, the resulting audio signal will have a phase relative to the original modulating signal which is the same as the phase difference between the two microwave signals, but which can be more readily measured. Phase modula-

tion of the reference channel can be achieved easily by the use of a rotary phase-shifter driven at an audio rate.

Single-sideband suppressed-carrier modulation [9] can also be used in place of phase modulation or the homodyne technique. In this system the resultant of the modulated and unmodulated signals, when detected, provides an audio signal of frequency equal to the difference between the sideband and carrier. Again the phase of this signal relative to the modulating signal is equal to the phase difference of the RF signals.

Finally, the coherent detection system [10], [11] provides another means of measuring the phase difference between two signals of unequal amplitude. A greatly simplified block diagram of a phase and amplitude recorder based on this principle is shown in Fig. 3. For phase measurements a reference signal derived from the RF source through a directional coupler is fed through a precision phase-shifter and thence to the *E* arm of a magic-T. The voltage from the probe is fed into the *H* arm of the magic-T as shown. The remaining arms of the T are identically terminated in matched crystals or bolometers whose outputs are fed into a bridge circuit, which measures their difference. A study of the vector relations in the magic-T circuit shows that the bridge output or difference voltage is zero when the signals fed into the magic-T differ in phase by 90° , regardless of their amplitude. For manual operation, it is necessary only to adjust the phase-shifter for zero output at each probe position, and to read the phase change introduced by the phase-shifter. For automatic operation, the bridge output can be used as the error voltage to operate a servomechanism which drives the phase-shifter to a balance.

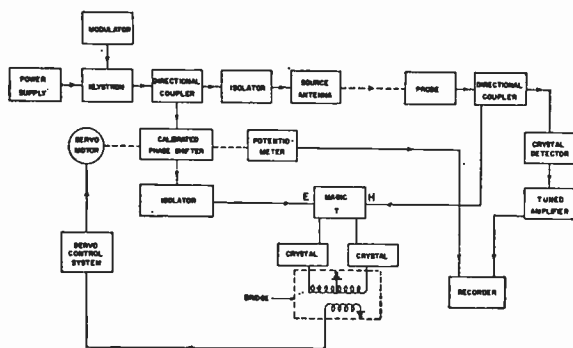


Fig. 3—An automatic recorder for phase and amplitude measurements.

For amplitude measurements, a portion of the probe signal is taken off through a directional coupler, detected, fed to an amplifier and then to a meter or automatic recording device.

As with the homodyne system, the accuracy of the equipment is extremely dependent on the known properties of the phase-shifter and on the constancy of electrical path length in the two channels. With reasonable precautions, phase measurements at *X* band can be made to about $\pm 0.5^\circ$, provided the probe can be posi-

tioned to an accuracy of approximately ± 0.001 inch.

The second type of diffraction-measuring equipment makes use of a single-image plane [19], and is suitable for objects having one plane of symmetry. For example, it lends itself to the study of fields diffracted by spheres, ellipsoids, spheroids, and cylinders, but is confined to a study of the normal field in the plane of the sheet. The design of the image plane and of the exciting antenna differs from that of the parallel-plate system for two reasons. First, the incident field in the vicinity of the obstacle must be uniform in amplitude and phase, not only in the direction parallel to the sheet, but normal to the sheet as well. Second, the source and scattered fields are only partially confined; therefore, attention must be given to the question of spurious reflected and scattered fields. These two factors demand careful design of the vertical aperture of the source antenna, for if it is made too directive, the first condition is not met, while if the directivity is too low, stray reflections become serious. In the case of line-source antennas, which have been described previously, vertical directivity is obtained by adding an *E*-plane horn, of either conventional design or lens-corrected. Similarly the point-source antennas are replaced with pyramidal horns, which are either used alone or to feed a two-dimensional dielectric lens.

The problem of reducing spurious fields is a difficult one, and can be solved either by placing the equipment in a microwave anechoic chamber of the type described in Section V, or in the open. The latter solution has serious limitations in that the measurement is dependent on the weather, while the former has the disadvantage of high cost.

The diffracted fields of three-dimensional objects having no plane of symmetry can be measured only in a free-space system; *i.e.*, one in which the obstacle is suspended in the field, remote from interfering objects. Such measurements are difficult, and are usually used only when image-plane techniques are not applicable. Spurious responses from all directions must be guarded against, including response from the structure used to support the obstacle, and from the microwave and electronic components which cannot be "hidden" behind a metal screen as is the case with the other methods. The design of the probe and the probe-positioning device is also more difficult, since the probe must be a balanced device, and since the field distortion due to the presence of the probe-supporting structure is far greater than that due to the probe itself. Finally, it is usually necessary to use flexible cable to connect the probe to the detector because of the length and complexity of a waveguide-transmission line. This limits the accuracy of phase measurements, as has been pointed out earlier.

One such free-space apparatus is described by Hamren [20], who did a series of measurements on spheres at *X* band. As a source antenna he used a paraboloid of aperture 30 inches, located some 200λ from the obstacles which were hung by threads from a wooden supporting structure. The field probing was limited: 1) to an ex-

ploration of the field along the longitudinal axis using a dipole and reflecting disk; 2) to an exploration in the transverse dimension using the same probe, and 3) to an exploration on the arc of a circle centered on the obstacle, using a horn probe. Flexible cable was used to connect the probe to a magic-T mixer in which the probe signal was compared with a reference signal whose amplitude and phase could be adjusted. The entire equipment was manually operated. Hamren not only measured the diffracted field in amplitude and phase, but for each probe position he also measured the amplitude and phase of the incident field, thus enabling him to subtract the two to obtain the scattered field.

A completely separate class of measurements involves the diffracted fields produced by two-dimensional obstacles consisting of thin conducting sheets of various configurations. Included are the knife-edge or half-plane, the half-wave slot, the long slit, and any other opening in an infinite sheet. The complementary obstacles, *i.e.*, the thin strip dipole, the long thin strip, the thin circular disk, etc., can be included with the obstacles already discussed insofar as measuring techniques are concerned, but openings in large metal plates generally require modifications to these techniques. To illustrate, consider a small circular hole in a thin infinite sheet. On the illuminated side there is a complete reflection of the incident wave over all regions except in the vicinity of the hole. This gives rise to interaction effects with the source antenna, including changes in its impedance. On the shadow side, the diffracted field, at least close to the plane of the sheet, is confined to the immediate vicinity of the hole, thus restricting the area to be explored. In many cases it is the field in the plane of the opening that is required, and this presents problems with regard to interaction between the probe and the screen.

Certain types of openings can be investigated using techniques essentially the same as those already described. For example, the half-plane illuminated by a field parallel to its edge can be conveniently studied in a parallel-plate region, as can the infinite slit in which the field is parallel to the slit [21], [22]. However, in both cases a strong reflected wave is caused by the conducting sheet, and interacts with the source. In most cases this difficulty can be resolved by limiting the size of the area to be probed, making the source-sheet spacings as large as possible, and by careful design of the source antenna. The question of coupling between the probe and the sheet or edge of the sheet, for close spacings, is a controversial one and merits further study. Down to approximately $1/10\lambda$ there appears to be little difficulty, but for closer spacings a variety of results has been reported by different authors [3].

A second class of openings, which have a line of symmetry and are assumed to be polarized perpendicular to this line, can be studied using the image-plane technique. This class includes the half-wave slot and the long slit with transverse polarization. In this method

the "infinite" sheet is presumed to be cut along the line of symmetry of the opening, and the resulting half-plane which contains one half of the opening is fastened normally to the image plane. The dimensions of the sheet are chosen so that the field arriving at the test area by diffraction around the sheet is negligible. This is easily checked by measuring the diffracted field with the aperture both opened and closed. In addition, the usual precautions with respect to reflection and interaction must be observed.

Free-space diffraction measurements involving circular, rectangular, elliptical, or irregular openings whose dimensions may be large in both the magnetic and electric planes, have received a considerable amount of attention in recent years [24]–[26]. They are similar in principle to the free-space diffraction measurements already described, but differ because of the presence of a large conducting plane which makes it easily possible to locate much of the microwave and electronic gear in the region of low field in the shadow side of the plane. This advantage is offset by the effects of interaction between screen and source, and between screen and probe. The former may be overcome either by using a large separation between source and screen, or by using as a source a large aperture antenna, such as a lens illuminated by a horn, whose collimated beam has a width approximately equal to that of the opening in the sheet [28], [29]. The latter disadvantage can be overcome only by limiting the minimum separation between the probe and the sheet, and by making the probe quite small. In cases where amplitude measurements alone are required, one technique is to probe the field directly with a crystal from which the brass mounting-tip has been removed. In cases where phase is also required, the probe design is much more complicated. Since these measurements are made with one of the phase-measuring techniques that have been described, it is necessary to feed the probe output into a hybrid-T. The hybrid must be kept clear of the field in the aperture, so a short length of miniature coaxial line is often used to link it to the probe. Electric-field probes are usually employed, because of the difficulty in obtaining from a loop a response that is entirely due to the magnetic field linking the loop [30], [31]. It is necessary that the probe be short (of the order of $\lambda/2$, or less) so that, in effect, it measures the field at a point rather than the average field over its length. This is particularly important when measuring the *E*-plane distribution near the edge of the sheet [32].

B. Scattering Measurements

An obstacle may scatter energy in all directions when placed in an electromagnetic field. If the obstacle is of finite length then, at great distances, the scattered energy is contained in a spherical wave whose amplitude and phase as a function of position represent the three-dimensional scatter pattern of the obstacle. For an obstacle infinite in length, the scattered energy is contained in a cylindrical wave, so the scatter pattern is a

function of position in one plane only. The chief difficulty encountered in measuring the scattered field lies in the fact that it must somehow be separated from the incident field. It can be measured indirectly, *i.e.*, by first measuring the incident field, then measuring the total or diffracted field, and finally subtracting the two, point-by-point, to obtain the scattered field. It can also be measured directly by various means which permit the separation or identification of the scattered component of the total field. Because of the difficulty of the measurement, it is very seldom that the complete scatter pattern for a given direction of arrival of an incident wave is determined. In both fundamental and applied studies, it is customary to measure the scattered field in a given direction relative to the direction of arrival of the incident wave, for various aspects of the obstacle. Furthermore, in both fundamental and applied studies it is customary to measure only the amplitude of the scattered field, and to relate it to that of the incident field in terms of the radar cross section, defined as follows. Consider a plane wave E_i incident on an obstacle from direction θ_i, ϕ_i . Let the scattered field in the direction θ_s, ϕ_s be $E_s(\theta, \phi)$. Then the radar cross section in that direction is given by

$$\sigma = \frac{p_{\text{isotropic}}^s}{S_i}$$

where $p_{\text{isotropic}}^s$ is the total power scattered by a hypothetical isotropic scatterer, producing a field strength E_s , and S_i is the Poynting vector associated with E_i . If the direction θ_s, ϕ_s coincides with that of the incident wave, as is the case in a monostatic radar system, σ is called the back-scattering cross section. This is the case of usual interest, except when dealing with bistatic radars in which case a knowledge is required of the radar cross section in a direction differing from that of the incident wave.

The measurement of either of these cross sections requires that the obstacle be placed in an incident field, and that the scattered field in the direction chosen be measured as a function of the aspect of the obstacle. Furthermore, the result can be most readily given as a radar cross section if the system is first calibrated in terms of a geometrical object of calculable cross section. Many obstacles of geometrical shape have planes of symmetry, or can be assumed infinite in one dimension, and so can be investigated on an image plane or on a parallel-plate region. These have the advantage of providing a field-free region in which auxiliary equipment can be placed, but unfortunately energy scattered from the edges of the ground-plane or parallel-plate system contributes to the background or residual field. This background signal is more troublesome than it is in the case of the diffraction measurements described previously, since the scattered field or echo area is usually measured in the far zone of the scatterer, and thus the desired quantity is much smaller.

One of the most commonly used techniques, termed

the CW balanced-bridge method [33] employs a single antenna and is shown schematically in Fig. 4. Energy from a carefully stabilized source is fed into the *H* arm of a magic-T. This energy divides between the side arms, one of which is terminated in a matched load, while the other is connected to a common transmitting and receiving antenna. The *E* arm is connected to a crystal detector through a ferrite isolator and ferrite modulator, which serve respectively to isolate the T from changes in detector impedance and to modulate the scatter signal. The detector output is then amplified and recorded in the usual manner on a logarithmic recorder.

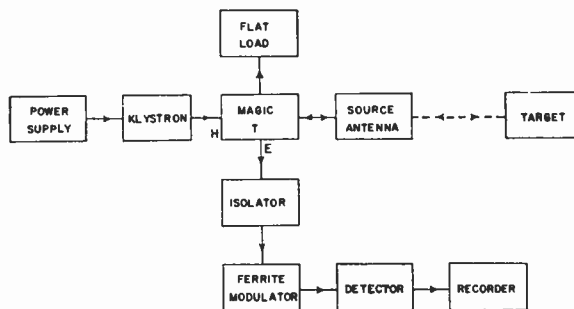


Fig. 4—A CW method of measuring back-scattered radiation.

In operation, the apparatus is first set up with no obstacle present. The magic-T is adjusted to produce minimum detected signal, a procedure which in effect uses leakage from transmitter to receiver through the T to cancel the effect of spurious fields scattered from nearby objects and from the edge of the ground plane. An object whose back-scattering cross section can be calculated, as for example, a hemisphere or cylinder, is then mounted on the ground plane, and the received signal noted. This effectively calibrates the instrument so that the cross section of other objects can be obtained directly. This type of instrument suffers from many limitations. First, the source-to-obstacle spacing is usually fairly large—perhaps 50λ —and is arrived at by the same considerations as were discussed when dealing with diffraction measurements. As a result, the ground plane must be of the order of 50λ by 100λ , and there may be path differences as great as 200λ between the signal reaching the detector by means of leakage through the hybrid and the signal reaching the detector via spurious reflections. Since these signals are cancelled out in the balancing procedure, extreme care must be taken when stabilizing the frequency of the source. In practice, automatic frequency control circuits are needed, together with temperature-controlled oil baths for the oscillators. Under typical conditions the undesired background signals are about 100 to 110 db below the transmitted signal, while the average signal scattered from a target is 70 to 80 db below this level. It is seen that very careful tuning of the hybrid is required in order to keep the resulting base level to an acceptable value, and this can usually be achieved for only short periods of time.

A variation of this method involves the use of a separate transmitting and receiving antenna [34]. Direct coupling between the two can be made quite low, and coupling by means of ground-plane reflections can be cancelled out by introducing a small amount of transmitter power into the receiving channel ahead of the modulator, through a directional coupler, an attenuator, and phase-shifter. This does not overcome the difficulty of frequency control, but does permit cancellation of reflections over a greater range of amplitudes than does the bridge method.

A basically different system has been used by Schmitt [35] and is shown in Fig. 5. Based on optical techniques, the direct and scattered signals are separated in space by means of a partially-reflecting dielectric mirror. In theory, the instrument provides considerable decoupling between the transmitter and receiver in the absence of the scatterer; however, in practice, spurious responses are quite troublesome. It will be seen from the figure that reflections from either the side or end of the ground plane can reach the receiving horn, so that absorbing material must be placed along these edges as shown. The result is an instrument with the same order of isolation as that described previously.

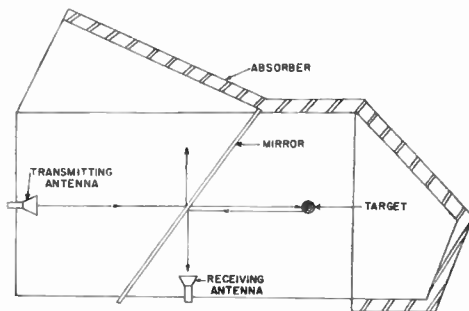


Fig. 5—A space-separation method of measuring the scattered field (after Schmitt).

A third system, used both by Scharfman and King [36] and by Tang [37] makes use of a Doppler frequency shift in the scattered signal to separate it from the leakage signal. Tang's apparatus consisted of a parallel-plate region 12 feet long by 6 feet wide, illuminated by an X-band horn. The obstacle, in this case a dielectric-coated cylinder, was mounted between two 34.5-inch-diameter disks, some 12 inches out from their centers, which were located three feet from the remote end of the plates. The parallel-plate region was terminated in lossy material to minimize reflections. The separation of source and obstacle was 80λ , so that amplitude variations over the test area were small. However, the phase front was cylindrical rather than plane.

In operation, a CW source is connected to the *H* arm of a magic-T and thence to the antenna via one of the branch arms. The disks carrying the obstacle are rotated at several hundred rpm, so that in the course of a revolution the phase of the scattered signal varies by several

wavelengths due to path length change. This scattered signal returns to the antenna where it reaches a crystal detector mounted on the *E* arm of the magic-T. Here it mixes with the leakage signal from the oscillator to produce an audio voltage which corresponds in frequency to the Doppler shift caused by the motion of the scatterer, and which varies linearly with the amplitude of the scattered field. This method is quite attractive because frequency and amplitude drift, as well as spurious-reflected signals, are practically eliminated. It does, however, suffer from one serious limitation. In the form described, only circularly symmetrical objects can be studied, since as the disks rotate, the aspect of the scatterer changes. This can be overcome by causing the obstacle to rotate relative to the disk at the same speed as that of the disk itself, but this is an obvious complication.

In describing various means of separating the incident and scattered fields, emphasis has been placed on techniques employing image planes or parallel-plate regions, which are only suitable for certain geometrical objects. It should be pointed out that these techniques are also applicable to the practical problem of determining the cross section of actual radar targets, which are usually quite large in terms of wavelength and cannot conveniently be studied full-size. We must therefore make use of the technique of electromagnetic modelling which is discussed in Section V. Scale factors of 5 or 10 to 1 are in common use and result in scaled frequencies lying in the millimeter wave region where available power is limited. These targets often do not have planes of perfect symmetry, and furthermore the scatter cross section is frequently required for incident fields polarized at various angles with respect to the target. These considerations dictate a free-space method in which the target is supported in the field, as far as possible from interfering objects. Since echoes from the supporting structure are a serious limitation, much care must be given to its design. It usually consists of a wooden frame, large compared with the target, from which the latter is suspended by means of nylon cords.

The image-plane, parallel-plate, and free-space methods can all be carried out indoors provided the range is not excessive, by locating the equipment in an anechoic chamber. The design of chambers for this purpose is quite critical and usually involves a great deal of adjustment in order to reduce reflections and spurious responses to an acceptable level.

The last method of back-scatter measurement to be discussed is most suitable for long ranges and is therefore applicable to large radar targets. It involves separation of the direct and scattered signals reaching the receiver by spacing them in time. This is accomplished by using a pulsed-radar technique such as the one described by Kennedy [34], in which two antennas are used, one for transmission and one for reception. Very short pulses of RF energy of 0.08- μ sec duration and a repetition frequency of 1500 cps are transmitted, and

after striking the obstacle, are received on the second antenna. A servo-driven attenuator placed in the receiver channel holds the receiver output constant, and controls the motion of a pen which records the attenuator setting. Permanent or background echoes from distant objects can be removed by a gated amplifier, but echoes from the supporting structure are troublesome because of their close spacing in time to the desired echoes.

One of the chief difficulties associated with the pulsed system is that ranges of 300 to 400 feet have often been required because of the width of the pulse, and thus the apparatus has had to be used outdoors. Further, it has been difficult to measure the echo area of small bodies. However, recent developments in the field of short pulse radars permit the generation of pulses 0.02 μsec in length, or at reduced power, only 0.002 μsec in length, so that short-range indoor test facilities are now possible.

The measurements so far described have involved the field scattered toward the source of the incident wave. In bistatic radar applications, and in certain fundamental studies, the field scattered in other directions is also of interest. The two-horn CW method is applicable to this measurement, as is the pulsed-radar technique. In either case the target is fixed with respect to the transmitting antenna, and the receiving antenna is moved in an arc while pointing constantly at the target. Energy reaching the receiver via undesired reflection paths can be tuned out if necessary by introducing signals from the transmitter directly into the receiver, or alternatively it can be gated out in time. An example of a pulsed system operating at 35 kmc is shown in Fig. 6. This equipment, which has the narrow pulse width mentioned previously, employs superheterodyne detection and uses S-band traveling-wave tubes as the IF amplifiers. Linear or circular polarization may be employed.

Additional means of measuring the echo area of various obstacles are described briefly in the survey paper of King and Wu [3], and are also noted in the bibliography.

III. TRANSMISSION AND REFLECTION

In addition to scattering and diffraction, reflection and refraction of electromagnetic waves by discontinuities in the propagation medium are of both fundamental interest and practical importance. Insofar as measurements are concerned, the fundamental interest lies in determining the electrical properties of materials by observing their reflecting or transmitting properties. On the applied side, the reflection coefficients of microwave mirrors, of the terrain adjacent to antennas, and of absorbing materials, as well as the transmission coefficients of radomes, are of importance.

The measurement of reflection and transmission coefficients is similar in many respects to the measurement of scattered and diffracted fields, since it involves

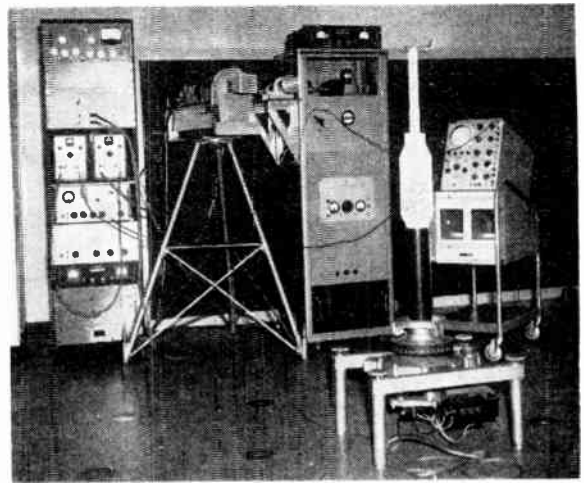


Fig. 6—Apparatus for bistatic cross-section measurements (courtesy of Defence Research Telecommunications Establishment, Department of National Defence, Ottawa, Canada).

a comparison between an initial or undisturbed wave and a wave which has been modified because of a change in medium. However, reflection and transmission are simpler phenomena to analyze and measure since they are associated with obstacles or discontinuities of a plane nature, or at least having radii of curvature large compared to the wavelength. Furthermore, they imply discontinuities whose transverse extent is large enough to intercept the entire incident wave, so that there are effectively no transverse discontinuities involved.

1. Reflection Coefficient

The specular-reflection coefficient of an interface is defined as the ratio of the reflected wave to the incident wave, with due regard for phase and polarization. It implies a smooth interface, or series of interfaces, between different media, where the maximum allowable roughness is usually described by the Rayleigh criterion, given by: $H < \lambda / (16 \cos \theta)$. Here, H is the height of the surface roughness, and θ , the angle between the incident wave and the normal to the surface. Measurement of the reflection coefficient can be made by methods very similar to those used for determining scattered fields. For example, measurement of reflection at normal incidence can be made using any of the techniques used for determining back-scatter cross section. However, for large reflection coefficients an even simpler method is possible, in which the energy reflected back into the source antenna is computed by observing the standing wave set up on the transmission line feeding it [46], [47]. In designing such an experiment, attention must be paid to multiple scattering or reflection between the antenna and interface [48]. Since the relative phase of, for instance, the first and second reflected waves reaching the antenna varies periodically with the separation between it and the interface, the effect of the higher order reflection can be eliminated by measuring the input-voltage standing-wave ratio as a function of spacing, and taking the mean between the maximum

and minimum VSWR as that due to the first reflection alone. Considering only this first reflection, it can be shown that the reflection coefficient of the surface is given by

$$|R| = \frac{8\pi d |\Gamma|}{G\lambda}$$

where $|R|$ is the reflection coefficient of the interface, d is the separation, $|\Gamma|$ is the voltage-reflection coefficient introduced in the transmission line, and G is the antenna gain for the separation used. This latter quantity can be calculated for small horn antennas following the method of Braun [49], who calculated the transmission between identical horn antennas for separations of up to eight times the usual far-zone distance of $d = 2a^2/\lambda$. An alternative method is to measure G directly by replacing the interface to be measured with a perfect mirror ($|R| = 1$). In both the calibration measurement and the reflection measurement, it is necessary that the transverse dimensions of the interface be large enough to justify the theoretical analysis, which for an incident plane wave assumes an interface infinite in extent. This condition can easily be met if the angle subtended by the sheet is large enough to include the main beam and major sidelobes of the antenna pattern.

The measurement of reflection at other than normal incidence is similar to the bistatic radar cross-section measurements described earlier. Two antennas are used, with their axes making angles of $\pm\theta$ with respect to the normal to the surface [50]. The sample size is chosen to intercept the principal energy from the transmitting antenna, and the antenna-to-sample spacing is made at least $2a^2/\lambda$, where a is the antenna aperture. The system is calibrated by noting the response when the interface is replaced by a perfect mirror. As grazing incidence is approached, this method fails because of direct coupling between the two antennas. This can be minimized by making the antennas highly directional. However, an alternative solution is to use antennas whose patterns in the plane of incidence are very broad, and to note the standing-wave pattern produced by the interference of the direct and reflected waves as one or both of the antennas is moved with respect to the interface. Such a scheme was used by Cumming [51] to measure the reflection coefficient of a snow-covered surface at X band, and is shown schematically in Fig. 7. A prepared area 40 feet by 12 feet was illuminated by a rectangular-waveguide array having slots in the narrow face. Metal plates were attached to the arrays, as shown, in order to reduce back radiation and broaden the elevation pattern. In operation, one antenna was fixed in height, while the position of the other was varied over a height of some two feet. Both antennas were then readjusted so that their beams were directed at the reflection point, and the process repeated.

Measurement of the phase angle of the reflection coefficient is more critical than measurement of the amplitude. In the normal incidence measurements using the

VSWR technique, the phase can be determined by noting the difference in the positions of the minima in the standing-wave pattern with the sample in place and with the reference mirror in place. However, because the separation between source and mirror is usually large, this measurement is very dependent on good frequency control. Similarly, in the oblique incidence measurement, the phase of the reflection can be calculated by comparing the position of the standing-wave pattern when the unknown is present, with its position when the mirror is present.

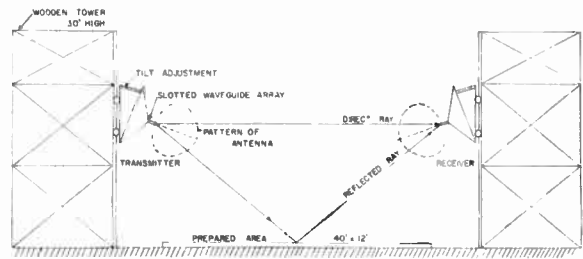


Fig. 7.—An interference method of measuring reflection coefficient (after Cumming).

B. Transmission Measurements

Measurement of the transmission coefficient of dielectric sheets, imperfect reflectors, or radomes, is analogous to the diffraction measurements described earlier, in the same way that reflection measurements are analogous to scattering measurements. The quantity to be determined is the ratio of the total field in the presence of the sheet to the total field in the absence of the sheet. The basic technique [46], [47] involves the use of a small transmitting antenna, usually a horn, arranged so that the angle of incidence may be varied. A similar horn is mounted on the other side of the sheet, and the amplitude and phase of the transmission field are measured by one of the techniques described earlier. Comparison of the result with that obtained in the absence of the sheet gives the transmission coefficient.

At the present time, transmission measurements of this type are most commonly made in connection with radome testing; furthermore, it is the phase shift through the radome, or the "insertion phase difference" which is of chief interest. This quantity is a measure of the difference in electrical path length between two points separated by the thickness of the radome, and is measured in one case with the radome present, and in the other with the radome removed. One technique is described by Kofoed [52], and is shown schematically in Fig. 8. The transmitting and receiving antennas are pointed at each other as shown in the figure, so that the angle formed by the transmission path and the surface of the radome is nearly equal to the Brewster angle of the dielectric material. In this way, the energy reflected back from the radome is reduced to a minimum. The entire surface of the radome is scanned by rotating the radome, while moving it slowly along its axis. Auxiliary

apparatus, not shown, is subsequently used to adjust the radome thickness in accordance with the measured changes in insertion-phase difference which occur from point to point. Under automatic operation, the instrument responds to changes of 0.15° , and the accuracy is stated as 0.5° .

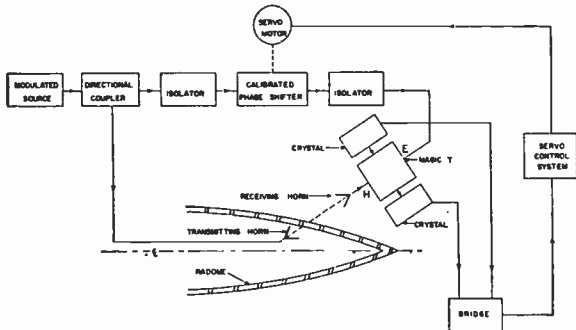


Fig. 8—The measurement of insertion phase difference (after Kofoid).

IV. ANTENNA DESIGN PARAMETERS

Turning now more directly to antennas, it is possible to describe their radiation characteristics in terms of a number of quantities, some of which are interrelated. These include aperture field, current distribution, radiation pattern, gain, polarization, and efficiency. In the case of an antenna-radome combination, the boresight error, losses, and scatter-lobes produced by the radome are also required in order to describe the performance of the radiating system. Of these various quantities the aperture field and current distribution are of interest solely to the antenna designer and may be called "primary characteristics." The remaining terms may be used to describe secondary characteristics which relate the performance of the antenna to that of the system.

A. Measurement of the Aperture Field

Measurement of the field existing in an aperture cut in a conducting sheet due to an incident wave has been discussed in Section II. Measurement of the field existing in the aperture plane of an electromagnetic horn or of a reflector illuminated by a primary source present a similar problem, somewhat modified because of the fact that in the case of a wave incident on an aperture in a sheet, the source antenna may be more loosely coupled to the aperture than is the case in a practical antenna. The various methods which have been described for field probing may be used, although the procedure may be complicated by the fact that practical antennas may have apertures many feet across. This increases the complexity of the transmission line connecting the probe to the receiver and recorder, as it dictates longer waveguide lengths and perhaps a greater number of rotating joints. Furthermore, in the case of a large antenna, it becomes exceedingly difficult to determine the probe position with an absolute accuracy of a few thousandths of an inch, as would be desired in some cases.

The generally close coupling between the aperture field of a practical antenna and its primary source makes it possible to measure the aperture field distribution by a scatter technique suggested by Justice and Rumsey [53], and by Cullen and Parr [54]. The basic method proposed by Justice and Rumsey involves the use of a small linear scatterer of some $\lambda/4$ in length, which is located in the aperture field by means of a fine nylon thread. The antenna under test is connected to one of the collinear arms of a magic-T, while the other collinear arm is terminated in a matched load. Energy is fed into the system from an unmodulated source through the *E* arm of the magic-T, while a ferrite modulator and receiver are placed on the *H* arm. It can easily be shown that if the electric field at the scatterer is E , the scattered signal V appearing at the detector is proportional to E^2 . Thus by recording the amplitude of V as the probe is moved across the aperture, the magnitude of the aperture field can be obtained. Phase information can be derived by deliberately unbalancing the bridge to produce a constant reference voltage R .

A second exploration of the region gives the quantity $V+R$ as a function of position of the scatterer. From the three quantities, namely V , $V+R$, and R , the phase of V and hence of E can be derived. An alternative method of obtaining phase is to make a direct comparison between the quantity V and a reference voltage taken from the source, using one of the methods previously described. In the case of large antennas, the decoupling between the aperture field and the input may be large, so that back-ground scatter becomes significant.

Richmond [55], Cullen and Parr [54], and Killick [56], have described systems in which the scatterer itself is modulated, thereby eliminating the difficulties due to incidental scattering.

A block diagram of the system proposed by Richmond [55] is shown in Fig. 9. Power is fed to the antenna from an unmodulated source through a magic-T. The linear scatterer is replaced by a germanium diode, of some 0.40 inch in length, which is supported by a pair of threads made slightly conducting by Aquadag. These serve to conduct current from a modulation source to the diode. As a result of the nonlinear impedance of the diode, the scattered signal is modulated at the frequency of the modulation source. The scattered signal reaching the magic-T via the antenna emerges from the *H* arm, where it is fed to the *H* arm of another magic-T, which forms part of a coherent detector [11]. An unmodulated reference signal taken from the source is fed into the *E* arm of this hybrid through a phase-shifter, as shown. The collinear arms of the T are terminated in matched crystals whose outputs are mixed in a bridge. Because of the properties of the coherent detector, it is possible to obtain outputs proportional to $V \cos \alpha$ and $V \sin \alpha$, depending on the phase of the reference signal, where $V \cos \alpha$ and $V \sin \alpha$ are quadrature components of the scatter signal. Thus it is possible to determine both the

magnitude and phase of the scattered signal, and of the aperture field.

The method used by Killick [56], and by Cullen and Parr [54], also makes use of a direct modulation of the scatterer in order to eliminate background signals. In both cases, the scatterer is supported at right angles to a nylon cord which is stretched across the aperture. This cord is driven by a motor in the manner of a rigid drive shaft, causing the scatterer to rotate in the plane of polarization. In this way a rotational frequency ω produces a modulation frequency 2ω . The apparatus employed by Cullen and Parr [54] does not use a coherent detector, but measures phase by making use of a calibrated phase-shifter in the antenna arm and a reference signal derived from a mismatch in the termination. The orientation of the polarization vector is also obtained by inserting a commutator between the output of the audio amplifier and the recorder. This commutator is driven at the rotational frequency of the scatterer, and the phase of the resulting modulation is controlled by changing the positions of the brushes on the commutator. By adjusting the phase for maximum output, the orientation of the field vector can be deduced from the brush setting.

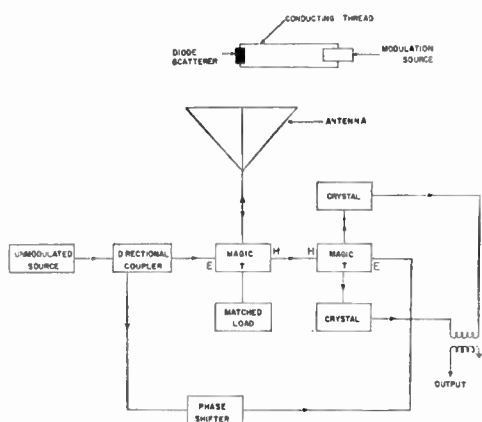


Fig. 9—The measurement of field distribution using a modulated scatterer (after Richmond).

It should be pointed out that aperture-field measurements are not necessarily confined to horn and reflector antennas. An excellent example of the usefulness of near-field probing in the design of arrays of linear electric-current radiators is given in a paper by Ehrenspeck and Kearns [59] dealing with a two-dimensional array of end-fire antennas.

B. Current Distribution

The radiation fields of most antennas can be derived from the amplitude and phase distribution of the electric conduction currents flowing on the antenna surfaces. However, in the case of many microwave antennas where the current flow is two-dimensional, in the form of a current sheet, it is more convenient to describe the far field in terms of the aperture field, as has been indicated.

The calculation or measurement of current distribution is therefore confined to combinations of linear radiators on which the current flow is predominantly longitudinal [61], or to certain specialized radiating structures of complex shape such as aircraft [62], [63] or land vehicles [64]. The measurement of current distribution was first made in connection with the design of broadcast antennas [65] using suitably scaled models of the towers. In more recent years such measurements have been applied to the fundamental investigation of folded dipoles, coupled antennas, loops, helices, and as has been mentioned, airframes and land vehicles.

The various techniques used in studying the longitudinal current flow on thin radiators are similar to the one described by Morita [66], whose apparatus is shown schematically in Fig. 10. By mounting a unipole

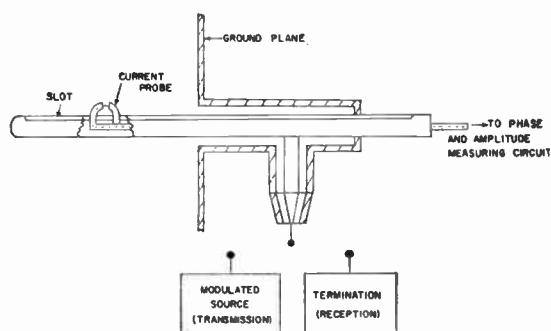


Fig. 10—The measurement of current distribution using an image plane (after Morita).

or other antenna above a ground plane in the manner shown, it is possible to provide both a direct mechanical connection and a separate electrical connection to the radiator. A longitudinal slot is provided through which a shielded loop projects. Except at the extremities of the antenna, this loop, which is quite small in terms of wavelength, produces an output which is proportional to the circumferential magnetic field, and hence to the longitudinal current flow on the radiator. The output of the loop is taken to a phase- and amplitude-measuring circuit through a coaxial line which passes through the hollow interior of the antenna. The antenna can be energized directly through the base connection in order to study its current distribution as a transmitting antenna, or indirectly from an external-source antenna, in order to study its properties as a receiving antenna. In the latter case, various terminations can be attached to the base connector, so that their effect on the current distribution can be determined. Measurements of this sort are usually carried out in the UHF band, between 200 and 1000 mc, which permits a reasonable compromise between ground-plane size and antenna and probe size.

Measurements of the current distribution on airframes and on land vehicles provide a valuable insight into the operation of antennas located on these structures. In both cases, it is customary to measure the current distribution on transmission, using a model of the

full-size structure. The probe cannot be contained within the radiator, as was described previously, but consists of a balanced loop manually held near the conducting surface. Because of the two-dimensional nature of the problem, the direction of total-current flow at a given point is also of interest, and can be determined by rotating the loop for maximum response. A high degree of balance is required in the loop; in practice it is possible to build loops with nulls 68 db below peak.

In the case of airframe measurements, the model is supported remote from the ground to simulate in-flight conditions. The airframe and antenna can either be energized from a self-contained battery-powered oscillator located within the fuselage, or from an external source. In the latter case, measurements made near the point of connection of the cable are not too reliable.

The current distribution on a land vehicle is influenced to a large extent by the presence of an imperfectly-conducting plane beneath the vehicle. In order to reproduce such a system accurately at, say a scale of 1/15, it would be necessary to model the ground constants [71, which is extremely difficult. A useful, if not completely accurate alternative, is to mount the scale-model vehicle over a perfectly-conducting plane. The RF signal can be led into the vehicle and thence to the antenna through a coaxial line which can be effectively decoupled from the ground plane by using the principle of the quarter-wave choke. This is accomplished by placing the coaxial line inside a metal sleeve, one end of which is mounted flush with the ground plane. A sliding short-circuit is provided between this sleeve and the outside of the line, and is adjusted to present a high impedance between the outside of the coaxial-feed line and the ground plane.

V. MEASUREMENT OF THE FAR FIELD

In the far zone of an antenna, at distances which are large relative to the aperture size, the field vectors are entirely transverse to the direction of propagation and vary inversely as the distance from the aperture. The spatial distribution of far field is therefore conveniently described as a function of an angular coordinate, at fixed radius. In the special case of a linearly polarized wave, a complete specification of the field can be given in terms of a three-dimensional radiation pattern and a three-dimensional contour of constant phase. While a number of proposals have been put forth to deal with the general case of an elliptically polarized wave [72], [73], the usual procedure is to resolve the electric- or magnetic-field vector into two orthogonal linear components whose amplitude and phase can be described separately. These two vectors are then related by specifying their ratio as a function of angular coordinates. This ratio is a phasor, whose modulus gives the relative amplitude of the two components, and whose argument gives their relative phase. The measurement of these three quantities—radiation pattern, phase pattern or phase contour, and polarization—will be dealt with separately.

A. Radiation Pattern Measurement

The radiation pattern of an antenna consisting of linear components is the same whether it be transmitting or receiving. The measurement of this pattern can be carried out either by probing the transmitted field of the antenna along an arc of fixed radius, or by measuring the angular variation of the response of the antenna to an incoming wave. The detailed technique of measurement depends on a number of factors, such as the size of the antenna and its operating frequency, as well as the environment in which it is to operate. Consideration of the latter factor leads to a logical division into two broad classes of measurement. The first of these may be termed “free-space measurements,” in which an attempt is made to isolate the antenna from its surroundings, measure its characteristics, and then to modify the result as necessary to include environmental effects. The second may be termed “*in situ* measurements,” in which the antenna and its surroundings are considered as an integrated radiating system.

1) *Free-Space Measurements*: The term “free-space pattern” in fact describes a hypothetical quantity which can rarely be actually measured in practice. This is because the structure used to support and position the antenna, the interconnecting cables used to connect it to the transmitter or receiver, and other nearby objects, invariably have some effect on the measurement. However, for many types of antenna this effect can be kept to an acceptable minimum by careful design of the experiment.

Radiation patterns are frequently measured by using the antenna under test as a receiver, as is shown schematically in Fig. 11. Here a source antenna, fed from a

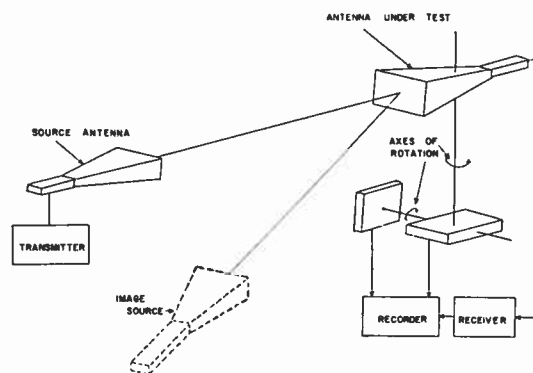


Fig. 11—Elements of an antenna pattern-measuring range.

suitable transmitter, is located some distance away from the antenna under test. Also shown is one of a number of possible virtual or image sources, which may be present because of reflections or scattering from nearby objects. The test antenna is mounted on a positioner, which in the general case permits rotation relative to the source on two orthogonal axes. The energy intercepted by the test antenna is fed into a receiver, and thence used to drive an automatic recording device.

Correlation of the received signal with orientation is accomplished by simultaneously providing the recorder with shaft position information.

In order to approach the ideal of a far-zone free-space pattern, the source antenna and the site must be chosen with two main factors in mind. First, the incident wave must be uniform in phase across the aperture of the test antenna, within arbitrary limits which depend on the type of antenna and the accuracy required of the measurement. Second, the wave must be uniform in amplitude across the aperture of the test antenna, also within somewhat arbitrary limits. To illustrate these points in an ideal case in which spurious responses are absent, consider a transmitting antenna of maximum aperture dimension d , and a test antenna of maximum aperture dimension D , separated by a distance R . The variations in phase and amplitude across D which can be tolerated depend to some extent on the field distribution assumed to exist across D in the transmitting case [74]; however, commonly accepted values are $\pi/8$ and $\frac{1}{4}$ db, respectively [75], [76]. Following the work of Rhodes [77], it can then be shown that in order to satisfy the phase requirement, R must be at least equal to $2D^2/\lambda$, and in order to satisfy the amplitude requirement when R has this limiting value, d must not be greater than D .

To illustrate the complications that arise when spurious signals are present at the test site, consider the geometry of Fig. 12. Here the source antenna and test antenna are raised to heights h_1 and h_2 above a specularly reflecting ground whose reflection coefficient may have any value. Energy then reaches the test site by paths r_1 and r_2 , with the interfering signal apparently originating from an image antenna located a distance $2h_1$ below d . The phase and amplitude of the image source are functions of the specular-reflection coefficient. If it is assumed for the moment that d and R have been chosen to satisfy the free-space criteria mentioned above, and further, if h_1 and h_2 are both much less than R , so that β , the angular separation between source and image, is small, then a wave-interference pattern of the type shown in Fig. 13 is formed at the test site. The amplitude, period, and phase of this pattern, together with the apparent angle of arrival of the resultant signal, depend on the geometry of the site, the reflection coefficient, and the source antenna pattern; while the effect these perturbations in the free-space source field have on the measurement depends on the characteristics of the test antenna. For example, if the test antenna is small compared with the period of the interference pattern, and if it is placed at one of the peaks of the interference pattern, then the requirements for a free-space pattern are effectively met, both in the azimuth and elevation planes, provided the apparent source position is taken into account. At the other extreme, in which the test antenna is very large compared with the interference pattern, a fairly reliable pattern may be obtained provided the amplitude of the interference pattern is small. In intermediate cases where the test an-

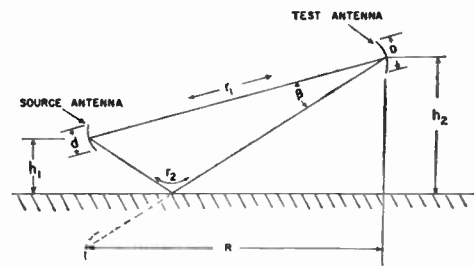


Fig. 12—Geometry of a pattern range over smooth earth.

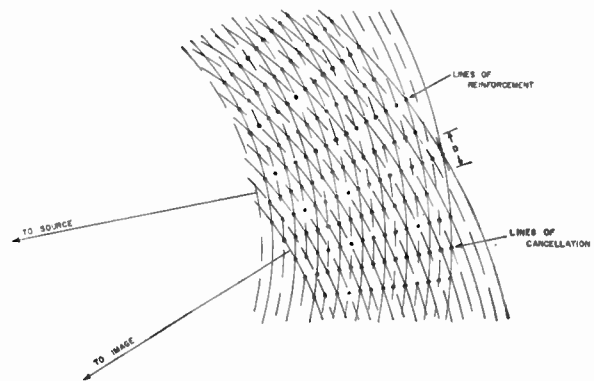


Fig. 13—Wave interference at the test site due to an image source.

tenna is comparable in size with the period of the interference pattern, variations across the aperture may lead to a completely erroneous radiation pattern. While the effect of interference phenomena or perturbations in the field at the test site is being discussed in terms of specular ground reflections, it must be stressed that equally serious spurious signals are possible because of scattering or reflection from trees, buildings, ridges, etc., lying either below or to the side of the radio path.

The design of a pattern range is in many cases made simpler if h_1 , h_2 , R , and the ground contour are chosen to produce the broad interference pattern described previously. This technique was pointed out by Cutler, King, and Kock [75], and has been used for a number of years at the Bell Telephone Laboratories. Following their analysis, it can be shown that if R is much greater than h_1 or h_2 , and if the reflection coefficient is -1 , the vertical variation of field at the test site is given by

$$E \propto \sin \frac{2\pi h_1 h_2}{R\lambda}$$

with a resulting distance between the first and second nulls of

$$\Delta h_2 = \frac{R\lambda}{2h_1} = \frac{\lambda}{\beta}$$

Furthermore, the height of the first lobe in the pattern occurs at

$$h_2^0 = \frac{R\lambda}{4h_1} = \frac{\lambda}{2\beta}$$

In designing such a range, the separation R and the size of the source antenna d are first chosen to satisfy at least the free-space criteria. The smallest values of h_1 and h_2 which will produce a sufficiently smooth interference pattern are then determined from the relations given above. To illustrate, suppose a range is to operate at 3.2-cm wavelength, and that the largest test antenna to be accommodated has a vertical dimension of 6 feet. Then the free-space requirement dictates a range at least 690 feet long. While there are many combinations of h_1 , h_2 , d , and R which will result in an amplitude distribution across the test site which is constant to within $\frac{1}{4}$ db, one possible design fixes R at 850 feet, d at 2.5 feet, h_1 at 1.25 feet, and h_2 at 18 feet.

In order that the intervening terrain produce a purely specular reflection, certain limits must be placed on the allowable roughness. The optical criterion of Lord Rayleigh is frequently used, in which the maximum difference in the path lengths associated with reflections from the mean terrain and from ridges or depressions is not allowed to exceed $\lambda/8$. For angles of incidence approaching grazing, this results in the height of such irregularities being limited to

$$H \lesssim \frac{\lambda}{16\psi}$$

where ψ is the angle of incidence. In the example cited, $\psi = 1.25/425 = 0.17^\circ$, and thus H must not be greater than 2.25 inches. In practice, test areas up to 1400 feet long by 200 feet wide have been precision graded to within ± 2 inches. If such an area is sown to grass and kept well mowed, it will produce an X -band reflection coefficient greater than 0.85 in magnitude at angles of incidence as great as 0.5° .

Since both the magnitude of the reflection coefficient, and to a greater extent the phase angle, are dependent on the moisture content of the ground, it is customary to check the shape and position of the interference pattern periodically. This can be done by probing the field at the test site with a small antenna mounted on a vertical track. An alternative test, suitable for day-to-day checks, is to mount the source antenna on a vertical track and to adjust its height for maximum received signal.

The chief disadvantages in the use of this method of solving the ground reflection problem lie in its sensitivity to frequency changes and climatic variations. The former is not too serious because the source height can be made variable, so that the working peak of the interference pattern can always be placed at the center of the test area. However, the latter disadvantage can prove quite costly in areas where frost and snow are prevalent. In the process of freezing, wet ground often heaves, destroying the smoothness of the prepared area, and of course snow must be cleared away after each storm.

An alternative solution, suitable for both snowy regions and for areas of rough terrain in which the cost

of levelling is high, involves the use of one or more diffraction screens set up normal to the ground and across the radio path. The purpose of such screens is to prevent energy from the source antenna from reaching the ground and being scattered to the test site. When viewed from the test site, the effect of such a diffuse reflection would be the same as that which would be produced by a large number of weak-image sources widely separated in space, and would result in random variations of large magnitude over the test antenna aperture. The energy diffracted by the screen or screens, on the other hand, interferes with the direct signal to produce a regular interference pattern of the type associated with a specular reflection, so that the test antenna may be located on the first broad maximum of this pattern. For an initially-chosen source antenna height, a single screen placed at midpath is made high enough so that ground-scattered signals cannot reach the test site in strength. The final height of the test site is then chosen to coincide with the peak of the first diffraction lobe. The field existing in the shadow and on the illuminated side of the line-of-sight can be conveniently calculated from the Fresnel integrals, or in graphical form from the Cornu Spiral [78].

The two methods which have been suggested to approximate a free-space test range are frequency sensitive, and rely on an image source. A third approach to the problem, and the one most commonly used in the design of both indoor and outdoor pattern ranges, attempts to eliminate all spurious fields, leaving only the direct wave from the source. This is accomplished by designing the source antenna so that the main reflection point lies in a null of the radiation pattern; by making h_1 and h_2 sufficiently high for the reflection point to be essentially outside the range of the radiation patterns of the source and test antennas; or by reducing the reflection coefficient to an acceptably low figure. The first of these methods is obviously frequency sensitive, but quite effective in cases where the intervening terrain is fairly smooth. It does, however, call for a detailed knowledge of the source antenna pattern and careful alignment of this pattern, or dictates that the field in the test area be probed for each frequency and antenna height used. In this connection, it should be mentioned that an alternative method of checking the field distribution at a test site is to measure the pattern of an antenna whose pattern has been calculated previously or measured on a pattern range of known characteristics. For a reliable check it is necessary, however, for this standard antenna to have dimensions comparable with those of the test antenna.

Where economic and geographic considerations permit, the elevated-site technique has many advantages. Chief among these is the fact that the field at the test site is not frequency sensitive to the extent that it is when using the other techniques, having only a limit based on the minimum source directivity required to avoid illuminating the reflection point. An excellent

example of such a range, usable down into the UHF band, is shown in Fig. 14. This photograph, taken at the Newhall pattern range of the Lockheed Aircraft Corporation, shows the test site on the slope of one mountain as viewed from the transmitting site located on a neighboring mountain. The separation is 2260 feet, with the source some 600 feet and the test site some 370 feet above the valley floor. The depression angles to the reflection point are 17° and 26° , respectively, so that even antennas of modest directivity provide a considerable degree of discrimination against reflections. Such large depression angles at a range of one half mile are obviously only possible in certain geographical regions. In most places, the best that can be done is to take advantage of the topography insofar as possible, to keep the range as short as is consistent with the phase-front requirement, and to elevate one or both antennas on towers. In this way, depression angles of about 5 to 10 degrees are possible, which is satisfactory at the high-microwave frequencies in *S* and *X* band. At lower frequencies, in the vicinity of *L* band and below, it is difficult to eliminate ground reflections, and they must therefore be accepted as forming part of the field at the test site.



Fig. 14—An antenna test site as viewed from the transmitter (courtesy of the Lockheed Aircraft Corporation, California Division).

An interesting method of avoiding reflection effects, which has had limited use, is to design the pattern range so that the radio path is vertical. The antenna under test is fastened to a rotating device mounted on a trolley which rides on a track fastened to a wooden pole perhaps 75 feet high. The source antenna which is mounted at ground level then points skyward, illuminating only the test antenna, the trolley, the tracks, and the pole. This method is limited by its short range and by the fact that the scattering from the objects mentioned may be serious at certain frequencies, particularly if the test antenna has poor directivity.

Reflections and other spurious responses pose even greater problems in the design of indoor pattern ranges, due to the presence of the walls and ceiling of the room, as well as the floor. As a result, such ranges must be carefully designed and tested for each measurement

problem, and are usually somewhat frequency sensitive. However, independence from climatic conditions is very valuable, and in the case of production testing for example, completely offsets any lack of flexibility in the facility.

Ideally, the space allotted for an indoor test range or anechoic chamber should be sufficiently large that the walls, ceiling, and floor at the geometrical reflection points would not be illuminated by the source antenna. Since this would result either in a very large chamber or one of limited usefulness, other techniques have to be employed. Because of the multiplicity of possible reflection paths and the resulting complexity of the interference pattern at the test site, spurious responses must be reduced to a low level, and in a room of practical size this can be done by confining the radiation of the source antenna to a sharp beam, and by reducing the reflection coefficients of all interfering objects to small values. In order that the test antenna be illuminated by an essentially plane wave, it may be placed in the near field of a large dielectric or metal plate lens, illuminated by a primary feed [79]. This technique does not appear to be used extensively, primarily because of the limited bandwidth of most lenses, and because of possible interaction between the test antenna and the lens. The design of most indoor ranges is based on the assumption that the test antenna is in the far field of the source, and so spurious reflections must be dealt with. This is done in a variety of ways, all of which employ lossy materials to absorb incident energy and to reduce specular reflection. Microwave absorbers [80]–[82] are of three basic types. The first consists of hair or fiber mats, available in thicknesses of from 2 inches to 8 inches, which are impregnated with rubber to which a lossy material, usually carbon, has been added. By making the loss vary with distance through the mat, an effective “free-space” termination is obtained. The second type in use employs a thin resistive film which takes the form of a large number of interconnected hollow pyramids whose dimensions are comparable with the wavelength. This product is constructed by first casting the pyramidal surface in rigid foamed plastic. The lossy film is then applied, and a second slab of foam cast on top, so the finished material has the appearance of a uniform slab of foamed plastic whose thickness depends on the operating wavelength. Both these types of absorber are capable of producing a power reflection coefficient of the order of 3 per cent at incidence angles to within 20° of grazing, and are broad-band in the sense that they have a lower frequency limit, above which they are effective. The third type of material gives a reflection that is somewhat more dependent on frequency and incidence angle, but it has the advantage of being thin and flexible. It consists of a flexible foamed plastic or rubber sheet with a layer of lossy material on one face. A number of methods have been suggested for using these absorbing materials. The most straightforward is to line walls, ceiling, and floor with the absorber, and following

the results of field probing, to use additional baffles or screens of absorbing material as needed. A photograph of one such design, based on this technique, is shown in Fig. 15. This room, designed primarily for *K*-band measurements, is some 70 feet long, 20 feet wide, and 20 feet high. A large number of auxiliary baffles are employed to keep field perturbations to a satisfactory level. Another technique, interesting because it does not require a separate room, is to surround the radio path between source and test site with a number of baffles of absorber, placed parallel to each other and transverse to the path. The resulting configuration is that of a number of rectangular "washers" of 2-inch or 4-inch hair absorber, whose inner dimensions are large enough to be out of the peak illumination of the source and whose outer dimensions and spacing are chosen so that energy diffracting from the inner edges of a given "washer" will not illuminate outside objects with sufficient strength to perturb the field at the test site. The entire assembly can be suspended from the ceiling so that ready access to the test area and radio path is possible.

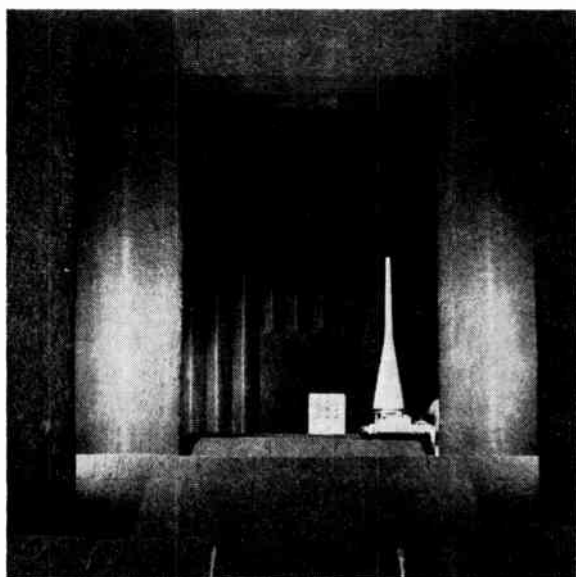


Fig. 15—An anechoic chamber viewed from the source antenna (courtesy of the Lockheed Aircraft Corporation, Missile Division).

Returning to test ranges which utilize a separate room, a novel design is shown in Fig. 16. In order to insure that spurious signals can reach the test area only by paths involving two or more reflections, the walls, floor, and ceiling are made pyramidal rather than flat. Since the spurious signal is attenuated, perhaps 15 db on each reflection, the total attenuation on a multihop path may be quite large.

Finally, it should be mentioned that at the lower very-high and ultra-high frequencies, current practice is to line the walls of a room with absorber in the form of a large number of pyramids projecting normally from the enclosing surfaces. These may be more than 4 feet long, so that a large enclosure is needed in order to pro-

vide a working area of reasonable size. So far as is known, the array of parallel "washers" has not been used at VHF or UHF, but should be considered as another alternative.

In the case of both indoor and outdoor measurements, the situation occasionally arises in which it is impractical to separate the source and the test antennas sufficiently to satisfy the far-zone criteria. This problem has received considerable attention in recent years [83]–[86], and two main solutions have been proposed. In one case, the phase and amplitude distribution at the test site is first determined, whereupon the radiation pattern is measured in the usual manner. The true far-zone pattern is then calculated from the measured near-zone pattern, a process which in most cases is quite tedious. However, it does permit the determination of the Fraunhofer pattern from a measurement made in the Fresnel region. The second scheme is applicable only to certain types of antenna, such as slotted arrays or reflectors illuminated by a primary source, and requires that the antenna under test be focused to the range of the pattern site rather than to infinity. In the case of the slotted array, this is accomplished by bending the antenna into a circular arc, in such a way that pathlength errors are compensated for. In the case of the reflector, the feed is displaced from the prime focus in such a way as to produce compensating phase variations across the aperture plane.

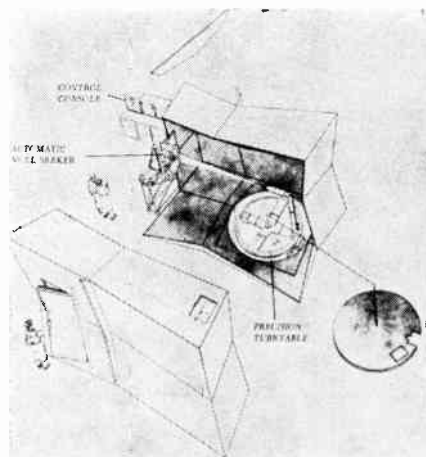


Fig. 16—The measurement of boresight error using a null-seeker (courtesy of Hughes Aircraft Company).

The turntable or mounting jig used to support the test antenna must be designed to permit measurement of the desired portion of the three-dimensional pattern without, at the same time, appreciably distorting the incident field. The amount of pattern information required depends to a large extent on the type and application of the antenna under test. For example, in the case of large pencil-beam or fan-beam microwave antennas, it is usually sufficient to measure patterns only in the principal planes. This can be accomplished most easily by using a turntable with an axis of rotation fixed in the vertical direction. For the *E*- and *H*-plane pat-

terns in succession, the antenna is mounted on the turntable, tilted in elevation for maximum received signal, and the pattern is measured. Provided the beamwidth is small, *i.e.*, not much greater than 10° , it is not necessary that the axis of rotation and the radio path to the source or virtual source be exactly perpendicular. A more sophisticated mounting arrangement is shown in Fig. 17, in which the main rotational axis of the antenna mount can be tilted from the vertical. This permits vertical alignment of the test antenna for maximum response, while at the same time insuring that the axis of rotation and the radio path are at right angles, as is desired for a principal-plane pattern. Furthermore, in cases where a conical section through the three-dimensional pattern is required, it is a simple matter to adjust the cone angle to any desired value. Finally, with this type of mount it is possible to measure the elevation-plane patterns directly by recording the received signal as a function of tilt-angle, provided however that the incident field is uniform, and due to only a single source, real or virtual. If two or more apparent sources are present, they must be separated by angles large compared with the beamwidth of the test antenna. Such dual-axis antenna mounts or positioners are extremely useful for measuring the pattern of shaped-beam antennas, antennas with significant off-axis side-lobes, or any other antennas for which more than principal-plane information is required.

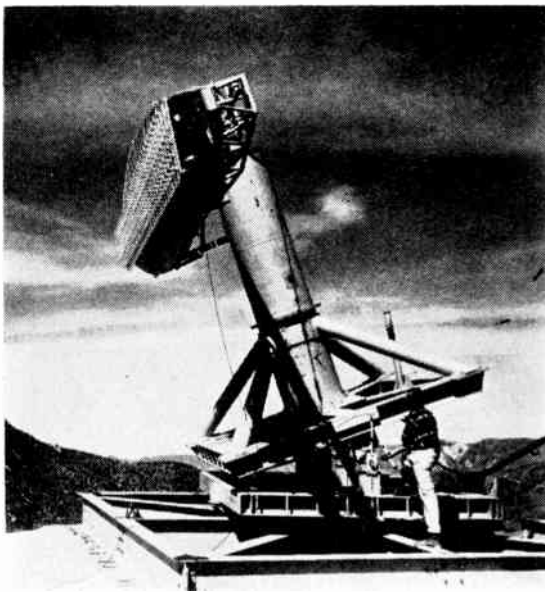


Fig. 17—A heavy-duty dual-axis antenna positioner with antenna installed (courtesy of the Lockheed Aircraft Corporation, California Division).

Mounting the test antenna in such a way that the positioning device has negligible effect on the pattern presents no problem in the case of high-gain microwave antennas. Since some prior knowledge of the pattern of the antenna is always available, it is usually sufficient that the supporting platform and associated brackets, cables, etc., be placed outside the angular sector occu-

ried by the main beam. In this way, energy reflected or scattered by these objects will not reach the receiver at a sufficiently high level to seriously affect the pattern. Low-gain antennas, on the other hand, present a serious problem insofar as mounting is concerned, because of the large solid angles occupied by their major lobes. In the extreme case of an antenna with a large backlobe, energy can reach the receiver by reflection from a support located behind the antenna. In such cases, expanded or foamed dielectric supports are quite useful, since they represent a minimum discontinuity in the medium and hence produce smaller perturbations than metallic or solid dielectric structures. In the VHF and UHF bands, wooden supporting structures are frequently used, since at these frequencies, where the wavelength is comparable with the dimensions of the structure, it is possible to induce resonant oscillations on metallic members, thereby enhancing the scattered field. Finally, in extreme cases, the test antenna can be positioned by means of nonconducting cables or cords, as has been pointed out earlier.

While the effect of the antenna-supporting structure may be reduced to an acceptable value by one of the methods suggested, the presence of a metallic cable connecting the antenna to the receiver may be equally serious. Various means of transmitting the signal picked up by the test antenna to a remote receiver have been suggested [87], [88], and include a modulated light or sound beam as well as an auxiliary radio link. These systems suffer from a common limitation, that of insufficient dynamic range. Since a square-law detector such as a crystal or a bolometer normally serves as the detector for the signal intercepted by the test antenna, a dynamic range of 80 db is required of the telemetry equipment if patterns are to be measured over a 40-db range. This is possible only with a large volume of auxiliary apparatus whose presence may in turn distort the pattern. The alternative solution of using a small battery-operated preamplifier, connected to the main amplifier by means of high-resistance leads, is an interesting one. These leads, in the form of a rubber-covered twisted pair of graphite-impregnated nylon, have a resistance of several thousand ohms per foot. The RF current induced in them is therefore small, so that they have a small effect on the incident field. At audio frequencies they operate as simple dropping resistors, whose loss is overcome by the preamplifier. Separations between the test antenna and main amplifier of up to 15 feet are possible using this technique.

The same precautions with regard to reflections and scattering from supporting structures must be observed at the transmitting site, although the problem is not so severe because of the limited movement required of the source antenna. At microwave frequencies, horn antennas are preferred because of the linearity of their polarization. Paraboloidal reflectors are used in cases where a high-gain source is desirable, but attention must be given to the relatively large off-axis cross-polarized

sidelobes which can exist under certain circumstances [75]. This cross-polarized energy may reach the test site by way of a reflection path, and thereby distort the measurement. At very-high and ultra-high frequencies, corner reflectors and Yagis serve very well as source antennas.

The power required for radiation-pattern measurements is considerably greater than that normally needed for impedance measurements because of the propagation loss and greater dynamic range inherent in the former. Values ranging from 100 mw to 1 watt CW are usually adequate, provided the range is kept near the minimum acceptable value of $2D^2/\lambda$. Fan-beam antennas pose the greatest problems since their high directivity in one plane dictates a long range, while their low directivity in the other plane results in low antenna gain. In certain cases these conflicting conditions can only be met by the use of high-powered klystrons, power oscillators, or pulsed magnetrons. In the majority of pattern measurements, however, the required power can be obtained from triode oscillators, low-powered klystrons, and more recently, traveling-wave-tube amplifiers and backward-wave oscillators. In pattern ranges of the type being discussed, in which the source antenna is remotely located, provision for remote control of the transmitter is useful [89]. The power and frequency stability required of the transmitter are to a large extent determined by the bandwidth of the antenna and receiver, and by the design of the recording equipment.

In recent years interest in low-level sidelobes, 40 db and more below the peak of the main beam, has grown considerably. As a result, pattern measurement facilities must be designed with a dynamic range of 40 db, and in some cases as much as 60 db. In order to stay within practical limits of transmitter power, this dictates that the receiver-detector-recorder combination have as great a sensitivity as possible, and sufficient dynamic range to meet the requirements.

In some systems the receiver and detector [90]–[95] may be considered more or less independently from the recorder. In this class are the low-level crystal receiver, the bolometer receiver, and the superheterodyne receiver, all of which may be used with or without phase detection or coherent detection. In the VHF, UHF, and microwave regions the crystal receiver has long been used in connection with pattern measurements, and consists of a RF tuner, a crystal detector, and a tuned audio amplifier. Such a combination, using a 1N21 crystal and an audio amplifier tuned to 1 kc, with a bandwidth of some 3 cps, has a sensitivity at X band of approximately 3×10^{-11} watts, and a dynamic range of 45 db. A bolometer detector used to drive the sample amplifier results in a comparable dynamic range, but a sensitivity of approximately 3×10^{-9} watts. In contrast, the crystal receiver followed by an AF phase detector results in a sensitivity of some 3×10^{-12} watts and a dynamic range of 55 db. This last scheme, however, has the drawback of requiring that a portion of the modu-

lating audio signal be made available at the test site to operate the phase detector.

Turning to a more elaborate system, a carefully designed superheterodyne receiver followed by a diode detector is capable of a sensitivity of some 3×10^{-14} watts and a dynamic range of 85 db. Such sensitivity demands very careful design of the receiver, and frequency and amplitude stabilization of the source. In addition, the local oscillator must be stabilized or fitted with automatic frequency control. One novel form of automatic frequency control in current use embodies a swept-frequency local oscillator whose sweeping range is somewhat greater than the IF bandwidth. A phase detector following the IF amplifier provides an AFC voltage which adjusts the sweeping range to insure that it includes the pass band of the IF amplifier.

Finally, the superheterodyne receiver followed by phase detection at either the intermediate or audio frequencies is capable of even greater sensitivity in return for increased complexity.

A variety of recording devices has been proposed to present the receiver output in graphical form. The earliest of these, used during and prior to World War II, was a simple graphic ammeter which plotted the radiation pattern in curvilinear coordinates. Since then various servo-driven graphic recorders have been designed [96]–[99] to plot either linearly or logarithmically on a polar or rectangular coordinate system, and a number of such instruments are available commercially. At the present time a linear presentation is usually used with polar coordinates, while a logarithmic plot is normally used with rectangular coordinates. This division is made because polar plots are used primarily for omnidirectional or low-gain antennas, while rectangular plots are used for high-gain antennas in which sidelobe levels are more important.

Since the output of a crystal or bolometer detector varies as the square of the input, special techniques are required to produce either a linear field strength plot or a logarithmic plot. The former, for example, can be achieved by the use of a square-root amplifier preceding a linear recorder, the only difficulty being in obtaining long-term stability of the amplifier. Alternatively, LeCaine and Katchky [98] proposed a nonlinear servomechanism incorporated within the recorder, using two ganged linear potentiometers with outputs derived from the received signal, and from a reference signal taken from the transmitter, respectively. In addition to employing simpler linear amplifiers, this two-channel system has the added advantage of being independent of changes in transmitter output. Logarithmic recorders may also employ nonlinear servomechanisms built around a logarithmic potentiometer [100]. The incoming signal is attenuated by the potentiometer, rectified and compared with a reference voltage. The resulting difference or error voltage is then passed through a chopper, amplified, and finally used to control a servomotor which drives the potentiometer to a balance. The re-

ording pen is mechanically linked to the potentiometer, so that as the servosystem responds to changes in input signal, a logarithmic plot is obtained.

An alternative method of obtaining a logarithmic plot makes use of a logarithmic amplifier [101] operating at AF or HF, and a linear recorder. Such an amplifier can be made to operate over a range of some 60 db, with an accuracy in response of 0.6 db. Finally, an integral receiver-recorder makes use of a superheterodyne receiver incorporating a piston attenuator in the IF amplifier. By using a servosystem to position the attenuator for constant IF output, and by mechanically linking a recording pen to the attenuator, a logarithmic plot is obtained.

Two specialized equipments used to present the radiation pattern of an antenna in graphical form are worthy of mention. The first of these is the "contour plotter" [102], used to present automatically contours of constant amplitude of an antenna pattern. In the form described by Dyson and Tucker, this instrument uses an indicating galvanometer to scan an opaque sheet in which are cut slits corresponding to predetermined field-strength intervals. Light from the galvanometer passes through a slit, is focused on a photocell, and causes a stylus to record a dot on the record sheet. By scanning the antenna in both azimuth and elevation, the entire beam may be covered, and joining the appropriate dots results in a complete contour map of the beam. The second is a rapid means of displaying an antenna pattern so that it may be either viewed or photographed, and makes use of a high-persistence cathode-ray tube on which the antenna pattern is displayed [103], much as is done with a plan-position indicator radar presentation. While not suitable for the study of low-level side-lobes, or for the accurate measurement of beamwidth, this method is useful in the preliminary stages of a test program.

2) *In Situ Measurements:* The effect of environment on the pattern of an antenna is in certain cases very complex, involving scattering, diffraction, reflection, and resonance phenomena. As a result the technique for determining the free-space pattern and modifying the result to allow for environmental influence becomes impractical, and the antenna pattern must be determined *in situ*. Such is the case when studying the patterns of certain airborne antennas, shipborne antennas, antennas on land vehicles, broadcast antennas, and more recently, large radar antennas. The patterns of the last two types of antenna are influenced primarily by the ground or terrain in their vicinity, while the patterns of the other types are determined by the configuration of the vehicle and the location of the antenna on the vehicle.

Airborne, shipborne, or vehicular antenna patterns can be measured on either the full-size system or on a scaled model. The latter technique is limited by the ability to construct an accurately scaled model; this requires close tolerances in linear dimensions, conductivity, and frequency, which if the scale factor is large,

may prove impractical. Consider for example a 1/48-scale model of a Yagi antenna operating at 200 mc full size. The element diameter in the full-size antenna might be 3/8 inch, in which case that in the model antenna would have to be 0.008 inch, an impractical dimension. Since the pattern of a Yagi antenna is very dependent on the self-reactance of the various elements, and hence on the element diameters, it becomes impossible to construct an accurate model. Full-scale measurements, on the other hand, are very costly, particularly in the case of airborne or shipborne antennas. Furthermore, precise control and determination of the vehicle's position is required. Perhaps the most serious drawback of all is that a full-size vehicle must be available, so that full-scale pattern measurements in the early design stages are not possible.

As a result of these considerations, model pattern measurements are used where possible to check the performance of proposed installations, while full-scale measurements are reserved for a final test which, because of the uncertainties mentioned, together with uncertainties in propagation conditions, must really be termed "systems tests."

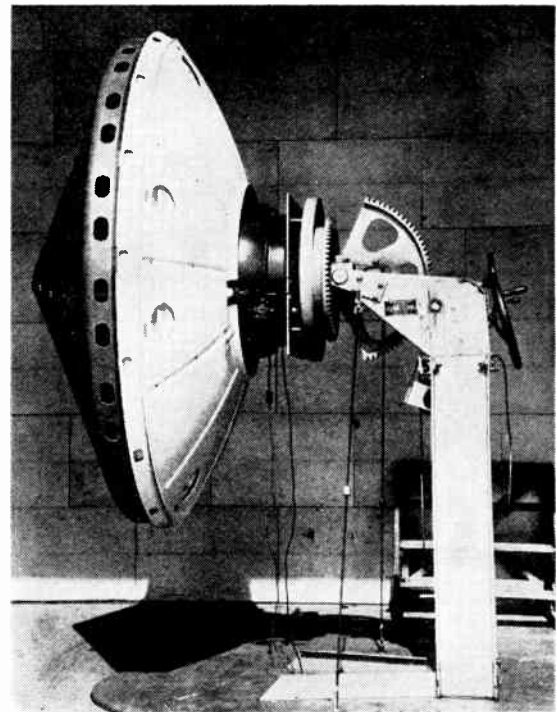


Fig. 18—A partially-enclosed test site for pattern measurements (courtesy of Bell Telephone Laboratories, Inc.).

Model aircraft measurements are employed on antennas operating in the LF, HF, VHF, and UHF bands. Above these frequencies scale-model antennas cannot be used because of their small size, and in any case, the effect of the airframe on the free-space antenna pattern can be calculated with reasonable certainty. The scale factors normally used lie between 1/5 and 1/50, and are a compromise between model aircraft size, model antenna

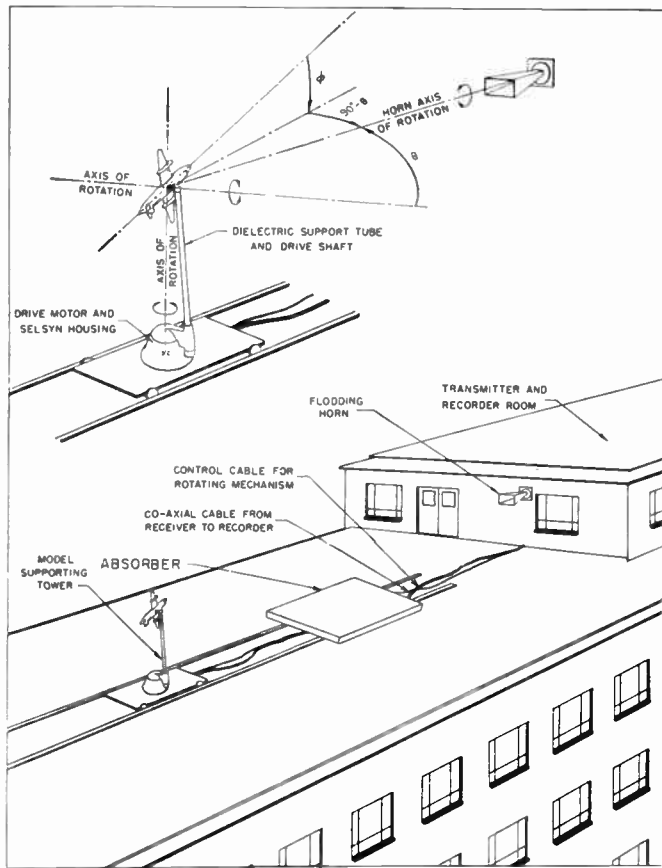


Fig. 19—Geometry of a typical model aircraft antenna range.

size, and model frequency. Scale-model aircraft are constructed either of sheetmetal, or more frequently, of wood or plastic which is metal-sprayed in order to satisfy the conductivity requirement. The models are hollow so that the receiver, usually a simple RF tuner and a crystal or bolometer detector, can be mounted inside. Since the analysis of an air-to-air or air-to-ground system can best be carried out with reference to a spherical coordinate system, the model is usually mounted on a supporting tower in the manner shown in Fig. 19. This dual-axis mounting arrangement permits plane cuts through the three-dimensional radiation pattern, as well as conical cuts defined by constant values of θ , the polar angle. The supporting tower must be designed so as to disturb the incident field as little as possible; this dictates that the size of the gears, drive-shaft, and torque tube be kept to a minimum. Scaled VHF and UHF antennas present the least difficulty because they are sufficiently directive that the supporting shaft can usually be placed in a low-intensity region of the antenna pattern. Scaled HF antennas, on the other hand, produce very broad patterns, so that interference from the tower and feed cable presents difficult problems. The latter problem is usually overcome by using the high-resistance audio leads referred to earlier.

A pattern range for scale-model aircraft measurements must be versatile with respect to frequency coverage and path length, in order to accommodate a variety

of scale factors and aircraft of different sizes. By placing the supporting tower on a trolley as shown, the length of the range may be adjusted easily, while broad frequency coverage can be obtained by using as source antennas a series of interchangeable horns, designed to fit a standard mount. This mount should permit rotation of the source antenna so that the polarization may be changed easily, and so that polarization patterns may be measured.

Midpath reflections present the usual problems and can be overcome by any of the methods described previously. Since model ranges are usually limited to the order of 100 feet in their maximum path length, it is feasible in many cases to use absorbing material as shown in the sketch. In addition, absorbing material may also be needed on the base of the supporting tower, as is shown in Fig. 20.



Fig. 20—A scale-model aircraft and antenna mounted for pattern measurements (National Research Council of Canada photograph).

Aircraft loop and sense antennas, used for LF direction finding, operate at wavelengths large compared with the airframe dimensions. As a result their patterns are essentially those of a doublet whose orientation is a function of the antenna location. In view of the relative size of the airframe and wavelength, the determination of this pattern, either by calculation or measurement, can be well approximated by treating the problem as a static one. Since the sense antenna is usually a short electric-current element, or its equivalent, the static pattern can be investigated by using a scale model located in an electrostatic cage [104]. For a given fixed field between the plates of the cage, the charge induced on the antenna can be measured as a function of orientation of the model. Similarly the LF loop antenna can be investigated by means of an electrolytic tank [104], [105]. Both these instruments have proven quite useful in extending model techniques to LF antennas whose patterns could not be measured on conventional model ranges because of severe environmental disturbances.

Full-scale measurements of aircraft antennas can be made reliably in certain cases, at least in one principal plane. For example, a fairly accurate azimuthal pattern of a UHF communications antenna, mounted on the

fuselage of an aircraft, could be measured by parking the aircraft on a runway, energizing the antenna, and probing the field at fixed radius. The method depends on the ground having an effect independent of azimuthal angle, and therefore implies a uniform ground and a fairly constant angle of incidence. Because of these limitations, more complete full-scale pattern tests are normally made in flight, and require that the aircraft be flown precisely on a prescribed path with respect to a ground transmitter. Because of difficulties encountered in correlating received signal with aircraft position such measurements are often repeatable to an accuracy of about 25 per cent.

Scale-model shipborne antenna measurements also involve antennas operating from the LF through to the UHF bands. The techniques used, however, differ considerably from those just described for three reasons. First, the physical and electrical dimensions involved are considerably larger; second, the antennas are mounted near the boundary between two media having widely different properties, and finally, the superstructure of a ship, consisting as it does of structural members of various shapes and sizes, does not give rise to the specular reflections associated with the smooth surfaces of an aircraft. At VHF and UHF, the effect of the superstructure is the most difficult of the three to analyze, and requires direct measurement. This is frequently carried out at 1/6 scale, and utilizes only that portion of the superstructure located within a distance of a few wavelengths of the antenna under test. Specular reflections from the sea and from the deck are treated as a separate problem, which can be solved by calculation.

At LF and HF the sea, as well as the hull and superstructure of the ship, affect the antenna pattern, so that all three have to be taken into account. In a scale-model study this is done by building a replica of the ship, usually to scale 1/48, down to the waterline. This model is soldered to a conducting turntable which is set flush into a conducting plane, as is shown in Fig. 21. This simple "model" of the smooth sea is possible because in the LF and HF bands, seawater is a reasonably good conductor, and when the conductivity is further increased by the scale factor of perhaps 48 times, the conductivity of typical metals is reached. The fixed area may take the form of an asphalt surface sprayed with a thin coating of lead or aluminum [106], [107]. Conductive coupling between the turntable and the fixed area may be obtained by attaching to the turntable a cylindrical lip which dips into a stationary trough of mercury forming the inner boundary of the sprayed asphalt area. An alternative solution [108] which appears to be quite satisfactory is to rely on capacitive coupling.

Pattern information in the principal azimuthal plane is obtained by transmitting from an antenna placed at one extremity of the model ocean and recording the response as the turntable is rotated. In order that conical patterns may be taken, provision is made to elevate the source antenna above the ground plane. This can be

achieved by moving the source antenna on a circular section of track passing over the center of the turntable, or by raising the source antenna on a long mast. In either case elevation patterns may be measured directly by recording the response as a function of elevation angle. It must be noted, however, that such elevation patterns are significant only insofar as the sea may be regarded as a good conductor.

As in the case of airborne antennas, full-scale pattern measurements can be carried out on shipborne antennas. At the lower frequencies, this can be done conveniently by steering the ship on a predetermined course and recording the signal received from a shore station. Pattern checks of UHF and VHF installations require the use of aircraft or helicopters in order to obtain the desired information.

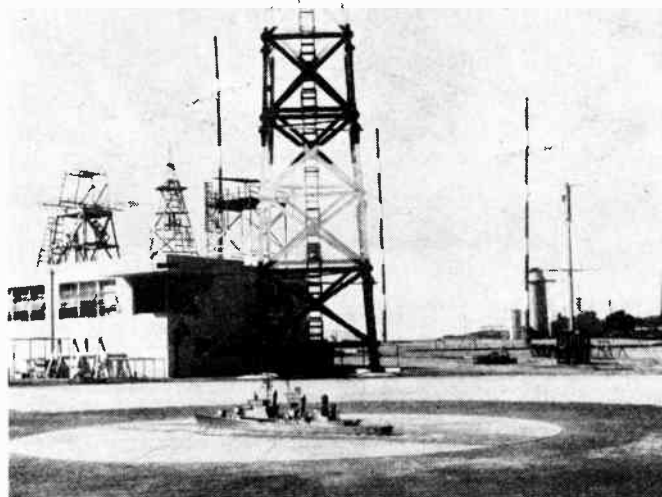


Fig. 21—A pattern range for studying HF shipborne antennas (official U. S. Navy photograph).

Finally, model and full-scale pattern measurements are useful when designing and assessing antennas for use on land vehicles. It has been pointed out in Section IV that the presence of ground beneath a vehicle carrying an antenna influences the current distribution on the vehicle's frame, and hence influences the radiation pattern. As a result some attempt must be made to take the ground into account when constructing a model. To exactly model the earth constants in a system operating at, for example, 1/15 scale, is extremely difficult. However, useful results are obtained if the ground is simulated by a metallic sheet, large compared with both the scaled vehicle and the wavelength. A facility such as that described for the study of shipborne antennas is well suited to the study of vehicular antennas, provided the scaled frequency is sufficiently high. Since those vehicular antennas whose patterns are most influenced by the vehicle's frame operate in the 30- to 70-mc band, scale factors of 1 in 10 to 20 are quite satisfactory. The undercarriage of the vehicle, which forms an important

part of the earth circuit, must be modelled very carefully in order to obtain reliable results.

Full-scale vehicular antenna pattern measurements are most reliably made by parking the vehicle in a clear area having typical ground constants, and by carrying a portable generator or field-strength meter around it at a constant radius. An alternative method which can be used with a lightweight vehicle, requires a swivel jack, so that the vehicle may be raised just clear of the ground and rotated.

The elevation pattern of a broadcast antenna operating over a homogeneous earth is a function of the earth constants. As a result, the azimuthal pattern of such an antenna operating over a nonhomogeneous earth is dependent on the radial and azimuthal variation of earth constants. While it is theoretically possible to modify the pattern of the antenna, as calculated for idealized conditions, in such a way as to take into account known ground conditions, it is usually more satisfactory to measure the equi-field-strength contours directly. This is done with a portable field-strength meter mounted on a vehicle, and involves sampling the field at a sufficient number of points to permit the construction of an accurate contour map. Model techniques have been used as an aid in the design of broadcast arrays, but provide only the pattern which would exist over uniform earth.

The measurement of the patterns of large communication antennas, radar antennas, and antennas for radio astronomy, poses special problems. First, because of the physical size of the antennas, as well as their electrical dimensions, it is in many cases impractical to obtain a free-space pattern. Second, in the case of the radar or communication antenna, it is doubtful whether a free-space pattern is really of value from the system point of view because of the influence of the site and surrounding terrain on the total pattern, effects which in many cases cannot be calculated with accuracy. As a result, these antennas are frequently tested under operating conditions using an aircraft or helicopter [109], [110]. Such pattern measurements serve a dual purpose. For example, consider *in situ* measurements on a large radar antenna. In the initial stages of operation they are useful in checking the antenna's performance, and in pointing out site deficiencies which may be corrected by relocation of the antenna. Subsequently, routine pattern checks may be necessary to insure continued performance of the antenna. Since the antenna is assumed to scan, the first measurement must provide the radiation pattern as a function of scan angle, if terrain effects are to be fully accounted for. This can be done either by fixing the beam in a given direction and measuring the field strength while flying through it, or by using a helicopter or balloon hovering in a fixed position and scanning the beam through this position. If the first of these methods is used it is essential that the aspect of the aircraft as well as its distance from the antenna be carefully maintained. This can be accomplished

with the aid of a gyroscopic stabilization device and RF distance-measuring equipment. In either case a large number of measurements may be necessary in order to cover the sector of interest and to insure repeatable results. The routine pattern check, on the other hand, need only be made for a single position in the scan, since its purpose is to check for continued antenna performance.

Radio astronomy antennas fall into two classes: those which are completely steerable, and those which in part rely on the rotation of the earth to provide the necessary scanning action [111]. The former, of necessity, have dimensions not over a few hundred feet, and since they are steerable, lend themselves to the methods used for radar antennas. The semifixed antennas, on the other hand, may have dimensions measured in thousands of feet if they are of the interferometer type, so that a measurement using an aircraft or helicopter becomes impractical if carried out in the far field. An alternative to a near-field measurement is to make use of extraterrestrial sources such as the sun or radio stars. Here special techniques are required because of the finite angular extent of such sources [86], which result in a so-called "drift curve" rather than a true radiation pattern. The two may be related provided some knowledge of the source distribution is available. For example, if a source has a reasonably uniform brightness with a known extent determined from prior measurement using a different antenna of known characteristics, a reasonable approximation to the true antenna pattern may be possible.

It should be emphasized that the field of measurements on very large antennas is a new one, and that techniques have by no means been perfected.

The effects of a number of environmental conditions on the performance of an antenna have been discussed. In addition, many airborne antennas and a large class of ground antennas are also influenced by the radome or housing used to cover them. In spite of careful design, a radome may affect the performance of an antenna in a number of ways. First, it may distort the wavefront emerging from the antenna, thereby modifying the gain, beamwidth, sidelobe level, and beam position. Second, it may absorb and reflect some of the energy passing through it. Until recently these effects were confined chiefly to radomes covering airborne antennas, where the electrical design must usually be compromised for aerodynamic reasons. However, with the advent of large rigid radomes for use with ground antennas, the scope of the problem has been increased.

The measurement of the effect of the radome on the pattern and on the power gain of an antenna is based on the techniques used to study the antenna alone, and requires a comparison of these quantities with and without the radome. Measurement of boresight error, on the other hand, requires special techniques, since certain systems demand that this quantity be held to the order of $1/100^\circ$ or better. This is particularly true of track-

ing antennas which use either a conical scan or a monopulse system in order to produce a well-defined minimum or null in the radiation pattern. In the case of airborne antennas of this type, it has become common practice to measure boresight error with the type of instrument shown diagrammatically in Fig. 16. The anechoic chamber in which the equipment is housed has been described earlier and needs no further comment. The antenna and associated radome are mounted on a special platform which makes it possible to simulate the antenna scan relative to the radome, by fixing the antenna and moving the radome. Considering the transmitting case, it is evident that as the radome moves, the position of the null axis of the antenna pattern may move in space, and it is the amount and direction of this shift relative to the position of the null axis in the absence of the radome that is called the boresight shift or boresight error. By measuring the linear displacement of the null at a range within the far-zone of the antenna, it is evident that the required angular sensitivity can be obtained. This is accomplished in one system by using a small receiving antenna as a field probe to determine the position in space of the null axis. The received signal from this probe is fed into a servosystem which automatically moves the probe on two orthogonal axes in such a way as to produce minimum received signal. By recording the position of the probe or null-seeker as a function of the relative orientation of the antenna and radome, the boresight error curve is obtained.

The direct application of the null-seeking technique to large ground antennas and radomes becomes impractical because of the large travel required of the null-seeker for even small boresight errors. Since the speed with which the null-seeker moves is limited, a large travel requires an excessive amount of time in order to make a measurement. For example, consider a 30-foot antenna operating at *S* band. In order to satisfy the usual criterion for pattern measurements, a range of some 5300 feet is required, and at this range a boresight error of perhaps $1/20^\circ$ represents a linear displacement of some 4.5 feet. A further complication in the testing of large radomes lies in the time required for assembly of the radome, which makes it very difficult to make a rapid comparison of the performance of the antenna with and without the radome. One interesting solution to these problems has been proposed by Miller and Lavrench [112]. The method used to permit rapid comparison of the patterns and gain with and without the radome is shown in Fig. 22. Here a tripod of curved members is erected on a carriage which is mounted on a circular track. The antenna to be used in the tests rests on a separate fixed platform within the circular track. The radome under investigation is erected in the usual manner on a base-ring fitted to the top of the carriage. It is then rigidly fastened to the tripod, so that a sector comprising one third of the surface area of the radome can be removed. Since the radome can be rotated by means of the carriage, rapid comparison of the antenna

parameters with and without the radome is possible. The radome shown in the photograph has an equatorial diameter of 55 feet and encloses an antenna of major dimensions 30 feet by 7.5 feet.

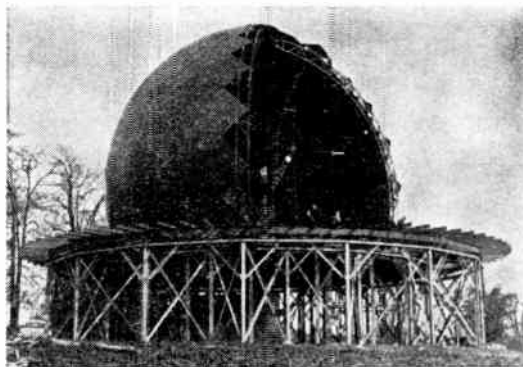


Fig. 22—A test site for large rigid radomes (National Research Council of Canada photograph).

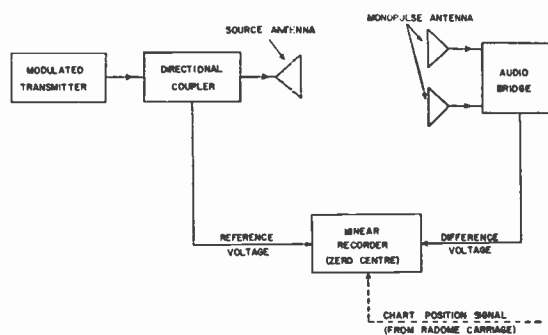


Fig. 23—Block diagram of boresight-measuring equipment.

A block diagram of the equipment used for automatic boresight measurements on a two-channel amplitude-comparison monopulse antenna is shown in Fig. 23. A remote transmitting antenna is fastened to a carriage which rides on a track which may be mounted either vertically or horizontally. A reference voltage obtained through a directional coupler is fed over land lines to a linear recorder of the zero-center type. The audio outputs of the detectors which terminate the monopulse feeds are subtracted in a bridge and the difference voltage applied to the recorder. Chart-positioning information is derived from selsyns connected to the carriage on which the radome is mounted. To use the instrument, the radome is rotated until there is a clear path between the source antenna and the test antenna. The system is then calibrated by first locating the source antenna at the center of its travel, and by aligning the test antenna so that its axis is directed at the source. Final balancing for a null is achieved by means of a potentiometer incorporated in the bridge circuit. Correlation between boresight error and pen position is then obtained by moving the source antenna in known steps and noting the pen position. After the system has been calibrated, a boresight-error curve for the plane in question is obtained by automatically recording the bridge output as

a function of radome position. Boresight errors of $1/200^\circ$ may be measured with this equipment, which is also adaptable to a phase comparison monopulse antenna. It should be pointed out, however, that when measuring vertical boresight error, extreme care must be exercised to avoid effects from ground reflections.

B. Polarization and Phase Measurements

Compared with radiation pattern measurements, polarization and phase measurements are of limited interest, chiefly because of the fact that these quantities are generally known with sufficient accuracy for measurements to be unnecessary. Insofar as polarization is concerned, the obvious exceptions to this include antennas designed to produce circular or nearly circular polarization, and antennas with nominally linear polarization which because of their geometry are known to produce cross-polarized components. Similarly, far-zone phase information, usually given by the position of the phase center, is required only for primary radiators and in connection with certain Doppler tracking systems and radar systems [113], [114].

An elliptically-polarized wave can be described in terms of two orthogonal linearly-polarized components, or in terms of two circularly-polarized components of opposite rotational sense. Logically then, measurements of the polarization of an elliptically polarized antenna can be made using either one or more linearly polarized source antennas or two circularly polarized source antennas [115], [116]. The method most commonly used employs a single linearly polarized source antenna, provided with an automatic rotating device. By measuring the response of the test antenna as a function of the rotation of the source antenna, a polarization pattern is obtained. The alignment of this figure with respect to the coordinate system and the ratio of its major-to-minor dimensions may then be used to derive the amplitudes and the relative phase of the two linear polarized components. While there remains an ambiguity in the sign of this phase angle, the method is attractive because it is compatible with a pattern-measuring installation, requiring in addition to the standard equipment only a rotating device for the source antenna and a polar recorder on which to record the polarization pattern. It is, however, a time-consuming method, since polarization patterns must in general be measured at a number of angular positions within the radiation pattern of the antenna. An alternative system which is much more rapid uses a PPI presentation of the polarization pattern so that the source antenna may be rotated at high speed.

If it is necessary to resolve the phase ambiguity associated with the rotating antenna method, two auxiliary source antennas are required, polarized left-hand circular and right-hand circular, respectively. By connecting each in turn to the transmitter and noting which produces the greater response in the test antenna, the sense of rotation and hence the sign of the phase angle

may be determined.

However, if it is necessary to resort to the use of the two circularly polarized antennas in order to obtain phase information, an alternative technique is available whereby these antennas may be used to completely determine the nature of the polarization. In this method the transmitter is connected alternatively to one and then to the other source antenna. Denoting the relative responses by E_R and E_L , the axial ratio of the polarization ellipse is then given by

$$\text{axial ratio} = \frac{E_R + E_L}{E_R - E_L}.$$

The tilt angle τ of the polarization ellipse with respect to the horizontal may be obtained by connecting the source antennas in parallel and rotating one of them for minimum response. If the rotation required is δ_1 relative to the position for minimum response with a horizontally polarized test antenna, then the value of τ is given by $\tau = \delta/2$.

Finally, the most direct approach to the problem is to use two orthogonal linearly polarized source antennas interconnected through a calibrated variable attenuator and phase-shifter. With this arrangement the amplitudes and the relative phase angle of the linear components can be determined directly, without ambiguity, by adjusting the attenuator and phase-shifter for maximum response and noting their settings. In the receiving case, the result is more readily obtained by first connecting the receiver to the two antennas in succession, thereby obtaining the ratio of the two components and then measuring their relative phase.

A number of methods of measuring the phase of an electromagnetic field were outlined in Section II. In these methods the emphasis lay in determining the phase contours of the near field, contours which, in general, are complex in shape. However, in the far-zone of an antenna the phase front is essentially spherical, at least over the region of the main beam; or in certain special cases, cylindrical. For example, a parabolic-cylinder antenna is designed so that the reflector is in the cylindrical wave zone of the feed.

In any case, far-zone phase measurements are almost always made in order to determine the location of the center of the spherical or cylindrical phase front; or at least of the sphere or cylinder which gives the best fit to the actual phase front. The techniques described previously lend themselves readily to this measurement in the case of primary feeds, since the separation between the feed under study and the test probe is small. The simplest method is to measure the phase as a function of angle by either rotating the antenna under test, or by moving the test probe in an arc. Simple trigonometry can then be used to determine the position of the phase center in terms of the measured phase and the radius of the test arc [47]. If more elaborate RF circuitry is used, so that the radial separation of the test probe and the antenna under test may be varied, the equiphase con-

tour may be traced out directly and hence the center of phase may also be directly determined. In determining the phase center of a large radar antenna, the distance to the test probe may be several hundred feet. This makes it necessary to use a microwave link to transmit the received signal or the comparison signal [114]. Apart from this complication, the techniques used are the same as those used for smaller antennas.

VI. GAIN, DIRECTIVITY, AND EFFICIENCY

The gain, directivity, and efficiency of an antenna are derived quantities, in contrast with the field-strength and radiation patterns which are more fundamental characteristics. To indicate the manner in which they are defined [117], consider an antenna which is supplied with input power P_{in} . If the radiated power density per unit solid angle is given by

$$P(\theta, \phi) = P_0 \{ F_\theta^2(\theta, \phi) + F_\phi^2(\theta, \phi) \}$$

where $F_\theta(\theta, \phi)$ and $F_\phi(\theta, \phi)$ are the normalized radiation patterns of the θ and ϕ components of the electric field vector, respectively, then the total radiated power W is given by

$$W = \int_{4\pi} P(\theta, \phi) d\Omega$$

where $d\Omega$ is the element of solid angle. The efficiency of the antenna is then defined as the ratio of the total radiated power to the input power, $\eta = W/P_{in}$. The "total" directive gain in a given direction is defined as 4π times the ratio of the total radiation intensity in that direction to the radiated power W . The "total" power gain is defined as 4π times the ratio of the total radiation intensity in that direction to the input power P_{in} .

In the usual case of an essentially linearly polarized antenna, the directive gain and power gain associated with the principal polarization are defined in the same manner, except that the radiation intensity referred to is that of the principal polarization component only. Thus the directive gain becomes

$$G_0 = \frac{4\pi P_0 F_\theta^2(\theta, \phi)}{\int_{4\pi} P_0 \{ F_\theta^2(\theta, \phi) + F_\phi^2(\theta, \phi) \} d\Omega}$$

and the directivity, defined as the maximum value of the directive gain, becomes

$$D = \frac{4\pi}{\int_{4\pi} \{ F_\theta^2(\theta, \phi) + F_\phi^2(\theta, \phi) \} d\Omega}$$

From the standpoint of measurements, it is evident that the directivity can be derived from the measured three-dimensional radiation pattern of the antenna, while the maximum power gain and efficiency must be obtained from a measurement which takes into account the ohmic losses of the antenna. At microwave frequen-

cies the efficiency is usually so high that the directivity and maximum power gain can be considered equal, but such is not the case for antennas operating in the VHF band and below. It is therefore convenient to divide this section into three parts dealing with: 1) the determination of directivity from radiation patterns; 2) the direct measurement of power gain, and if the efficiency is high, of directivity; and 3) the measurement of efficiency.

If the pattern functions $F_\theta(\theta, \phi)$ and $F_\phi(\theta, \phi)$ of an antenna can be expressed analytically rather than graphically, it is frequently possible to perform the integration necessary to compute the directivity. It follows that if the measured patterns of an antenna are similar in shape, within arbitrary limits, to those of some standard antenna whose gain can be computed, then the approximate gain of the antenna can be determined from beamwidths of the measured patterns [115]. For example, if an antenna is found by measurement to have essentially a single-lobed pattern, the approximate directivity is given by

$$D = \frac{41,253}{\phi_1^\circ \phi_1^\circ},$$

where θ_1° and ϕ_1° are the half-power beamwidths in degrees. Similarly, an antenna having a broadcast or doughnut-type pattern has a directivity given by

$$D = \frac{1}{\sin \frac{\theta_1^\circ}{2}}.$$

These simple expressions relating the directivity and the beamwidth are obviously limited in their usefulness, and if it is necessary to determine the directivity of an antenna having numerous sidelobes or a main beam of unusual shape, more sophisticated methods must be employed [118], [120] in which the integral

$$\int_{4\pi} \{ F_\theta^2(\theta, \phi) + F_\phi^2(\theta, \phi) \} d\Omega$$

is evaluated by graphical means or by a computer. In either case the amount of information that is required concerning the three-dimensional pattern depends on the nature of that pattern. To illustrate, consider a pencil-beam antenna (one having essentially equal beamwidths in the E and II planes), and suppose that this beam is aligned parallel to the Z axis of a conventional spherical coordinate system. If the antenna has a single component of polarization it is seen at once that because of the circular symmetry of the beam the integral can be written,

$$\begin{aligned} \int_{4\pi} F_\theta^2(\theta, \phi) d\Omega &= 2\pi \int_{\theta=0}^{\pi} F_\theta^2(\theta) \sin \theta d\theta \\ &= -2\pi \int_{\theta=0}^{\pi} F_\theta^2(\theta) d(\cos \theta). \end{aligned}$$

Thus in order to obtain the gain by graphical means it is necessary only to measure a single pattern, $F_{\theta}(\theta)$, and to plot its square $F_{\theta}^2(\theta)$ on rectangular coordinates as a function of $\cos \theta$. The integral can then be evaluated by the usual means, *i.e.*, Simpson's rule, planimeter, etc. On the other hand, suppose the antenna, while having only one component of polarization, lacks circular symmetry. If the pattern data are available in the form of conical sections, *i.e.*, $F_{\theta}(\theta, \phi)$ vs ϕ for a series of values of θ , the integral may be written,

$$\begin{aligned} \int_{4\pi} F_{\theta}^2(\theta, \phi) d\Omega &= \int_{2\pi} d(\cos \phi) \int_0^{2\pi} F_{\theta}^2(\theta, \phi) d\phi \\ &= 2 \int_0^{\pi} d(\cos \theta) \int_0^{2\pi} 1/2 F_{\theta}^2(\theta, \phi) d\phi. \end{aligned}$$

Since the patterns are plotted in polar form it is evident that the area of a given polar plot is given by the second integral, namely:

$$A(\theta) = \int_0^{2\pi} 1/2 F_{\theta}^2(\theta, \phi) d\phi$$

and so the complete integral may be written,

$$\int_{4\pi} F_{\theta}^2(\theta, \phi) d\Omega = \int_0^{\pi} A(\theta) d(\cos \theta).$$

Therefore, in order to obtain the directivity, first a series of conical patterns are measured with intervals of θ sufficiently small to include the pattern detail. The area $A(\theta)$ of each polar plot is determined with a planimeter, and plotted on rectangular coordinates against $\cos \theta$. The area of the resulting figure, measured with a planimeter, represents the antenna's directivity. This method is well adapted to obtaining the directivity of aircraft antennas, since the pattern data are usually provided in the form of conical cuts for ease in interpretation. Furthermore, the operations of measuring the area of each polar pattern, and of performing the summation in θ are easily carried out automatically by using a ball-and-disk or an electronic integrator. By making provision in the positioning device for the angle θ to sweep slowly through 180° as the angle ϕ changes continuously at a faster rate, a fully automatic instrument can be constructed, in which the data are recorded and processed simultaneously.

Regardless of whether the integration is performed graphically or with the aid of a computer, the final figure for directivity is based on the antenna pattern. With highly directional antennas, the variation in field strength with angle becomes quite rapid, and an accurate determination of directivity by these means is limited by the accuracy with which the beamwidth or pattern can be measured. Furthermore, in most cases a knowledge of power gain is more important from the point of view of operation of the system than is a knowledge of the directivity. As a result, measurements of the former quantity are more frequently made when assessing the performance of an antenna.

The definition of "power gain" which has been given can be interpreted to represent the power density at a given distance from an antenna as compared with that which would be produced by a hypothetical antenna radiating the same input power, but radiating it isotropically, or uniformly in all directions. Since gain is a dimensionless ratio, this suggests that it might be measured by comparing one antenna with another of known characteristics. Various methods have been proposed for making this comparison, and for establishing a reference antenna of known gain with respect to an isotropic antenna.

In the microwave-frequency range, horn antennas are most commonly used as gain standards. They are relatively easy to construct, and are efficient, so that almost always the directivity and power gain are equal, and in addition, their directivity and gain can be calculated or measured fairly readily. The size of a gain standard is chosen with two factors in mind. First, it must have suitable physical dimensions, which means it must be small enough to mount conveniently, and to orient. Second, its gain must not be too different from that of the antenna to be measured, otherwise the comparison will be inaccurate. These requirements often conflict so that a compromise must be made, and gain-standard horns are usually chosen to have a gain in the region 10 to 25 db above isotropic. The gain of a standard horn can be determined: 1) by calculation, 2) direct measurement using two or more antennas, and 3) direct measurement using a single antenna [47], [48], [121]–[126]. Provided the horn is not too small, *i.e.*, aperture dimensions greater than approximately three wavelengths, the gain can be calculated accurately to within some ± 0.25 db for a horn having a gain of 20 db. However, in the case of a small horn, radiation from currents on the outside surfaces of the horn becomes significant and is difficult to calculate. Also, if the horn is not symmetrical, higher modes may exist in the aperture, resulting in a gain different from that calculated. In these cases the gain must be determined by measurement.

The two-antenna method requires that two antennas be constructed as identically as possible. After careful matching one antenna is connected to a transmitter which is fitted with a monitoring device to indicate output. The second antenna is located a distance R away from the first, and terminated in a matched receiver. As has been shown by Braun [49], R must be very large if the true far-zone gain of the antennas is to be measured directly. Even at distances as great as $R = 16D^2/\lambda$, or eight times the usual far-zone criterion for pattern measurements, the gain may be 0.1 db below that for $R = \infty$. This is entirely due to path-length variations from a given point in one aperture to a given point in the other aperture. Fortunately the magnitude of the effect may be calculated, so that if the gain is measured at a given range, say $R = 2D^2/\lambda$, it may be calculated for $R = \infty$. If the test site has been chosen carefully so that only a single propagation path exists between the two

antennas, the power delivered to the receiver is given by

$$P_R = \frac{\lambda^2 G^2}{(4\pi R)^2} P_t,$$

where G is the gain of each of the identical antennas, and P_t is the transmitted power. The ratio P_R/P_t may be determined most accurately by first connecting the receiver to the receiving antenna, as has been mentioned, and noting the reading. The receiver is then connected directly to the transmitter through a precision attenuator which is adjusted until the same reading is again obtained. The amount of attenuation inserted is then a measure of P_R/P_t , from which the gain may be computed, since

$$G = \frac{4\pi R}{\lambda} \sqrt{P_R/P_t}.$$

In some cases it may be doubtful whether or not the two antennas have identical gain. Suppose, for example, one has gain G_1 and the other gain G_2 . Then the single measurement yields the result,

$$\sqrt{G_1 G_2} = \frac{4\pi R}{\lambda} \sqrt{P_R/P_t}.$$

The ambiguity may be resolved by using a third antenna as a source, and by comparing the relative power received by antennas no. 1 and no. 2 at the test site. If, for example, antenna no. 1 receives power P_1 , and antenna no. 2 receives power P_2 , then it is easy to see that

$$G_1 = \sqrt{P_1/P_2} \cdot G \text{ and } G_2 = \sqrt{P_2/P_1} \cdot G.$$

It must be emphasized that a measurement of this sort requires extreme care, and that the site must be free from scattering or reflecting obstacles. If all necessary precautions are observed, a 20 db-gain standard in the microwave region may be established to within ± 0.10 db.

An alternative to the three-antenna method was first described by Purcell [126]. A carefully matched antenna is connected to a transmitter through a standing-wave detector. At a distance $R/2$ from the antenna, a large plane mirror is placed. Energy striking the mirror will be reflected back to the antenna where a portion will be absorbed and a portion reradiated. The net result of this multiple-scattering process is that a reflected wave is set up on the transmission line feeding the antenna. A number of precautions are necessary when making this measurement. First, R must be chosen to be at least equal to $2D^2/\lambda$, and the measured gain corrected as described earlier, to obtain the gain at infinity. Second, the mirror must be perpendicular to the propagation path so that the reflection coefficient has a magnitude of unity. Finally, the interaction effects must be taken into account. As was shown by Pippard, Burrell, and Cromie [48], if the measured standing-wave ratio is plotted as a function of R , the resulting curve will have

a cyclical variation due to multiple scattering. If the mean curve is used it can be shown that the gain is given by

$$G = \frac{4\pi R}{\lambda} \left(\frac{\rho - 1}{\rho + 1} \right),$$

where ρ is the mean standing-wave ratio associated with the distance R . Purcell's method is attractive since only one antenna is needed, but it is limited to small antennas, since the mirror must be large enough to intercept all the energy from the antenna.

Below the microwave region, horn antennas become too large for use as gain standards, and other types must be used. The corner reflector has proven useful in the UHF and upper-VHF portion of the spectrum [127], while at lower frequencies, the half-wave dipole or quarter-wave unipole over ground is used. Since the gain of these last two antennas can be calculated, it is not usually necessary to measure their gain, unless it is suspected that their efficiencies are low. Once a gain standard has been established, the gain of an unknown antenna can be determined by comparison. The technique is to set up a source antenna as is done for pattern measurements, and to compare the measured responses of the unknown antenna and of the gain standard, using a matched receiver. To prevent the receiver characteristics from influencing the measurements, a precision calibrated attenuator may be used ahead of the receiver to maintain constant output.

The precautions mentioned in connection with determining the gain of the standard antenna must also be observed in making gain-comparison measurements. First, the relation between the measured gain and the separation between source and test antennas must be carefully examined if a figure for the true far-zone gain is to be obtained. Second, the site must be carefully chosen to insure that the response of the unknown and of the reference antennas are compared with respect to the same incident wave. If the site is of the elevated type, in which ground reflections are reduced to negligible proportions, this presents no great problem. In the case of a site in which the reflected wave forms part of the total incident field, it is important that either the vertical variation of field across the test antenna aperture be small, or else that the test antenna and standard antenna have similar vertical apertures, so that field variations will be common to both.

A method in which small-field variations are averaged out, and which uses an additional "reference" antenna, is described by Cottony [128]. In his method, a remote source antenna is mounted some three wavelengths above ground. At the test site approximately one half mile away, a reference antenna is located, at approximately the same height above ground as the source antenna. In order to carry out a gain measurement, a standard antenna (commonly a half-wave dipole) is erected at the height and location at which the

unknown antenna is to be erected. This standard antenna is moved over the region to be occupied by the unknown antenna in steps of 0.1 wavelength, and its output relative to that of the reference antenna is noted. The comparison in signal level is made at IF using a piston attenuator as a standard. Following the complete exploration of the test region, the average difference in response is calculated, and the standard antenna is replaced by the unknown. The attenuator is then used to adjust the output of the unknown until the same difference is maintained with respect to the reference antenna.

This method of measuring the gain is interesting, since it takes into account minor field variations, and further, makes it simpler to deal with equipment instability. An accuracy of the order of 0.1 db in the 300 to 400-mc region is claimed.

Turning briefly to the question of efficiency, it is apparent that if the power gain of an antenna has been measured and if the directivity has been determined from the radiation pattern, the efficiency can be obtained from the ratio of the two quantities. An interesting alternative method of measuring efficiency has been proposed by Crowley [129], in which two antennas are constructed which are identical in all respects except the conductivity of the metal used. By measuring the power gain of the two antennas it is possible to compute the efficiency without independently determining the directivity.

ACKNOWLEDGMENT

The author is indebted to his colleagues in the laboratories of the National Research Council for their helpful advice and criticism, and to the many individuals in other laboratories who contributed directly or indirectly to this paper.

BIBLIOGRAPHY

- [1] J. F. Ramsay, "Microwave antenna and waveguide techniques before 1900," *Proc. IRE*, vol. 46, pp. 405-415; February, 1958.
- [2] W. A. Cumming, C. P. Wang, and S. C. Loh, "Analysis and reduction of scattering from the feed of a cheese antenna," 1958 IRE CANADIAN CONVENTION RECORD (in preparation).
- [3] R. King and T. T. Wu, "The Reflection of Electromagnetic Waves from Surfaces of Complex Shape, I—Experimental Studies," Cruft Lab., Harvard University, Cambridge, Mass., Sci. Rep. No. 12, Contract AF19(604)-786; November, 1957.
- [4] R. V. Row, "Microwave diffraction measurement in a parallel-plate region," *J. Appl. Phys.*, vol. 24, pp. 1448-1452; December, 1953.
- [5] G. Bekefi, "Studies in Microwave Optics," The Eaton Electronics Res. Lab., McGill University, Montreal, Can., Tech. Rep. No. 38; March, 1957.
- [6] R. M. Barrett and M. H. Barnes, "Automatic antenna wave-front plotter," *Electronics*, vol. 25, pp. 120-125; January, 1952.
- [7] F. L. Vernon, Jr., "Application of the microwave homodyne," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-4, pp. 110-116; December, 1952.
- [8] H. R. Worthington, "Measurements of Phase in Microwave Antenna Fields by Phase Modulation Method," M.I.T. Rad. Lab., Cambridge, Mass., Rep. No. 966; March, 1946.
- [9] J. C. Cacheris and H. A. Dropkin, "Compact microwave single-sideband modulator using ferrites," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-4, pp. 152-155; July, 1956.
- [10] J. Bacon, "An X band phase plotter," *Proc. Natl. Electronics Conf.*, vol. 10, pp. 256-263; October, 1954.
- [11] J. H. Richmond, "Measurement of time-quadrature components of microwave signals," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-3, pp. 13-15; April, 1955.
- [12] H. A. Dropkin, "Direct reading microwave phase-meter," 1958 IRE NATIONAL CONVENTION RECORD, pt. I, pp. 57-63.
- [13] J. S. Ajioka, "A microwave phase contour plotter," *Proc. IRE*, vol. 43, pp. 1088-1090; September, 1955.
- [14] J. N. Hines and C. H. Bochner, "Measurement of the phase of radiation from antennas," *Proc. Natl. Electronics Conf.*, vol. 4, pp. 487-495; October, 1948.
- [15] E. W. Hamlin and W. E. Gordon, "Comparison of calculated and measured phase difference at 3.2 centimeters wavelength," *Proc. IRE*, vol. 36, pp. 1218-1223; October, 1948.
- [16] W. Sichak and D. J. Levine, "Microwave high-speed continuous phase shifter," *Proc. IRE*, vol. 43, pp. 1661-1663; November, 1955.
- [17] S. D. Robertson, "A method of measuring phase at microwave frequencies," *Bell Sys. Tech. J.*, vol. 28, pp. 99-103; January, 1949.
- [18] Y. P. Yu, "Measuring phase at RF and video frequencies," *Electronics*, vol. 29, pp. 138-140; January, 1956.
- [19] R. D. Kodis, "An Experimental Investigation of Microwave Diffraction," Cruft Lab., Harvard University, Cambridge, Mass., Tech. Rep. No. 105; June, 1950.
- [20] S. D. Hamren, "Scattering from Spheres," Antenna Lab., University of California, Berkeley, Calif., Tech. Rep. No. 171; June, 1950.
- [21] B. N. Harden, "Diffraction of electromagnetic waves by two parallel half-planes," *Proc. IEE*, vol. 100, pp. 348-350; November, 1953.
- [22] L. R. Allredge, "Tandem Slit Diffraction Measurements," Cruft Lab., Harvard University, Cambridge, Mass., Tech. Rep. No. 176; May, 1953.
- [23] C. Huang and R. D. Kodis, "Diffraction by Spheres and Edges at 1.25 CMS," Cruft Lab., Harvard University, Cambridge, Mass., Tech. Rep. No. 138; November 30, 1951.
- [24] H. L. Robinson, "Diffraction in circular apertures less than one wavelength in diameter," *J. Appl. Phys.*, vol. 24, pp. 35-38; January, 1953.
- [25] C. L. Andrews, "Diffraction patterns in a circular aperture measured in the microwave region," *J. Appl. Phys.*, vol. 21, pp. 761-767; August, 1950.
- [26] G. Bekefi, "Diffraction of electromagnetic waves by an aperture in an infinite screen," *J. Appl. Phys.*, vol. 23, p. 1403; December, 1952.
- [27] D. C. Hogg, "The electromagnetic field in the plane of a circular aperture due to incident spherical waves," *J. Appl. Phys.*, vol. 24, pp. 110-111; January, 1953.
- [28] D. C. Hogg, "The electromagnetic field near a dielectric lens," *J. Appl. Phys.*, vol. 25, p. 542; April, 1954.
- [29] J. R. Mentzer, "The use of dielectric lenses in reflection measurements," *Proc. IRE*, vol. 41, pp. 252-256; February, 1953.
- [30] R. W. P. King, "The Loop Antenna as a Probe in Arbitrary Electromagnetic Fields," Cruft Lab., Harvard University, Cambridge, Mass., Tech. Rep. No. 262; May, 1957.
- [31] Ronald King, "The Rectangular Loop Antenna as a Dipole," Cruft Lab., Harvard University, Cambridge, Mass., Tech. Rep. No. 263; May, 1957.
- [32] R. E. Houston and R. N. Noble, "Edge effects in circular apertures," *J. Appl. Phys.*, vol. 22, p. 1295; October, 1951.
- [33] R. G. Kouyoumjian, "Back Scattering from a Circular Loop," Eng. Exper. Station, The Ohio State University, Columbus, Ohio; Bull. 162; November, 1956.
- [34] P. D. Kennedy, "Equipment and techniques for the measurement of radar reflections from model targets," 1957 WESCON CONVENTION RECORD, pt. 1, pp. 208-215.
- [35] H. J. Schmitt, "Back-Scattering Measurements with a Space-Separation Method," Cruft Lab., Harvard University, Cambridge, Mass., Sci. Rep. No. 14; November, 1957.
- [36] H. Scharfman and D. D. King, "Antenna-scattering measurements by modulation of the scatterer," *Proc. IRE*, vol. 42, pp. 854-858; May, 1954.
- [37] C. C. H. Tang, "Backscattering from dielectric-coated infinite cylindrical obstacles," *J. Appl. Phys.*, vol. 28, pp. 628-633; May, 1957.
- [38] C. C. H. Tang, "Back-Scattering Measurements with a Pulse Method," Cruft Lab., Harvard University, Cambridge, Mass., Sci. Rep. No. 15; 1958.
- [39] J. M. Minkowski and E. S. Cassidy, "Cross-section of colinear arrays at normal incidence," *J. Appl. Phys.*, vol. 27, pp. 313-316; April, 1956.
- [40] S. H. Dike and D. D. King, "The absorption gain and back-scattering cross section of cylindrical antennas," *Proc. IRE*, vol. 40, pp. 853-860; July, 1952.

- [41] D. D. King, "The measurement and interpretation of antenna scattering," *PROC. IRE*, vol. 37, pp. 770-777; July, 1949.
- [42] M. H. Cohen and R. C. Fisher, "A dual-standard for radar echo measurements," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-3, pp. 108-110; July, 1955.
- [43] C. R. Grant and B. S. Yaplee, "Back-scattering from water and land at centimeter and millimeter wavelengths," *Proc. IRE*, vol. 45, pp. 976-982; July, 1957.
- [44] J. Sevick, "Experimental and Theoretical Results on the Back-Scattering Cross Section of Coupled Antennas," *Cruft Lab., Harvard University, Cambridge, Mass., Tech. Rep. No. 150*; May, 1952.
- [45] J. Honda, L. J. Black, and F. D. Clapp, "Techniques of Measuring Scattered Fields," *Inst. Eng. Res., University of California, Berkeley, Calif., Ser. No. 60, Issue No. 127*; October, 1954.
- [46] R. M. Redheffer, "The Techniques of Microwave Measurements," *M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 11*, pp. 591-606; 1947.
- [47] D. D. King, "Measurements at Centimeter Wavelengths," *D. Van Nostrand Co., Inc., New York, N. Y.*, pp. 291-294; 1952.
- [48] A. B. Pippard, O. J. Burrell, and E. E. Cromie, "The influence of re-radiation on measurements of the power gain of an aerial," *J. IEE*, vol. 93, pt. 111A, pp. 720-722; 1946.
- [49] E. H. Braun, "Gain of electromagnetic horns," *Proc. IRE*, vol. 41, pp. 109-115; January, 1953.
- [50] H. M. Barlow and A. L. Cullen, "Microwave Measurements," *Constable and Co., Ltd., London, Eng.*, pp. 299-301; 1950.
- [51] W. A. Cumming, "The dielectric properties of ice and snow at 3.2 centimeters," *J. Appl. Phys.*, vol. 23, pp. 768-772; July, 1952.
- [52] M. J. Kofoid, "Automatic measurement of phase retardation for radome analysis," *Rev. Sci. Instr.*, vol. 27, pp. 450-452; July, 1956.
- [53] R. Justice and V. H. Rumsey, "Measurement of electric field distributions," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-3, pp. 177-180; October, 1955.
- [54] A. L. Cullen and J. C. Parr, "A new perturbation method for measuring microwave fields in free space," *Proc. IEE*, vol. 102, pt. B, pp. 836-844; November, 1955.
- [55] J. H. Richmond, "A modulated scattering technique for measurement of field distributions," *IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-3, pp. 13-15; July, 1955.
- [56] E. A. Killick, "Phase and amplitude measurements by means of the reflection from a rotating dipole," *Admiralty Signals and Radar Establ., Portsmouth, Eng. ASRE Tech. Note AX-55-15*; September, 1955.
- [57] J. H. Richmond and T. E. Tice, "Probes for microwave near-field measurements," *IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-3, pp. 32-34; April, 1955.
- [58] V. H. Rumsey, "Reaction concept in electromagnetic theory," *Phys. Rev.*, vol. 94, pp. 1483-1491; June 15, 1954.
- [59] H. W. Ehrenspeck and W. J. Kearns, "Two-dimensional end-fire array with increased gain and sidelobe reduction," *1957 WESCON CONVENTION RECORD*, pt. 1, pp. 217-221.
- [60] R. Mitra, "An automatic phase-measuring circuit at microwaves," *IRE TRANS. ON INSTRUMENTATION*, vol. I-6, pp. 238-240; December, 1957.
- [61] R. W. P. King, "The Theory of Linear Antennas," *Harvard University Press, Cambridge, Mass.*, pp. 127-141; 1956.
- [62] J. V. N. Granger and T. Morita, "Radio-frequency current distribution on aircraft structures," *Proc. IRE*, vol. 39, pp. 932-938; August, 1951.
- [63] I. Carswell, "Current distribution on wing-cap and tail-cap antennas," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-3, pp. 207-212; October, 1955.
- [64] R. E. Webster, "20-70 mc monopole antennas on ground-based vehicles," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 363-368, October, 1957.
- [65] H. E. Gihring and G. H. Brown, "General consideration of tower antennas for broadcast use," *Proc. IRE*, vol. 23, pp. 311-356; April, 1935.
- [66] T. Morita, "Current distributions on transmitting and receiving antennas," *Proc. IRE*, vol. 38, pp. 898-904; August, 1950.
- [67] T. Morita and C. E. Farley, "The measurement of current distributions along coupled antennas and folded dipoles," *Proc. IRE*, vol. 39, pp. 1561-1565; December, 1951.
- [68] P. A. Kennedy, "Loop antenna measurements," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-4, pp. 610-618; October, 1956.
- [69] G. Barzilai, "Experimental determination of the distribution of current and charge along cylindrical antennas," *Proc. IRE*, vol. 37, pp. 825-829; July, 1949.
- [70] J. A. Marsh, "Current distributions on helical antennas," *Proc. IRE*, vol. 39, pp. 668-675; June, 1951.
- [71] G. Sinclair, "Theory of models of electromagnetic systems," *Proc. IRE*, vol. 36, pp. 1364-1370; November, 1948.
- [72] H. G. Booker, V. H. Rumsey, G. A. Deschamps, M. L. Kales, and J. I. Bohnert, "Techniques for handling elliptically polarized waves with special reference to antennas," *Proc. IRE*, vol. 39, pp. 533-552; May, 1951.
- [73] G. Sinclair, "The transmission and reception of elliptically polarized waves," *Proc. IRE*, vol. 38, pp. 148-151; February, 1950.
- [74] D. K. Cheng, "Microwave aerial testing at reduced ranges," *Wireless Eng.*, vol. 33, pp. 234-237; October, 1956.
- [75] C. C. Cutler, A. P. King, and W. E. Kock, "Microwave antenna measurement," *Proc. IRE*, vol. 35, pp. 1462-1471; December, 1947.
- [76] S. Silver, "Microwave Antenna Theory and Design," *M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 12*, pp. 543-613; 1949.
- [77] D. R. Rhodes, "On minimum range for radiation patterns," *Proc. IRE*, vol. 42, pp. 1408-1410; September, 1954.
- [78] E. C. Jordan, "Electromagnetic Waves and Radiating Systems," *Prentice-Hall, Inc., New York, N. Y.*, sec. 15.06, p. 574; 1950.
- [79] G. A. Woonton, R. B. Borts, and J. A. Carruthers, "I. Indoor measurement of microwave antenna radiation patterns by means of a metal lens," *J. Appl. Phys.*, vol. 21, pp. 428-430; May, 1950.
- [80] W. H. Emerson, A. G. Sands, and M. V. McDowell, "Development of Broadband Absorbing Materials for Frequencies as Low as 500 MC/S," *Naval Res. Lab., Washington, D. C., Memo. Rep. No. 300*; May, 1954.
- [81] L. E. Swarts, "Absorbing Screens for Microwave Radiation," *Antenna Lab., Dept. Eng., University of California, Berkeley, Calif., Rep. No. 165*; January, 1950.
- [82] A. J. Simmons and W. H. Emerson, "An anechoic chamber making use of a new broadband absorbing material," *1953 IRE NATIONAL CONVENTION RECORD*, pt. 2, pp. 34-41.
- [83] D. K. Cheng, "On the simulation of Fraunhofer radiation patterns in the Fresnel region," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 399-402; October, 1957.
- [84] R. H. T. Bates and J. Elliott, "The determination of the true side-lobe level of long broadside arrays from radiation pattern measurements made in the Fresnel region," *Proc. IEE*, monograph 169R, pt. C, vol. 103, pp. 307-312; September, 1956.
- [85] D. H. Shinn, "Mis-focusing and the near-field of microwave aerials," *Marconi Rev.*, vol. 19, pp. 141-149; 1956.
- [86] S. Matt and J. D. Kraus, "The effect of source distribution on antenna patterns," *Proc. IRE*, vol. 43, pp. 821-825; July, 1955.
- [87] G. Sinclair, E. C. Jordan, and E. W. Vaughan, "Measurement of aircraft-antenna patterns using models," *Proc. IRE*, vol. 35, pp. 1451-1462; December, 1947.
- [88] L. Peters, Jr., "Use of High Resistance Audio Cable for Pattern Measurements," *Res. Foundation, The Ohio State University, Columbus, Ohio, Eng. Rep. No. 487-4*; September, 1953.
- [89] L. Young and G. M. Ward, "Telephone remote control circuit for an antenna test site," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 385-386; October, 1958.
- [90] R. I. Primich, "Relative Merits of Various Types of Microwave Receivers," *Radiophys. Lab., Def. Res., Board, Ottawa, Can., Proj. Rep. No. 38-3-5*; February, 1956.
- [91] S. N. Van Voorhis, "Microwave Receivers," *M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y.*, vol. 23; 1948.
- [92] R. Kitai, "Coherent and incoherent detectors," *Electronic and Radio Eng.*, vol. 34, pp. 96-99; March, 1957.
- [93] H. V. Schurmer, "Noise performance of a three-stage microwave receiver," *Electronic and Radio Eng.*, vol. 35, pp. 271-274; July, 1958.
- [94] C. T. McCoy, "Present and future capabilities of microwave crystal receivers," *Proc. IRE*, vol. 46, pp. 61-66; January, 1958.
- [95] R. A. Smith, "The Relative Advantages of Coherent and Inherent Detectors: A Study of their Output Noise Spectra Under Various Conditions," *Telecommun. Res. Establ., Great Malvern, Eng., T.R.E. Tech. Note 95*; September, 1950.
- [96] E. G. Hamer, "V.H.F. aerial radiation pattern measurements," *Electronic Eng.*, vol. 25, pp. 427-431; October, 1953.
- [97] E. G. Hamer and J. B. L. Foot, "An automatic recorder of aerial radiation diagrams," *J. Brit. IRE*, vol. 14, pp. 33-42; January, 1954.
- [98] H. Le Caine and M. Katchky, "Microwave antenna beam evaluator," *Electronics*, vol. 20, pp. 116-120; August, 1947.
- [99] A. H. Beck and S. A. Tibbs, "An automatic polar diagram recorder," *Electronic Eng.*, vol. 20, pp. 17-19; January, 1948.
- [100] J. Bacon, "An improved design for audio-type exponential attenuators," *IRE TRANS. ON INSTRUMENTATION*, vol. I-6, pp. 23-26; March, 1957.

- [101] T. H. Chambers and I. H. Page, "The high-accuracy logarithmic receiver," *PROC. IRE*, vol. 42, pp. 1307-1314; August, 1954.
- [102] J. Dyson and B. A. C. Tucker, "An automatic contour plotter for the investigation of radiation patterns of microwave antennae," *J. IEE*, pt. IIIA, vol. 93, pp. 1403-1406; 1946.
- [103] T. T. Ling, "Aerial radiation patterns: apparatus for cathode-ray tube display," *Wireless Eng.*, vol. 30, pp. 192-195; August, 1953.
- [104] J. T. Bolljahn and R. F. Reese, "Electrically small antennas and the low-frequency aircraft antenna problem," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-1, pp. 46-54; October, 1953.
- [105] P. A. Kennedy and G. Kent, "Electrolytic tank, design and applications," *Rev. Sci. Instr.*, vol. 27, pp. 916-927; November, 1956.
- [106] V. Smith and C. M. Hatcher, "Ship models predict antenna patterns," *Electronics*, vol. 27, pp. 162-163; April, 1954.
- [107] A. W. Andrews, "Point Loma models help in developing shipboard systems," *Bur. of Ships J.*, vol. 3, pp. 2-6; November, 1954.
- [108] J. Y. Wong and J. C. Barnes, "Design and construction of a pattern range for testing high frequency shipborne antennas," *Trans. Eng. Inst. Can.*, vol. 2, pp. 15-21; January, 1958.
- [109] J. P. Shanklin, "Pattern measurements of large fixed antennas," *IRE TRANS. ON INSTRUMENTATION*, vol. I-4, pp. 16-22; October, 1955.
- [110] H. Brueckmann, "Helicopter measures antenna patterns," *Electronics*, vol. 28, pp. 134-136; November, 1955.
- [111] Radio Astronomy Issue, *PROC. IRE*, vol. 46; January, 1958.
- [112] G. A. Miller and W. Lavrench, *Natl. Res. Council of Canada*, Ottawa, Can., private communication.
- [113] T. Morita, "Determination of Phase Centers and Amplitude Characteristics of Radiating Structures," Stanford Res. Inst., Menlo Park, Calif., Tech. Rep. No. 1, SRI Proj. 898; March, 1955.
- [114] D. B. Anderson, "A microwave technique to reduce platform motion and scanning noise in airborne moving target radar," 1958 WESCON CONVENTION RECORD, pt. 1, pp. 202-211.
- [115] J. D. Kraus, "Antennas," McGraw-Hill Book Co., Inc., New York, N. Y.; 1950.
- [116] J. I. Bohnert, "Measurements on elliptically polarized antennas," *Proc. IRE*, vol. 39, pp. 549-552; May, 1951.
- [117] "IRE Standards on Antennas, Modulation Systems and Transmitters: Definitions of Terms"; 1948.
- [118] J. A. Saxton, "Determination of aerial gain from its polar diagram," *Wireless Eng.*, vol. 25, pp. 110-116; April, 1948.
- [119] J. S. Brown, "Determination of Antenna Gain from Measured Radiation Patterns," Elec. Eng. Dept., Syracuse University, Syracuse, N. Y., Final Rep. on Contract AF30(602)-300-EE-17; September 30, 1953.
- [120] C. E. Fisher, "A Computer for Use with Antenna Model Ranges," Stanford Res. Inst., Menlo Park, Calif., Tech. Rep. No. 33, SRI Proj. 591; 1953.
- [121] W. T. Slayton, "Design of microwave gain-standard horns," *Electronics*, vol. 28, pp. 150-154; July, 1955.
- [122] W. C. Jakes, Jr., "Gain of electromagnetic horns," *PROC. IRE*, vol. 39, pp. 160-162; February, 1951.
- [123] J. D. Lawson, "Some methods for determining the power gain of microwave aeriels," *J. IEE*, pt. III, vol. 95, pp. 205-209; July, 1948.
- [124] E. H. Braun, "Calculation of the gain of small horns," *PROC. IRE*, vol. 41, pp. 1785-1786; December, 1953.
- [125] E. H. Braun, "Some data for the design of electromagnetic horns," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-4, pp. 29-31; January, 1956.
- [126] E. M. Purcell, "A Method for Measuring the Absolute Gain of Microwave Antennas," Rad. Lab., M.I.T., Cambridge, Mass., Rep. No. 41-9; January, 1943.
- [127] J. J. Epis, "Corner reflector gain standards for 800-1600 mc/s.," Inst. Eng. Res., University of California, Berkeley, Calif., Ser. No. 60, Issue No. 118; July, 1954.
- [128] H. V. Cottony, "Methods for Accurate Measurement of Antenna Gain," Natl. Bur. of Standards, Boulder Labs., Boulder, Colo., Rep. No. 5539; November 18, 1957.
- [129] T. H. Crowley, "Measurement of Antenna Efficiency by Using Metals with Different Surface Resistivities," Res. Foundation, The Ohio State University, Columbus, Ohio, Rep. No. 478-21; November, 1953.

Submarine Communication Antenna Systems*

ROBERT W. TURNER†, MEMBER, IRE

Summary—Submarine communication antenna systems are divided into three systems: VLF, HF, and UHF. The evolution of each system from World War II to the present is described, and some of the problems associated with the antennas in each frequency band are reviewed. HF transmission line systems and their losses are discussed, and various environmental problems are considered. Specific examples of stacked and combined antennas are given.

INTRODUCTION

THE designer of submarine antennas is confronted not only with problems common to antennas in general but also with many problems peculiar to submarines. For example, the hydrodynamic forces involved place a severe limitation on the length of the HF antenna and since reliable long-range communication is essential, fairly high-powered transmitters are required.

High power and electrically short antennas, of course, mean high voltages. Hence, voltages of over 5 kv at end seals and hull fittings are not uncommon. Submarine antennas and system components are exposed to wide temperature variations and high hydrostatic pressures and are subject to high hydrodynamic shocks resulting from depth charges. They are also affected by salt deposits, exposure to sunlight, and icing conditions. In fact, it has recently been reported that submarines in transit beneath the polar ice pack determine whether there is ice above by observing through the periscope whether the HF antenna is bent.

During World War II, the submarine antenna system problem received sporadic and inadequate attention. It was felt that, with radio silence being the rule rather than the exception, the antenna system problem was not serious enough to warrant high priority effort. This opinion was not shared by the submariners, however.

* Original manuscript received by the IRE, January 19, 1959.

† U. S. Navy Underwater Sound Lab., Fort Trumbull, New London, Conn.

After futilely calling their advance base, they found it annoying, to say the least, to have the enemy answer them.

Largely through the efforts of the submariners, a submarine antenna system research and development program was initiated in 1945. The work, which was to be conducted on a continuing basis, was assigned to the U. S. Navy Underwater Sound Laboratory because of its unique facilities, scientific talent, and its proximity to the Atlantic Fleet submarine headquarters. This laboratory is under the direction of the Bureau of Ships.

Submarine communication antenna systems can be divided into VLF, HF, and UHF. These will be discussed in turn, with emphasis on their present status but with sufficient historical background to show the evolution of the systems. Various mechanical problems will then be discussed, and finally a brief picture of what is planned for the near future will be given.

VLF ANTENNA SYSTEM

Communicating from a submerged position has always been one of the submariner's major goals. Since sea water is a lossy medium, however, radio waves are rapidly attenuated. Fortunately, at the low end of the radio spectrum, the loss for moderate depths is not excessive. At 20 kc, for example, the attenuation is about 1.5 db per foot (see Fig. 1), but it becomes 15 db per foot at 2.0 mc. Consequently, if transmission through sea water is desired, even for short distances, very low frequencies must be used. The relatively low attenuation and infrequent blackouts at these frequencies provide reliable ground wave and sky wave propagation conditions.

A disadvantage, however, is the very low antenna efficiency at such long wavelengths, unless antennas of the order of 1000 feet are used with large tuning coils. Such antenna installations on submarines obviously are not feasible. As a result, VLF communication is limited to one-way transmission—from shore station to submarine—and is ideally suited to administrative traffic. Since the bandwidth of the transmitting station is very narrow, CW transmission is used rather than voice. It is interesting to note that since the refractive index is very high, the waves entering the water from the air travel nearly directly downward, regardless of the angle of arrival.

The early VLF submarine antenna consisted of a large, triangular, single-turn loop with the corners at bow, stern, and a point high on the periscope shears at the center. This antenna later was replaced by a 20-inch-diameter, multiturn loop mounted in the superstructure. Both of these antennas had the disadvantage of being bidirectional with nulls on fixed bearings. Soon after World War II, the Naval Research Laboratory developed an omnidirectional, crossed-loop antenna that was embedded in plastic in a streamlined configuration. This antenna, which is the one currently used, is mounted on a retractable mast.

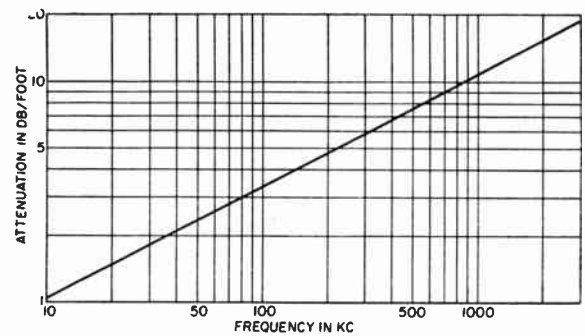


Fig. 1—Attenuation of plane wave in sea water (conductivity, $\sigma = 4$ mhos/m).

HF ANTENNA SYSTEM

By far the bulk of submarine communications, both short range and long range, takes place in the 2- to 30-mc band. Prior to World War II, HIF antennas were almost exclusively horizontal wires. For HF ground wave propagation, horizontally polarized waves over sea water experience very high attenuation. Accordingly, the horizontal wire merely acted as top loading for the vertical feeder. During the war years, the need to communicate from periscope depth became increasingly urgent. As a first step toward fulfilling this requirement, a PT boat whip antenna was attached to the periscope shears and was hinged so that it could be raised from or lowered to a horizontal position. It was spring-loaded and would erect to a vertical position when released. A deflector plate was attached to the antenna, causing the water to fold it back when the submarine dove. Unfortunately, the submarine had to surface so that a man could go out on deck and release the antenna. The submarine could then submerge to a depth that would still leave the feed-through insulator out of water. Although this was an improvement over the horizontal wires, it still did not provide communications from periscope depth. Late in the war, the first combination antenna was developed by the Naval Research Laboratory. It replaced one of the periscopes and radiated in a selected portion of the HF band and in the VHF band.

With the advent of the snorkel, communications from snorkel and periscope depths became virtually a "must." The horizontal wires were on their way out, and the retractable-mast whip antenna came into being. With this system (see Fig. 2), the submarine can remain submerged. This antenna, which is at present 25 feet long, folds back when the mast is lowered and thus reduces drag forces on it at high submerged speeds.

Most submarines need a second HF antenna. Since only one mast is normally available for HIF communications, a second whip is hinged, like the earlier whip, from the side of the upper portion of the submarine (see Fig. 3). There is one notable difference, however, in that the insulator is near the middle of the antenna. As a result, the insulator and the top of the whip, when in the erect position, are out of water when the submarine is at

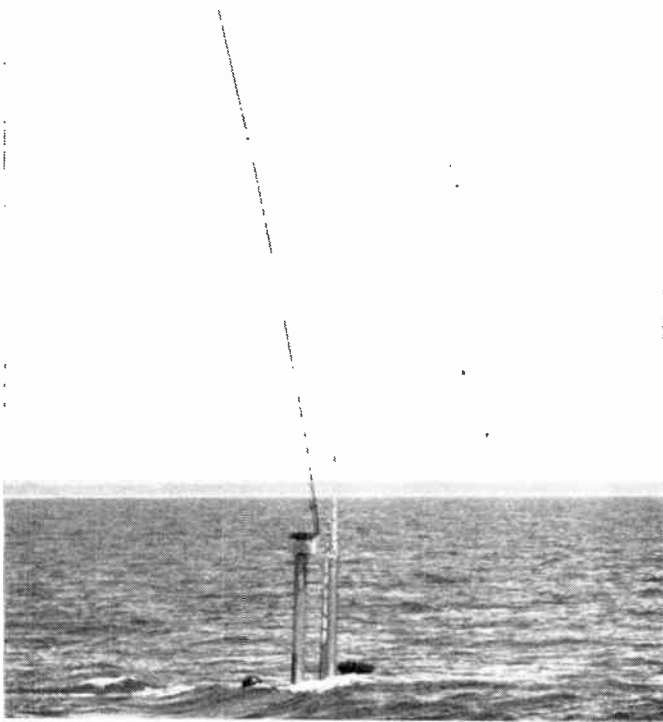


Fig. 2—Retractable mast antenna.

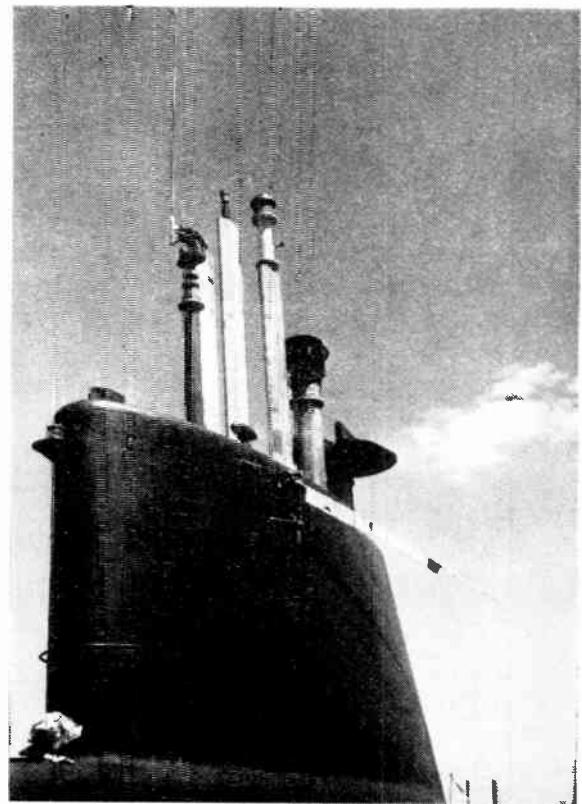


Fig. 3—Submarine superstructure showing retractable mast antenna and midfed antenna.

periscope depth. This antenna is generally referred to as the midfed antenna. The section of the whip below the insulator is rigid, and on it are mounted streamlined fairings that are free to rotate to conform to the water flow. These fairings greatly reduce the drag on the whip and the tendency to vibrate. The top portion of the whip is 12 feet long and is highly flexible. Consequently, if the submarine should dive with the antenna left in the vertical position, the top of the whip will bend over without breaking and thus will reduce the moment at the antenna base. The antenna is raised and lowered hydraulically.

In addition to the work which was done on the retractable and midfed antennas, development of telescoping antennas also was undertaken. However, the resulting antennas lacked the mechanical flexibility required at sea, and for that reason, further work on telescoping antennas was temporarily discontinued.

A radical change has taken place in the submarine's IIF transmission system since 1945. At that time the principal portion of the transmission line was a copper-lined rectangular trunk, 12 to 16 inches on a side and a few feet long. Several center conductors ran through the same trunk to the various antennas. The trunk was pressurized by bowl insulators at the outboard end and by a small hatch on the inboard end. This hatch had to be closed when the submarine submerged since it was a secondary barrier that protected the submarine in case a bowl insulator was damaged. The center conductors were connected to flexible leads that were disconnected when the hatch was closed. Since the center conductors were unshielded from the transmitter to the trunk, the

radio room was hot. Redesign of this somewhat crude transmission line system was assigned a high priority, and the system soon evolved into a completely coaxial line which extended from transmitter to antenna and utilized standard polyethylene cables. Special pressure-proof hull fittings and end seals were developed to keep water out of the transmission line and the submarine.

At the low end of the IIF band, submarine antennas are electrically quite short and consequently present a poor match to the transmission line. This mismatch causes a large loss within the transmission line, and the change from the large trunk and open-wire system to the coaxial cable increased this loss. However, the over-all system loss, including the transmitter output and coupling circuit, was not appreciably increased. This was due to the fact that the higher loss transmission line presents a better match to the transmitter. Thus, the small increase in over-all loss was a small price to pay for the advantage of the completely coaxial system.

Curves of system loss vs frequency for the 25-foot retractable whip and the 12-foot section of the midfed whip are shown in Fig. 4. The most significant step that has been taken to reduce system loss is the incorporation of an automatic tuner at the base of the antenna. In the case of the retractable mast, the tuner is contained in a cylindrical container which fits inside the mast. This tuner is controlled from the transmitter and is automatically adjusted to match the antenna. A curve of system loss vs frequency for the system with a tuner is also shown in Fig. 4. It is seen from these curves that

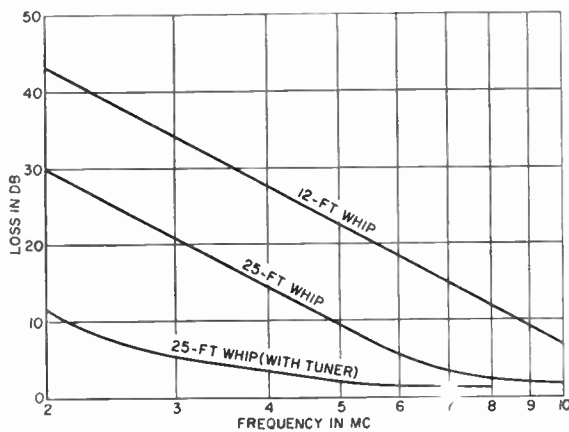


Fig. 4—System loss for 12-foot and 25-foot whip antennas with 50 feet of RG-17/U cable.

the loss is considerably reduced if a high rather than a low frequency is used. A compromise must be made, however, because of the higher ground wave attenuation at the higher frequencies.

Another obvious method of reducing system losses is that of lengthening the antenna. This approach will be discussed later in the description of a combined antenna.

Ground (sea water) and insulator losses can be significant at the low end of the HF band, especially with a tuner. When no tuner is used, these losses cause the antenna to present a better match to the transmission line, so that they are nearly balanced by a decrease in transmission-line loss. The ground loss can be reduced by using a larger-diameter antenna base (below the insulator), and the insulator loss can be kept small by using a low-loss dielectric with high water repellency and also by cleaning the insulators at frequent intervals.

UHF ANTENNA SYSTEM

The requirement for submarines to communicate at frequencies above 30 mc developed early in World War II when it became necessary to communicate over short ranges with aircraft and ships in task force operations. Initially frequencies from 100 to 156 mc were used, but soon after the war a shift was made to the 225- to 390-mc band. An around-the-mast type of antenna was designed, first for VHF and later for UHF. With this design several antennas can be stacked vertically as the mast need not be broken for an insulator. Isolation of the radiating elements from the mast is accomplished by quarter-wave sleeves, and the feed point is similarly isolated. Fig. 5 illustrates the principle used, and Fig. 6 shows an example of stacked antennas. When the antennas are stacked in this manner, the mutual interaction between antennas is significantly reduced and fewer masts are required.

ENVIRONMENTAL PROBLEMS

Submarine antenna systems are subjected to a variety of unique environmental conditions. Consider, for example, that every foot of depth produces an additional 0.4 psi of hydrostatic pressure. Thus, the pressure ex-

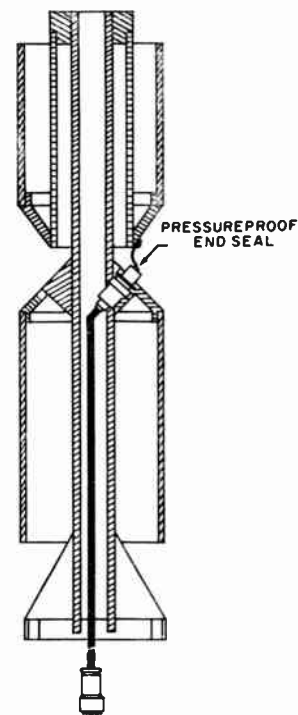


Fig. 5—Around-the-mast type antenna.



Fig. 6—Typical stacked antenna assembly employing around-the-mast type antennas.

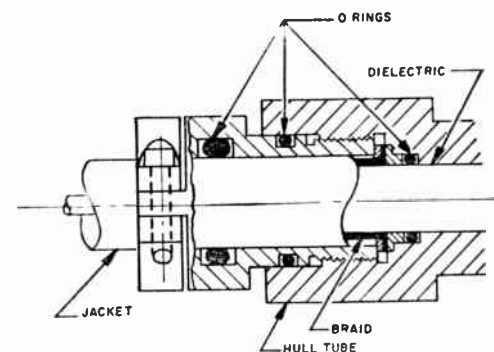


Fig. 7—Section of hull fitting illustrating use of O rings.

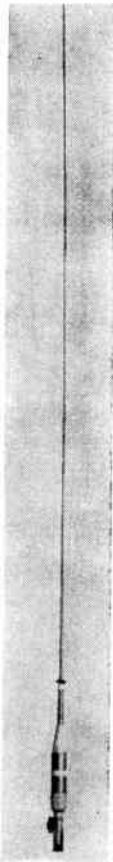


Fig. 8—Combined HF-UHF-IFF antenna assembly.

erted on components at several hundred feet beneath the surface is considerable. In addition, a design must take into consideration possible depth charge attacks. The items most affected by these pressures are end seals, connectors, hull fittings, and transmission lines. The most reliable method of solving the pressure problem is the O ring (see Fig. 7).

The forces resulting from drag and wave action must also be reckoned with. Wave action is a factor only when the submarine is on or near the surface, but it is serious enough to require that a large safety factor be built into antennas. Drag forces are becoming of increased importance as the underwater speed of submarines is raised. One way to solve this problem is to house completely all antennas when the submarine is submerged. This has not always been possible, however, especially with long whip antennas. Streamlining of all retractable antenna masts is becoming standard.

Corrosive and galvanic action are ever present and must be minimized. Since a sufficient choice of materials and protective coatings is usually available, these problems are not serious.

NEW ANTENNA SYSTEM DEVELOPMENTS

A combined HF-UHF-IFF antenna assembly that has recently been developed is illustrated in Figs. 8 and 9. The basic design objectives that led to the development of this antenna assembly were 1) to obtain a longer HF antenna, and 2) to eliminate a retractable mast

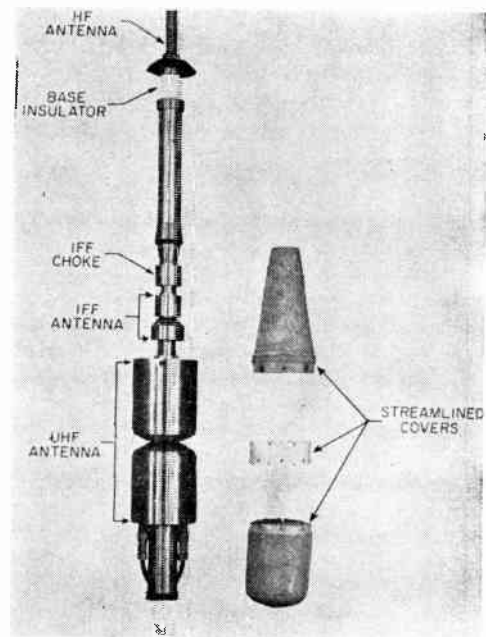


Fig. 9—Enlargement of UHF-IFF portion of the combined HF-UHF-IFF antenna assembly.

by combining the HF antenna with the UHF-IFF antenna assembly. The HF antenna is 30 feet long, but with the two antennas below it an effective length of 36 feet is obtained. The UHF and IFF antennas are of the around-the-mast type. Considerable difficulty was experienced in the development of the IFF antenna. On one hand, it was desirable to keep the diameter of the antenna small in terms of wavelengths, while on the other hand, the diameter of the mast running through it was determined by the hydrodynamic forces on the HF antenna. A satisfactory solution was obtained by using a very high strength steel for the mast and by accepting a moderate distortion in the horizontal pattern of the IFF antenna.

Telescopic antennas are again under development by Government activities and contractors but are only in the experimental stages. It is confidently expected, however, that the submarine-mounted communication antennas of the future will be completely housed when the submarine is deeply submerged. In addition, other antenna systems are being developed and new concepts are being investigated in the over-all effort to keep pace with the continued evolution of the submarine. Unfortunately, further discussion of this work is precluded because of security restrictions.

ACKNOWLEDGMENT

The author wishes to thank C. M. Dunn, Head of the Electromagnetics Development Division, for his suggestions and information on the World War II submarine antenna systems. Also, thanks are due the members of the Submarine Antenna Branch and the Submarine Communications Branch for their part in the developments with which this article is concerned.

Models of the Atmospheric Radio Refractive Index*

B. R. BEAN†, ASSOCIATE MEMBER, IRE, AND G. D. THAYER†

Summary—Evaluation of atmospheric refraction effects on UHF-VHF radio propagation has long been accomplished with the convenient four-thirds earth concept of Schelling, Burrows, and Ferrell. This method has proven particularly useful in evaluating performance of point-to-point radio communications systems. However, relatively new long-range applications have demanded a model of atmospheric radio refractive index more representative of observed refractive index profiles than the simple linear decay inherent in the four-thirds earth approach.

This paper introduces two models of atmospheric radio refractive index which can be used to predict refraction effects from the value of the refractive index at the transmitting point. Both models offer considerable improvement over the four-thirds earth model, particularly for applications at long distances and high elevations in the atmosphere. Further, both models may be adjusted to represent mean conditions at different times of year and in different geographical locations. A new method of predicting radio-ray refraction at very low initial elevation angles is introduced which utilizes both the initial value and the initial height-gradient of the refractive index over roughly the first 100 meters above the earth's surface. This method, which is dependent only upon the first two radiosonde reporting levels or simple tower measurements of the common meteorological elements, results in a considerable improvement of the values of ray-refraction predicted by the model.

FOREWORD

THIS paper is concerned with model or standard atmospheres. Although one does not expect the real atmosphere actually ever to correspond to the model atmosphere, such atmospheres are usually designed to be representative of, say, average summer or winter conditions. This is not to say that the model atmosphere will represent any particular summer or winter but rather represents the average of many summers or winters. It is in this sense that the present study is directed towards a model of the atmosphere that yields average radio-ray refraction for any location for which surface weather observations have been compiled for a long time, whether these compilations be for a particular month, season or region. In practice, when it is necessary to estimate future or past refraction effects for short periods of time, such as a few days, the models offered in the present study may not be strictly applicable but will be of more use than previous models since they are more representative of the refractive index structure of the atmosphere.

INTRODUCTION

Refraction is one of the more important effects of the earth's atmosphere on radio propagation, particularly at frequencies of 30 mc and above. Perhaps the simplest treatment of radio refraction, and certainly the most convenient for calculation, is the classical effective

earth's radius model advanced by Schelling, Burrows and Ferrell.¹ This concept, inherently a linear model of refractivity, has become known familiarly as "4/3 earth" refraction and has the advantage of specifying a mean of refraction effects that is particularly useful for calculation of field strengths in point-to-point radio communications systems.

In 1953, it was shown that the effective earth's radius factors determined from actual meteorological observations varied systematically as a function of climate.² Later it was shown that annual variations in the monthly means of observed refractive index gradient (and thus the effective earth's radius factor) were highly correlated with the corresponding monthly medians of observed radio field strengths for many different paths in the United States.³ This, in turn, led to the development of a prediction method for determining radio field strengths as a function of a variable effective earth's radius factor.⁴

Although it has been known for some time that the effective earth's radius approach can be misleading for applications involving large elevation differences,⁵ it was not until the advent of high precision, long-range radar techniques that the need for rectification of these errors, particularly range and elevation angle errors, became pressing.⁶⁻¹⁰ As a result of this need for a more accurate representation of actual atmospheric conditions, a reference atmosphere based on a simple exponential decay of refractive index with height was recommended for adoption as an international standard by the International Consultative Committee on Radio (CCIR) meeting at Geneva, Switzerland, in the summer of 1958.

¹ J. C. Schelling, C. R. Burrows, and E. B. Ferrell, "Ultra-short-wave propagation," *Proc. IRE*, vol. 21, pp. 427-463; March, 1933.

² B. R. Bean, "The geographical and height distribution of the gradient of refractive index," *Proc. IRE*, vol. 41, pp. 549-550; April, 1953.

³ B. R. Bean and F. M. Meaney, "Some applications of the monthly median refractivity gradient in tropospheric propagation," *Proc. IRE*, vol. 43, pp. 1419-1431; October, 1955.

⁴ K. A. Norton, P. L. Rice, and L. E. Vogler, "Use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain," *Proc. IRE*, vol. 43, pp. 1488-1526; October, 1955.

⁵ W. Miller, "Effective earth's radius for radio wave propagation beyond the horizon," *J. Atmos. Phys.*, vol. 22, pp. 55-62; January, 1951.

⁶ M. Schulkin, "Average radio-ray refraction in the lower atmosphere," *Proc. IRE*, vol. 40, pp. 554-561; May, 1952.

⁷ B. M. Fannin and K. H. Jehn, "A study of radar elevation angle error due to atmospheric refraction," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 71-77; January, 1957.

⁸ B. R. Bean and B. A. Cahoon, "The use of surface weather observations to predict the total atmospheric bending of radio waves at small elevation angles," *Proc. IRE*, vol. 45, pp. 1545-1546; November, 1957.

⁹ L. J. Anderson, "Tropospheric bending of radio waves," *Trans. AGU*, vol. 39, pp. 208-212; April, 1958.

¹⁰ G. H. Millman, "Atmospheric effects on VHF and UHF propagation," *Proc. IRE*, vol. 46, pp. 1492-1501; August, 1958.

* Original manuscript received by the IRE, December 9, 1958; revised manuscript received February 16, 1959.

† Natl. Bur. Standards, Boulder, Colo.

It is the purpose of this paper to present an examination of the problem of radio refraction, with special attention being given both to the background of material which led to the choice of various models of refractive index structure and to the improvement which these models offer over the standard effective earth's radius approach.

In addition, a method of predicting the effective earth's radius factor from surface weather observations that will improve the effectiveness of that approach will be suggested.

As criteria for judging the usefulness of the various refractive index models presented, two comparisons will be employed: 1) the relative agreement between the model and the average refractive index structure of the atmosphere, and 2) the relative agreement between the amount of refraction (ray bending), predicted by the model and the amount actually calculated for observed refractive index profiles representing both mean and individual observations of extreme conditions from many geographically and climatically diverse locations.

These studies of atmospheric ray bending have clearly shown the extreme sensitivity of refraction effects to the initial gradient of the refractive index at very small initial elevation angles. This realization has led to the development of a new method of predicting the bending of radio rays in terms of both the initial refractive index gradient and the value of the refractive index at the earth's surface. This new method of prediction is several orders of magnitude more accurate than the various model atmospheres alone.

BACKGROUND

The classical method of accounting for the effects of atmospheric refraction when considering radio propagation problems is to assume an effective earth's radius, $a_e = ka$ where k is the effective earth's radius factor and a is the true radius of the earth. This method, advanced by Schelling, Burrows and Ferrell,¹ assumes an earth suitably larger than the actual earth and allows that radio rays be drawn as straight lines over this earth rather than as curved rays over an earth with the true radius a . The value of k usually assumed in the United States is 4/3. This method of accounting for atmospheric refraction permits a tremendous simplification in the computation of radio field strengths although the distribution of refractive index implied by this method is not a very realistic representation of the average refractive index structure of the atmosphere. In the sections that follow we shall consider the refractive index structure implied by the effective earth's radius model and how this differs from the observed refractive index structure of the atmosphere. Further, we will explore the limits of applicability of the effective earth's radius approach and will suggest a model of atmospheric refractive index that will be a more realistic model under those conditions where the effective earth's radius is most in error.

The effective earth's radius factor is given for radio rays leaving the transmitter at low angles to the horizon by the expression:

$$k = \frac{1}{1 + \frac{a'}{n} \frac{dn}{dh}}, \quad (1)$$

where n is the refractive index and is of the order 1.0003 and dn/dh is the vertical gradient of the refractive index. The sea level radius of the earth, a , is not used in (1) but rather the local radius, a' , which includes the altitude of the local terrain above sea level. It is usually assumed for mathematical simplicity that $a' \approx 3960$ miles. By noting that dn/dh is normally negative and of the order of $-40 \times 10^{-6}/\text{km}$ and assuming that this gradient is constant with height, one may obtain the familiar value of 4/3 for k . Note, however, that the assumption that dn/dh is constant with height implies that n decreases linearly with height. This assumption is patently in disagreement with the observed refractive index structure of the atmosphere except at low heights, say less than 1 km, as may be seen on Fig. 1. For convenience in Fig. 1 and throughout the remainder of the paper, the refractive index is given in terms of its scaled-up value, N , commonly called the refractivity:

$$N = (n - 1)10^6. \quad (2)$$

N may be determined from standard weather observations by use of the Smith-Weintraub determination,¹¹

$$N = 77.6 \frac{P}{T} + 3.73 \times 10^5 \frac{e}{T^2}, \quad (3)$$

where P is the total atmospheric pressure in millibars, e is the water-vapor pressure in millibars and T is the absolute temperature in degrees Kelvin. When utilizing data from radiosonde ascents, the water-vapor pressure is calculated as the saturation-vapor pressure at temperature T , multiplied by the relative humidity (expressed as a decimal).

The data given on Fig. 1 represent the average of individual radiosonde observations over a 5-year period at several locations chosen to represent the extremes of refractive index profile conditions within the United States. The Miami, Fla., profile is typical of warm, humid sea-level stations that tend to have maximum refraction effects while the Portland, Me., profile is associated with nearly minimum sea-level refraction conditions. Although Ely, Nev., has a much smaller surface N value than either Miami or Portland, it is significant that when its N profile is plotted in terms of altitude above sea level, it falls within the limits of the maximum and minimum sea level profiles. It is to take advantage of this simplification that altitude above sea level rather than height above ground is used through-

¹¹ E. K. Smith and S. Weintraub, "The constants in the equation for atmospheric refractive index at radio frequencies," *Proc. IRE*, vol. 41, pp. 1035-1037; August, 1953.

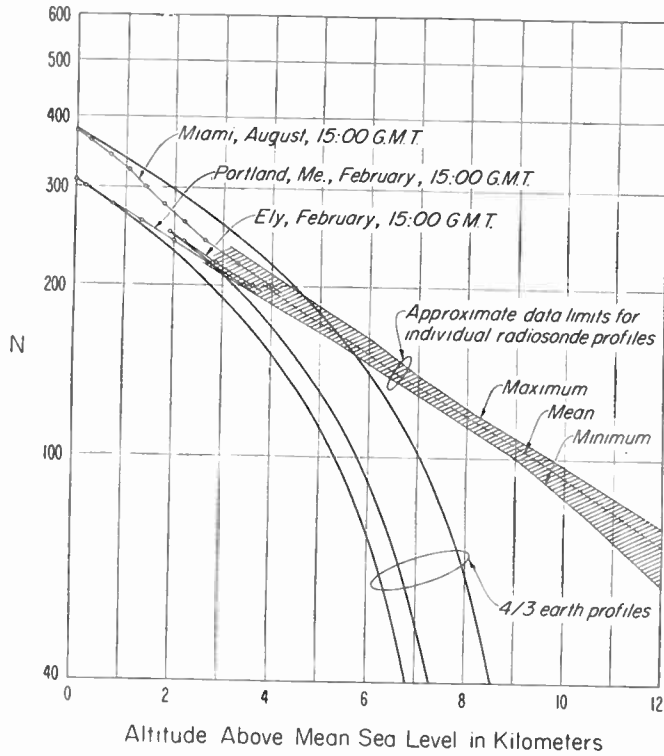


Fig. 1—Typical N vs height distributions.

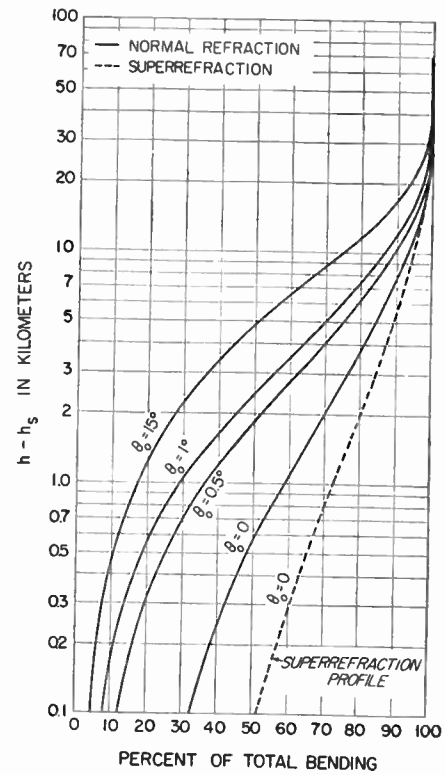


Fig. 2—Percentage of total bending as a function of height in the earth's atmosphere.

out this study. The N distribution for the $4/3$ earth atmosphere is also shown on Fig. 1. It is quite evident that the $4/3$ earth distribution has about the correct slope in the first kilometer above the earth's surface but decreases much too rapidly above that height. It is also seen that the observed refractivity distribution is more nearly an exponential function of height than a linear function of height as assumed by the $4/3$ earth atmosphere. One might expect the refractivity to decrease exponentially with height since the first term of (3), involving P/T , comprises at least 70 per cent of the total and is proportional to air density, a well-known exponential function of height.¹²

One might wonder, in the light of the data of Fig. 1, why the effective earth's radius approach has served so well for so many years. It appears that this success is due to the $4/3$ earth model being in essential agreement with the average N structure near the earth's surface. The importance of this agreement can be seen in Fig. 2 where the percentage of the total bending of a radio ray is given in terms of the height above the earth's surface, $h - h_s$, to which the radio ray has traveled and where h is altitude above sea level and h_s is the altitude of the station. The total bending is defined as the angle through which the radio ray turns as it passes from the earth's surface completely through the atmosphere. The bending, designated by the angle τ is shown on Fig. 3. Returning to Fig. 2, we note that fully 30 per cent of the total bending occurs in the first 100 meters of the atmos-

phere for a ray leaving the earth tangentially, *i.e.*, the initial elevation angle θ_0 is equal to zero. Similarly, 60 per cent of the total bending occurs in the first kilometer above the earth's surface. It is seen then that the disagreement between the N structure of the effective earth's radius approach and the actual atmosphere is minimized by their agreement in the lower layers of the atmosphere where the greater percentage of the bending occurs. As will be seen, it is only at very high altitudes and long distances that the disagreement in N structure produces marked error in estimation of atmospheric refraction of radio waves.

Fig. 2 also offers a possible explanation of the observation that the total bending may be well represented by a linear function of the surface value of the refractivity⁸ N_s for values of θ_0 as small as 10 milliradians. One might surmise that the reason for this correlation of total bending and N_s is that, on the average, N_s itself is well correlated with N gradient. That this is indeed the case in some instances may be seen from Fig. 4 where 6- to 8-year averages of daily observations of N_s are compared with similar values of ΔN , the difference between N_s and the value of N at 1 km above the earth's surface:

$$-\Delta N = N_s - N(1 \text{ km}). \quad (4)$$

It is evident that, for average conditions, a relationship exists between ΔN and N_s . One such relationship is the least squares determination:

$$-\overline{\Delta N} = 7.32 \exp \{0.005577 \overline{N_s}\}, \quad (5)$$

¹² B. Haurwitz, "Dynamic Meteorology," McGraw-Hill Book Co., New York, N. Y., pp. 11-12; 1941.

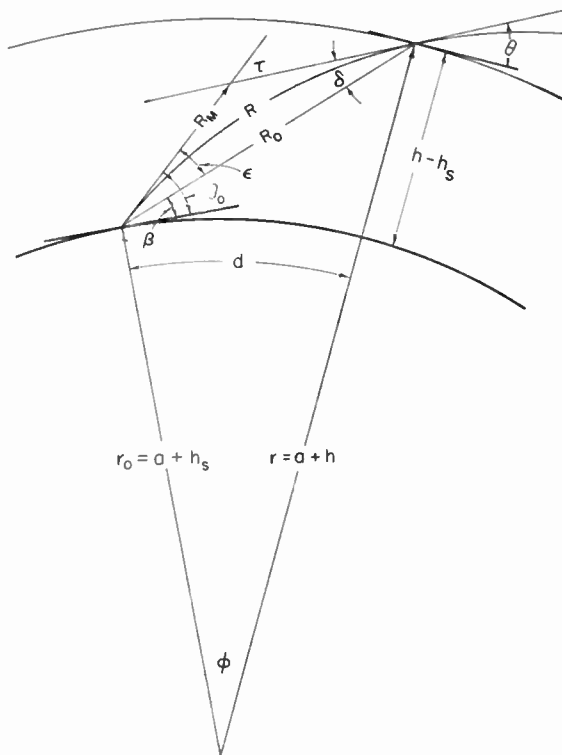


Fig. 3—Geometry of the refraction of radio waves. Note that the sea level radius of the earth is taken as $a = 3960$ miles.

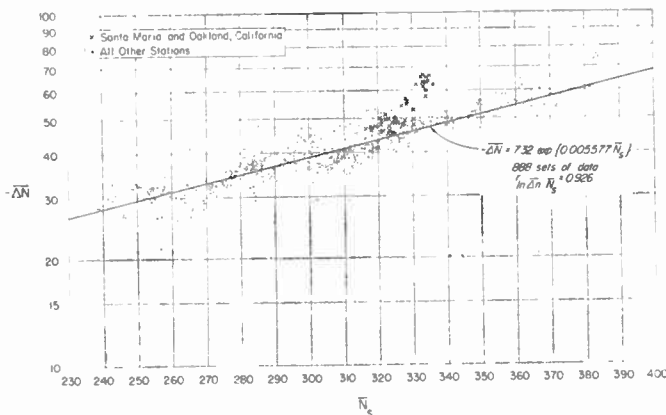


Fig. 4—Regression of $\ln \overline{\Delta N}$ upon \overline{N}_s for the hours 0300 and 1500 GMT.

based upon 888 sets of monthly mean values of $\overline{\Delta N}$ and \overline{N}_s from 45 United States weather stations representing many diverse climates. This relationship between $\overline{\Delta N}$ and \overline{N}_s is expected to represent the best estimate of a majority of individual profiles and certainly will closely agree with average conditions for the United States with one notable exception, Southern California in the summer. These data, although shown on Fig. 4, were excluded from the least squares determination due to their singular large range of ΔN compared to their small range of N_s which resulted in a marked "finger" of points rising from the main body of the data. This obvious departure of data points (24 points out of a total of 912) plus the well-known unique nature of the Southern

California summer climate were taken as sufficient justification for ignoring these points, although, as shall be seen, the ray bendings based upon the ΔN obtained from (5) are in rather good agreement with the values calculated from actual N profiles, even those observed in Southern California.

Eq. (5) offers a convenient method of specifying various models of the refractivity structure of the atmosphere, since it allows an estimation of the value of N at 1 km in addition to the two values already known, *i.e.*, N_s and $N=0$ at $h = \infty$.

Several models of the refractivity decay with height will be developed in the following sections. As a critical test of these models, the ray bending calculated from their N structure will be compared with similarly obtained values from five-year average N profiles.

AVERAGE N STRUCTURE OF THE ATMOSPHERE

It would seem that the deficiency of the effective earth's radius approach could be lessened by modifying that theory in the light of the average N structure of the atmosphere. An indication of the average N structure was obtained by examining a variety of N profiles which were carefully selected from 39 station-years of individual radiosonde observations to represent the range of N profile conditions during summer and winter at 13 climatically diverse locations.⁸ The results of this examination are given in Table I.

TABLE I
REFRACTIVITY STATISTICS AS A FUNCTION OF ALTITUDE ABOVE SEA LEVEL AS DERIVED FROM INDIVIDUAL RADIOSONDE OBSERVATIONS

Altitude in kilometers	\overline{N}	Maximum N	Minimum N	Range*
4	197.1	209.5	186.5	23.0
5	172.3	184.0	165.0	19.0
6	151.4	161.0	146.0	15.0
7	134.0	139.5	129.5	10.0
8	118.4	121.5	113.3	8.2
9	104.8	108.0	100.0	8.0
10	92.4	97.0	86.0	11.0
11	81.2	86.0	70.0	16.0
12	70.7	76.0	60.5	15.5
14	53.2	60.0	44.5	15.5

* Range = Maximum N - Minimum N .

It is interesting to note that the range of N values has a minimum at 8 to 9 km above sea level but is systematically greater above and below that altitude. The average value of 104.8 at 9 km corresponds to a similar value reported by Stickland¹³ as typical of the United Kingdom. Further, the altitude of 8 km corresponds to the altitude reported by Humphreys¹⁴ where the atmospheric density is nearly constant regardless of season

¹³ A. C. Stickland, "Refraction in the lower atmosphere and its application to the propagation of radio waves," in "Meteorological Factors in Radio-Wave Propagation," The Physical Society, London, Eng., pp. 253-267; 1946.

¹⁴ W. J. Humphreys, "Physics of the Air," McGraw-Hill Book Co., Inc., New York, N. Y., p. 82; 1940.

or geographical location. Since the first term in the expression for refractivity is proportional to air density, and the water-vapor term is negligible at an altitude of 9 km, the refractivity also tends to be constant at this altitude. It seems quite reasonable, then, to adopt a constant value of $N=105$ for 9 km, thus further facilitating the specification of model atmospheres. Further, when the values of Table I are plotted on semilogarithmic paper, such as on Fig. 1, it is seen that the data strongly suggest that N may be represented by an exponential function of height of the form:

$$N(h) = N_0 \exp \{-bh\},$$

in the altitude range of 1 to 9 km above sea level.

The preceding background discussion has presented the material necessary for the main purpose of this paper: the consideration of the suitability of various models of refractivity to describe atmospheric refraction of radio waves. As a guide to what follows, let us ask what a logical sequence of models (or assumptions) would be to describe the effects of atmospheric refraction. One such sequence might be:

1) Assume an invariant model that is near to the actual average conditions and facilitates the calculation of radio field strengths. This has been done by the 4/3 earth model.

2) Assume a variable effective earth's radius factor for the calculation of radio field strengths in various climatic regions. This approach has been followed by Norton, Rice, and Vogler.⁴

When it has become apparent that the effective earth's radius approach is inadequate one might proceed by:

3) Correcting the effective earth's radius model by assuming a more realistic N structure in the region where that model is most in error. This "modified effective earth's radius" model would then maintain, for some applications, the advantages of the original model.

4) Assume an entirely new model of N structure guided by the average N structure of the atmosphere.

It is assumed that models 3) and 4) would allow for seasonal and climatic changes of the average refractive index structure of the atmosphere.

In the sections that follow, models 3) and 4) will be developed and tested by their relative agreement with the ray bendings obtained from actual long-term average N profiles.

DEFINITION OF MODEL ATMOSPHERES

The first model of atmospheric refractivity that will be considered is based upon the effective earth's radius concept in the first kilometer. In this atmosphere N is assumed to decay linearly with height from the surface h_s to 1 km above the surface h_s+1 . This linear decay is given by

$$N(h) = N_s + (h - h_s)\Delta N, \quad h_s \leq h \leq h_s + 1, \quad (6)$$

where

$$-\Delta N = 7.32 \exp(0.005577N_s). \quad (5)$$

Note that throughout this paper the letter h designates the altitude above mean sea level.

It may be further assumed that N decreases exponentially from h_s+1 to a constant value of 105 at 9 km above sea level. In this altitude range N is defined by:

$$N = N_1 \exp \{-c(h - h_s - 1)\}, \quad h_s + 1 \leq h \leq 9 \text{ km}, \quad (7)$$

where

$$c = \frac{1}{8 - h_s} \ln \frac{N_1}{105},$$

and N_1 is the value of N at 1 km above the surface.

Above the altitude of 9 km, where less than 10 per cent of the total bending occurs even for rays of $\theta_0=0$, a single exponential decrease of N may be assumed. The coefficients in the exponential expression:

$$N = 105 \exp \{-0.1424(h - 9)\}, \quad h \geq 9 \text{ km}, \quad (8)$$

were determined by a least squares analysis of the Rocket Panel data.¹⁵ This expression is also in agreement with the ARDC Model Atmosphere-1956¹⁶ and Dubin's¹⁷ conclusion that a standard density-distribution may be used to determine the refractivity distribution at altitudes in excess of 20,000 feet.

The three-part model of the atmosphere expressed by (6)–(8) has the advantage of the effective earth's radius approach, particularly for such applications as point-to-point radio relaying over distances up to, say, 100 miles, where the radio energy is generally confined to the first kilometer, plus being in reasonably good agreement with the average N structure of the atmosphere. The reader is cautioned, however, that application of this model to mode-theory calculations would be misleading, since the resultant diffraction region fields would be enhanced by the addition of strong reflections from the n -gradient discontinuities at h_s+1 and at 9 km. The specific combinations of N_s , h_s and ΔN , that define the CRPL Reference Atmosphere-1958 are given in Table II.

The station elevations corresponding to given combinations of N_s and ΔN were chosen to correspond with an average decay of N with station elevation. Although the error in neglecting this height dependence has been estimated to be no more than a few per cent, it could be important in such high precision applications as radar tracking of earth satellites. It should be remembered in subsequent applications that a unique feature of these reference atmospheres is the dependence of N_s on the altitude of the surface above sea level. This feature was built in so that the reference atmospheres would be completely specified by the single parameter N_s .

¹⁵ The Rocket Panel, "Pressures, densities and temperatures in the upper atmosphere," *Phys. Rev.*, vol. 88, pp. 1027–1032; December 1, 1952.

¹⁶ R. A. Minzner and W. S. Ripley, "The ARDC Model Atmosphere, 1956," *Air Force Surveys in Geophysics*, No. 86; December, 1956.

¹⁷ M. Dubin, "Index of refraction above 20,000 feet," *J. Geophys. Res.*, vol. 59, pp. 339–344; September, 1954.

The second model of the atmosphere to be considered may be specified by assuming a single exponential distribution of N :

$$N = N_s \exp \{ -c_e(h - h_s) \}, \quad (9)$$

where

$$c_e = \ln \frac{N_s}{N(1 \text{ km})} = \ln \frac{N_s}{N_s + \Delta N} \quad (10)$$

and these equations are used to determine N at all heights. This model of atmospheric refractivity is a close representation of the average refractivity structure within the first 3 km. Further, the single exponential model has the advantage of being a continuous function, and therefore may be used for theoretical studies. This model of the atmosphere has been adopted for use within the National Bureau of Standards with the specific values of the constants in (9) and (10). These constants are given in Table III and specify the CRPL Exponential Reference Atmosphere-1958.

TABLE II
CONSTANTS FOR THE CRPL REFERENCE ATMOSPHERE—1958

N_s	h_s feet	a' miles	$-\Delta N$	k	a_e miles	c per km
0	0	3960.0000	0	1.00000	3960.00	0
200	10,000	3961.8939	22.3318	1.16599	4619.53	0.106211
250	5,000	3960.9470	29.5124	1.23165	4878.50	0.114559
301	1,000	3960.1894	39.2320	1.33327	5280.00	0.118710
313	700	3960.1324	41.9388	1.36479	5403.88	0.121796
350	0	3960.0000	51.5530	1.48905	5896.66	0.130579
400	0	3960.0000	68.1295	1.76684	6996.67	0.143848
450	0	3960.0000	90.0406	2.34506	9286.44	0.154004

Note: a_e is the effective earth's radius and is equal to $a'k$, $a' = a + h_s$, where h_s is the altitude of the earth's surface above sea level.

$a = 3960$ miles.

$$c = \frac{1}{8 - h_s} \ln \frac{N_1}{105}$$

TABLE III

TABLE OF THE CONSTANT c_e FOR THE CRPL EXPONENTIAL RADIO REFRACTIVITY ATMOSPHERES*

$$N = N_s \exp \{ -c_e(h - h_s) \}$$

ΔN	N_s	c_e per km
0	0	0
22.3318	200.0	0.118400
29.5124	250.0	0.125625
30.0000	252.9	0.126255
39.2320	301.0	0.139632
41.9388	313.0	0.143859
50.0000	344.5	0.156805
51.5530	350.0	0.159336
60.0000	377.2	0.173233
68.1295	400.0	0.186720
70.0000	404.9	0.189829
90.0406	450.0	0.223256

* The International Radio Consultative Committee of the International Telecommunication Union meeting in Geneva during July, 1958, recommended the adoption of the basic reference atmosphere:

$$N(h) = 300 \exp \{ -0.139h \}$$

where h is the height above the ground in kilometers.

Fig. 5 compares the N structure of the above two models plus the 4/3 earth model. It can be seen that the 4/3 earth assumption agrees with the reference atmosphere in the first kilometer, which is to be expected since $N_s = 301$ is the value required to yield the 4/3 gradient from (5). Fig. 5 illustrates the essential agreement of the reference atmosphere with the Rocket Panel and RDC data. The exponential reference atmosphere is also shown on Fig. 5 for $N_s = 313$, the average value for the United States. The exponential reference atmosphere appears to be a reasonable single line representation of N throughout the height interval shown. The differences between the various models become more apparent by examining their agreement with observed N profiles over the first 10 km as in Fig. 6.

The reference and exponential reference atmospheres are given for the N profiles corresponding to near-maximum N_s (Lake Charles, La.) and near-minimum-at-sea-level N_s conditions (Caribou, Me.). The two reference atmospheres were determined solely from the N_s values of each profile. Several observations can be made of these data. First, the 4/3 earth model closely represents the slope of the minimal N_s profile over the first kilometer, but then decreases too rapidly with height. Note, however, that the 4/3 earth model with its constant decay of 39.2 N units per kilometer would be a very poor representation of the maximum profile which decreases over 66 N units in the first kilometer. The exponential reference atmosphere is in good agreement with the initial N distribution but tends to give values systematically low above approximately 3 km. At first glance, the exponential reference atmosphere does not appear to be as good a representation of the two observed profiles as the reference atmosphere, particularly above approximately 5 km. Subsequent analysis of the refraction obtained from the two model atmospheres will show that this systematic disagreement of the exponential reference atmosphere in the 5- to 20-km interval is a minor defect of the model compared to its closer agreement with observed N distributions over the first 1 to 3 km. This is particularly true for the higher values of N_s such as that for Lake Charles.

The above models are more in agreement with long-term mean N profiles than is the 4/3 earth model. The application at hand would aid in deciding which of the reference atmospheres would be most useful. To aid in distinguishing between the various models, the following sections will be concerned with comparing the ray bending between the models. To do this, it will be necessary first to develop the pertinent equations of ray theory.

CALCULATION OF ATMOSPHERIC BENDING OF RADIO RAYS

If it is assumed that the refractive index is spherically stratified with respect to the earth, then Snell's Law may be expressed in cylindrical coordinates as:

$$nr \cos \theta = n_0 r_0 \cos \theta_0 = \text{constant}$$

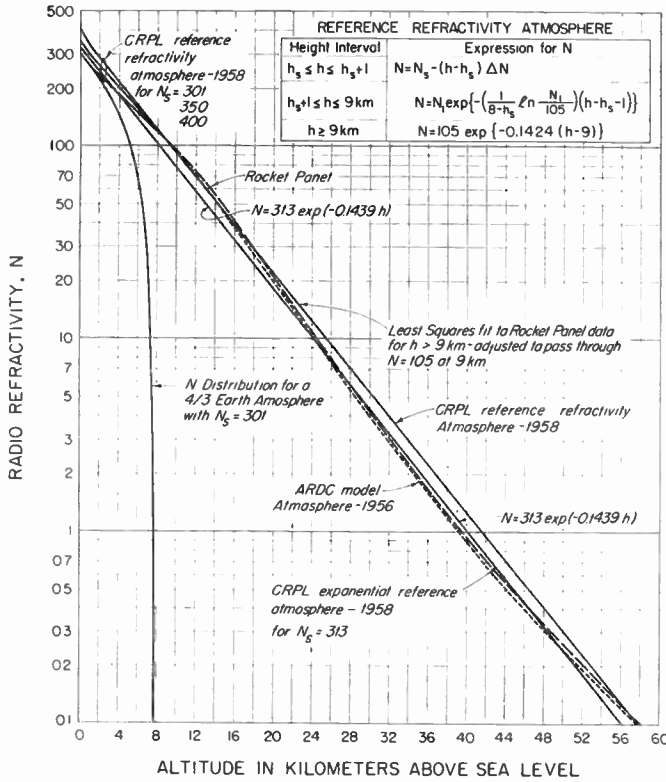


Fig. 5—CRPL Reference Refractivity Atmospheres-1958.

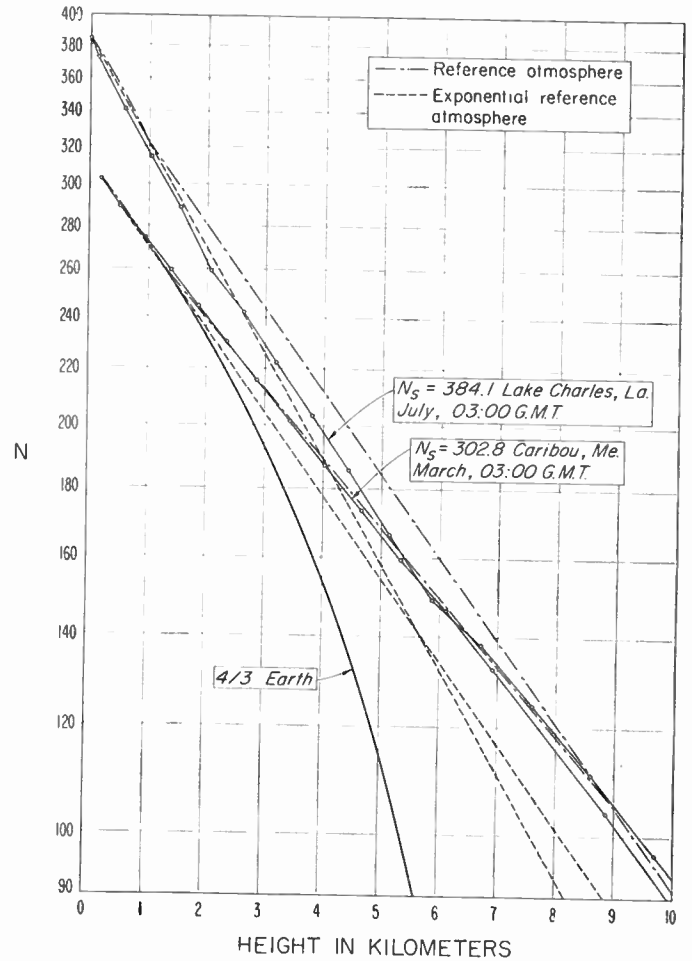


Fig. 6—Comparison of reference atmospheres with observed N profiles.

or

$$(1 + N \times 10^{-6})(a + h) \cos \theta = (1 + N_s \times 10^{-6})(a + h_s) \cos \theta_0, \quad (11)$$

where the radial distance from the center of the earth, r , equals $a + h$. The zero or s subscripts in (11) refer to the initial values at the earth's surface. The geometry of the problem was illustrated by Fig. 3. Eq. (11) may be used to calculate the local elevation angle, θ , at any point along the ray and thus affords a complete description of the ray path. If

$$n_2 r_2 \cos \theta_2 = n_1 r_1 \cos \theta_1,$$

and

$$\begin{aligned} n_2 &= n_1 + \Delta n, \\ r_2 &= r_1 + \Delta r, \end{aligned}$$

and

$$\theta_2 = \theta_1 + \Delta \theta,$$

where Δn , Δr , and $\Delta \theta$ are infinitesimals, then

$$n_1 r_1 \cos \theta_1 = (n_1 + \Delta n)(r_1 + \Delta r) \cos(\theta_1 + \Delta \theta), \quad (12)$$

and in the limit as Δr approaches zero, this becomes

$$n_1 r_1 \cos \theta_1 = (n_1 + dn)(r_1 + dr)(\cos \theta_1 - d\theta \sin \theta_1). \quad (13)$$

If (13) is expanded and products of differentials are omitted, then one obtains the differential equation

$$\tan \theta d\theta = \frac{dn}{n} + \frac{dr}{r} \quad (14)$$

or

$$d\theta = \frac{dn}{n} \cot \theta + \frac{dr}{r} \cot \theta. \quad (15)$$

Noting that $dr \cot \theta / r = r d\phi / r = d\phi$ and, from geometry, $d\theta = d\phi - d\tau$, one obtains the classic expression for the bending of a radio ray^{6,18}

$$d\tau = -\frac{dn}{n} \cot \theta$$

or

$$\tau_{1,2} = -\int_{n_1}^{n_2} \frac{dn}{n} \cot \theta. \quad (16)$$

Note that (16) defines downward bending as a positive quantity.

Eq. (16) is the basic measure of the effects of atmospheric refraction used throughout this study. Although these effects may be expressed in many ways, the bending offers a convenient method often used in classical astronomy, and is simply that angle formed by the intersection of the tangents to the ray at the two points

¹⁸ W. M. Smart, "Spherical Astronomy," Cambridge University Press, London, Eng., ch. 3; 1931.

being considered. The bending of radio rays in the model atmospheres was obtained by numerical evaluation of (16).

Once values of θ and τ have been obtained, they may be used to derive all other refraction variables. The elevation angle error, ϵ , was obtained from:

$$\epsilon = \tau - \text{Arc tan} \left[\frac{\frac{n_s}{n} - \cos \tau - \sin \tau \tan \theta_0}{\sin \tau - \cos \tau \tan \theta_0 + \frac{n_s}{n} \tan \theta} \right]. \quad (17)$$

This expression was suggested by K. A. Norton and allows for a more accurate determination of ϵ than can be obtained from a law of sines formulation for the true elevation angle. The true elevation angle, β , is then found from:

$$\beta = \theta_0 - \epsilon. \quad (18)$$

Further, the angular distance, ϕ , and the distance along the earth's surface, d , are obtained from

$$\phi = \frac{d}{a + h_s} = \tau + \theta - \theta_0, \quad (19)$$

while the true range, R_0 , is obtained from the law of sines as:

$$R_0 = \frac{(a + h) \sin \phi}{\cos \beta}. \quad (20)$$

The geometric path length, R , was obtained by summation of the chord lengths from point to point along the ray path:

$$\Delta(R) = \sqrt{r_2^2 + r_1^2 - 2r_1r_2 \cos(\phi_2 - \phi_1)} \quad (21)$$

where $r \equiv a + h$.

The "optical" or effective radio path length was obtained by numerical evaluation of:

$$R_e = \int_0^R ndR. \quad (22)$$

Finally the radio-range error, ΔR_e , and the geometric path length difference, ΔR , were evaluated as:

$$\Delta R_e = R_e - R_0, \quad (23)$$

and

$$\Delta R = R - R_0. \quad (24)$$

It must be remembered that all of the above refraction results are obtained under the assumption that the refractive index is horizontally homogeneous and is subject to the usual two restrictions of ray tracing: 1) the refractive index does not change appreciably in a wavelength, and 2) the fractional change in the spacing between neighboring rays must be small in a wavelength.

These conditions are met in the model atmosphere for radio frequencies greater than, say, 100 kc, since N varies more slowly with height than in the ducting case and caustics do not occur since the bending is calculated for ground-based antennas over a spherical earth. It is conceivable that for some air-to-air refraction calculations, the discontinuity in N gradient at h_s+1 and at 9 km would violate these assumptions.

COMPARISON OF MODEL ATMOSPHERES WITH OBSERVATION

A comparison of the ray paths in the reference atmospheres and in the 4/3 earth model will illustrate the systematic differences between these models. Such a comparison is given in Fig. 7 for a distance of 200 miles and a height of 14,000 feet and in Fig. 8 for a distance of 800 miles and a height of 240,000 feet. The particular graphical presentation used in Figs. 7 and 8 shows the 4/3 earth rays as straight lines. It is noted that the 4/3 earth ray at $\theta_0=0$ is in relatively good agreement with the values from the reference atmosphere for distances out to 200 miles and heights up to 14,000 feet, but systematically departs from the reference atmosphere for greater distances and heights. For a range of 600 miles, where the ray reaches heights of about 200,000 feet, the 4/3 earth ray is some 9000 feet lower than the $N_s=400$ reference atmosphere and 36,000 feet lower than the $N_s=250$ reference atmosphere. This height discrepancy is due to the 4/3 earth model's unrealistically large N gradient at high heights with resultant increased bending.

As a further comparison, the bending in the 4/3 earth atmosphere is compared with that in the exponential reference atmosphere in Fig. 9. The bending in an "average" atmosphere is also given. This average atmosphere is a composite of the 5-year mean profiles for both summer and winter at the eleven U. S. radiosonde stations enumerated in the following paragraph, and was used as a readily-available measure of average conditions. The important point made by Fig. 9 is that the 4/3 earth model is systematically in disagreement with average bending: at low heights it gives too little bending, while at high altitudes it gives too much bending. The exponential reference atmosphere does not appear to be systematically biased, and deviates less than 5 per cent from the average atmosphere. It is significant that the exponential reference and the average atmosphere are in essential agreement as to the shape of the τ -height curve.

It would now be instructive to compare the bendings obtained from the various models with values obtained from each of the 5-year mean N profiles from different climatic regions. The 5-year mean N profiles were obtained for both summer and winter for a variety of climates as represented by the states of Florida, Texas, Maine, Illinois, Nevada, California, North Dakota, Washington, Nebraska, Wyoming, and the District of Columbia.

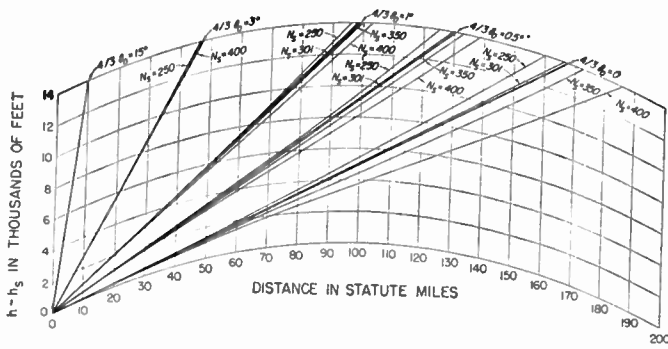


Fig. 7—Comparison of rays in the CRPL Reference Refractivity Atmosphere-1958 and the 4/3 earth atmosphere, 0 to 200 miles.

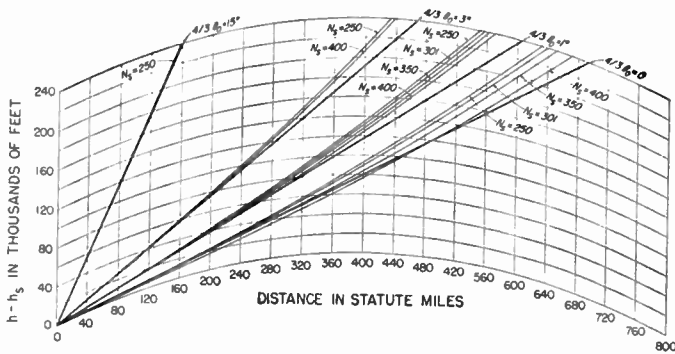


Fig. 8—Comparison of rays in the CRPL Reference Refractivity Atmosphere-1958 and the 4/3 earth atmosphere, 0 to 800 miles.

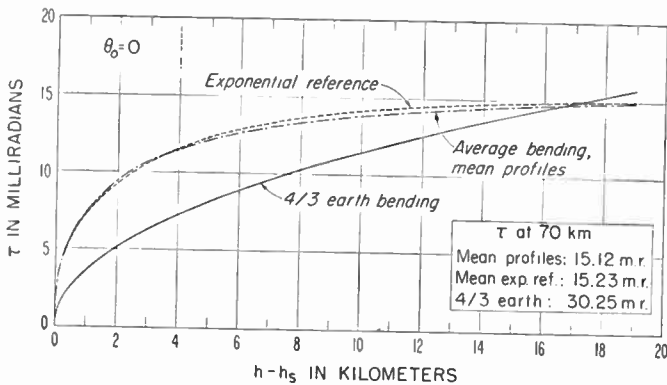


Fig. 9—Comparison of long-term mean bending with the values predicted by the exponential reference atmosphere and the 4/3 earth atmosphere.

Comparisons of the bending obtained from the 4/3 earth model and the bendings obtained from the 5-year mean N profiles with the reference atmospheres are shown in Fig. 10(a) and 10(b). This figure was selected to illustrate the range of agreement between the models and the expected long-term average bendings. Fig. 10(a) gives a comparison for a small initial elevation angle, $\theta_0 = 0$, and a small height increment, $h - h_s = 3$ km, and shows that both reference atmospheres tend to set a lower limit to the bendings. In this case, the exponential reference atmosphere appears to be in better agreement with the expected long-term mean bendings than does the reference atmosphere. The numbered data points for Washington, D.C., Omaha, Neb., and Santa Maria, Calif., are of special interest. Washington and Omaha have the only

long-term mean N profiles with initial N gradients (*i.e.*, $-112/\text{km}$ and $-106/\text{km}$, respectively) that are significantly greater than would be expected from the ΔN vs N_s relationship. Both of these stations have an unusually large humidity decrease near the ground. The third point, Santa Maria, Calif., is of interest since it is in relatively good agreement with the reference atmospheres, even though it represents the Southern California summer climate which was excluded from the original ΔN vs N_s relationship. This agreement is attributed to the reference atmospheres being a good representation of the N distribution below the California elevated inversion and to the fact that a majority of the bending is accomplished below the elevated inversion height of about 500 meters. Further, it can be easily shown that the bending integral is increasingly insensitive to strong N discontinuities as the height increases. Fig. 10(b) shows a similar comparison for a high initial elevation angle, $\theta_0 = 15^\circ$ and a large height increment, $h - h_s = 70$ km. This comparison shows that both of the reference atmospheres are in closer agreement with the long-term mean bendings than are the 4/3 earth bendings. Note that, whether τ is predicted from N_s or ΔN , the 4/3 earth model gives but a single value of bending that is outside the limits of the values of τ obtained from the long-term mean profiles.

In considering the comparisons of Fig. 10(a) and 10(b), one might ask if they reflected the form of the basic equation for bending; namely, at low angles is τ determined by the N gradient throughout the N profile, and at high angles is τ essentially a function of the value of N at both ends of the N profile, *i.e.*, the limits of integration? Thus one might expect the deviations to be smaller if the comparisons were made on the basis of a function of the N gradient such as ΔN , particularly for small values of θ_0 . Such a comparison is given by Fig. 10(c) and 10(d) for the same initial elevation angles and height increment as before. It is seen that the ΔN -specified reference atmospheres improve the agreement for the low-angle case, but decidedly decrease the agreement for the high-angle case.

A numerical evaluation of the root mean square (rms) deviation of the long-term mean bendings from both the reference atmospheres determined as a function of both ΔN and N_s was made for a variety of initial elevation angles for the height increments 3 and 70 km. Root mean square deviations were not calculated for the 4/3 earth model since it was felt that this model was obviously in marked disagreement with the long-term mean bendings under these conditions. Fig. 11 summarizes the rms deviations for the $h - h_s = 3$ -km case. It is seen that for $\theta_0 < 10$ milliradians (about 0.5°), the ΔN -specified reference atmospheres have the smaller rms deviations. Also, the exponential reference atmospheres, whether specified by ΔN or N_s , have smaller rms deviations than the reference atmosphere.

It is seen for the 70-km case, Fig. 12, that the N_s -specified reference atmospheres have a significantly

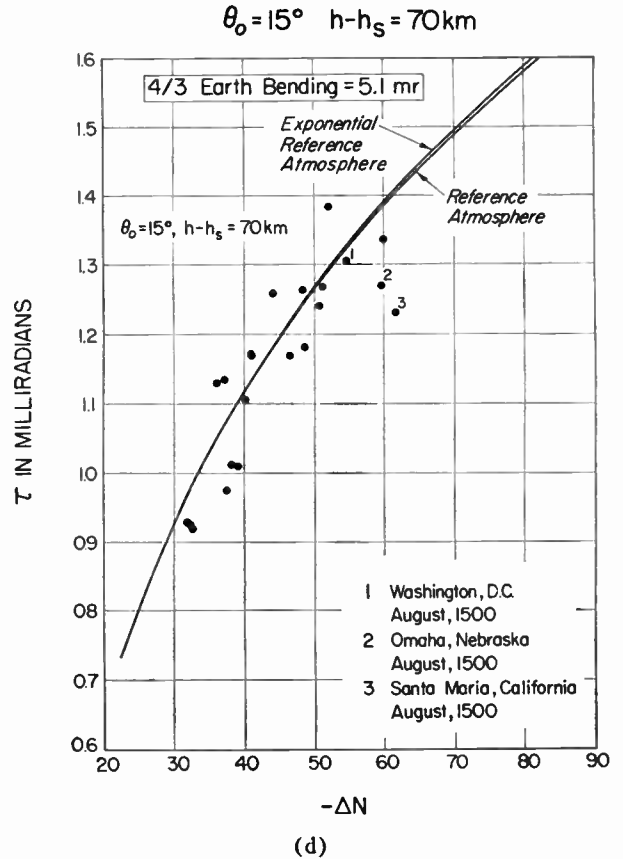
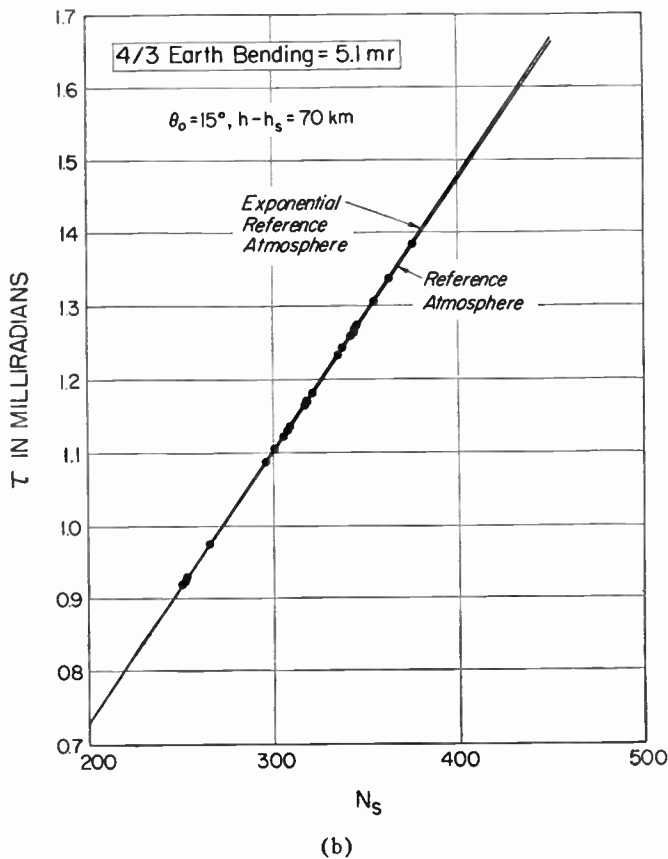
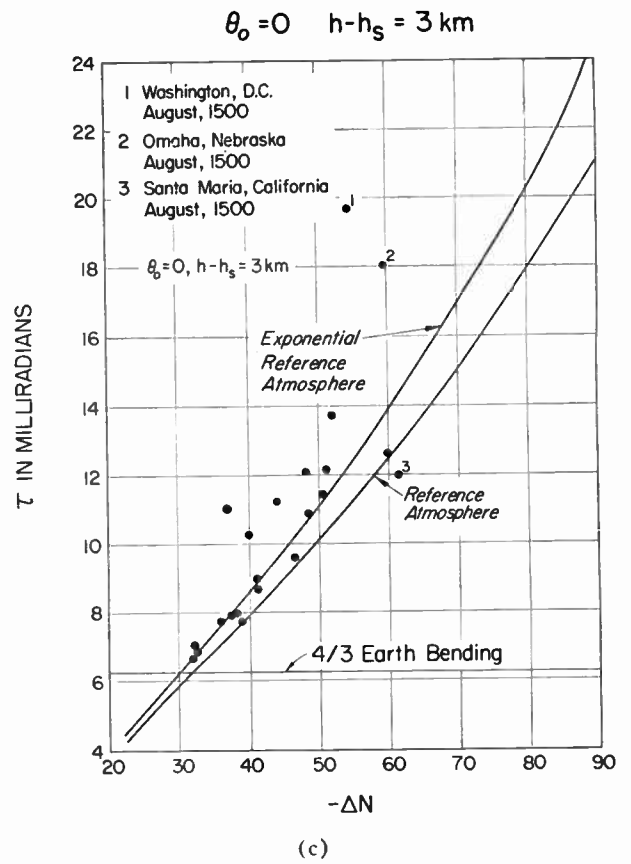
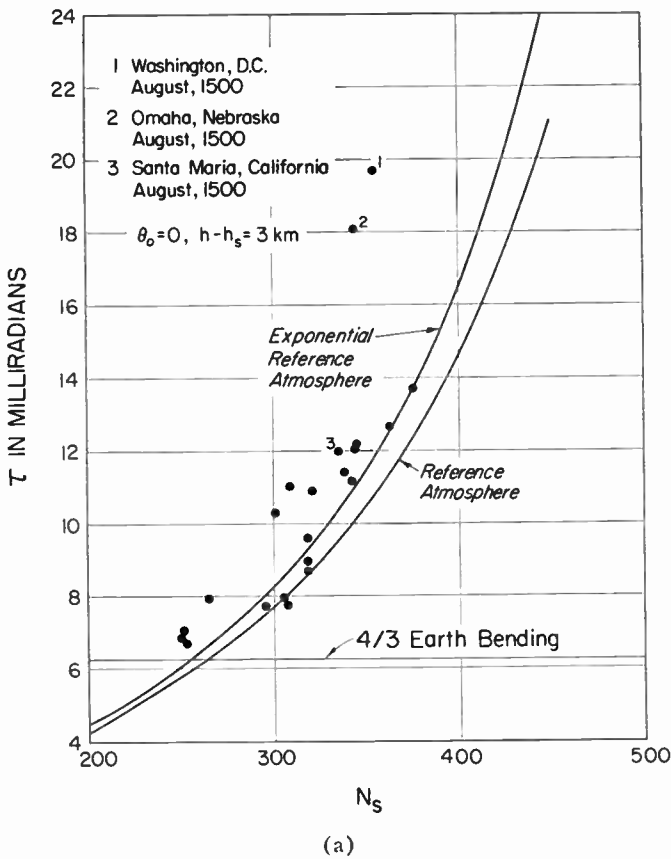


Fig. 10—Comparison of τ vs N_s and ΔN as obtained from the CRPL reference atmospheres and 5-year mean radiosonde data: (a) $\theta_0 = 0$, $h - h_s = 3$ km, (b) $\theta_0 = 15^\circ$, $h - h_s = 70$ km, (c) $\theta_0 = 0$, $h - h_s = 3$ km, and (d) $\theta_0 = 15^\circ$, $h - h_s = 70$ km.

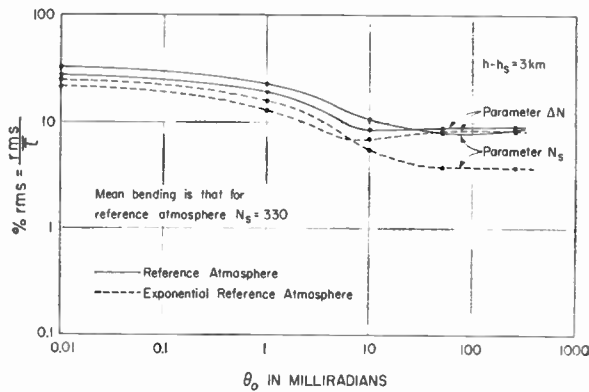


Fig. 11—Comparison of per cent rms deviations of 5-year mean profile bendings about the CRPL reference atmospheres using two parameters, N_s and ΔN ; $h-h_s=3$ km.

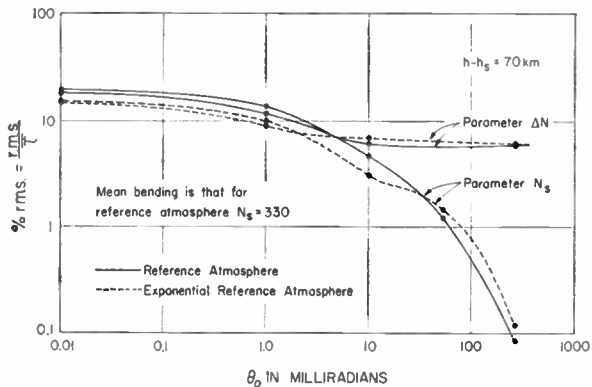


Fig. 12—Comparison of per cent rms deviations of 5-year mean profile bendings about the CRPL reference atmospheres using two parameters, N_s and ΔN ; $h-h_s=70$ km.

smaller rms deviation than the ΔN -specified atmospheres for $\theta_0 > 5$ milliradians. Again it is seen that the exponential reference atmosphere generally has the smaller rms deviation for values of θ_0 less than 10 milliradians. However, the slightly smaller rms deviations associated with the reference atmosphere for $\theta_0 > 10$ milliradians is felt to be a reflection of that model's closer agreement with the actual N structure of the atmosphere at high heights.

The above analysis appears to indicate that the exponential reference atmosphere allows a more accurate determination of the effects of atmospheric refraction than does either the reference atmosphere or the 4/3 earth atmosphere, and thus we are led, in the next section, to express these effects in terms of the exponential reference atmosphere.

APPLICATIONS OF THE EXPONENTIAL REFERENCE ATMOSPHERE

There are many aspects of radio engineering and propagation studies to which the present reference atmospheres can be applied. Some of these applications will be mentioned below. It is quite evident that a thorough treatment of any one application would require more detail than is warranted in the present paper.

A radar set indicates the apparent range and angular

elevation of a target. These apparent ranges may be corrected by the difference in the apparent range and the true geometric range, *i.e.*,

$$R_0 = R_e - \Delta R_e. \quad (25)$$

Values of ΔR_e as determined from the exponential reference atmosphere are given for various initial elevation angles and radio ranges on Figs. 13 and 14. By a similar procedure the true elevation angle, β , may be obtained from

$$\beta = \theta_0 - \epsilon, \quad (18)$$

where ϵ may be determined as a function of θ_0 and R_e from Figs. 15 and 16.

The above application of the reference atmosphere to the radar range and elevation angle error problem is but one of many applications that might have been chosen to illustrate the use of this new model. For example, refraction affects the calculation of the radio horizon distance with a resultant modification to diffraction fields beyond the horizon. In addition, tropospheric and ionospheric scatter-propagation calculations would be modified to account for the difference of the angle of incidence in the effective earth's radius and the exponential models of the atmosphere. Many of these effects have already been considered in forthcoming articles by Norton,¹⁹ Millington,²⁰ and Rice, Longley, and Norton.²¹

In addition, the present authors are preparing extensive tables and graphs of the various refraction variables of the exponential reference atmosphere that should allow the reader to evaluate his own applications.

A NEW METHOD OF PREDICTING THE BENDING OF RADIO RAYS

The extreme importance of the initial gradient of refractivity in determining the bending at very low elevation angles has already been discussed. In view of this, it seemed desirable to explore the possibility of estimating τ as a function of both N_s and the initial gradient of N , hereafter denoted as $\delta N'$.

As a basic postulate it was assumed that the exponential reference atmosphere defined the expected bending under any refractive index profile conditions and that most of the error of prediction arises from the fact that $\delta N'$ differs from the value predicted by the model atmosphere. Since it was observed that the error in predicting τ became constant after the first 100 meters or so above the earth's surface, it was further assumed that this error of prediction could be reduced by simply adding the bending due to the difference in initial gradients.

¹⁹ K. A. Norton, "Low and medium frequency radio propagation," Proc. Internatl. Conf. on Electromagnetic Wave Propagation. (To be published in book form by the Academic Press, New York, N. Y., Spring of 1959.)

²⁰ G. Millington, "Propagation at great heights in the atmosphere," *Marconi Rev.*, vol. 21, pp. 143-159; November, 1958.

²¹ P. L. Rice, A. G. Longley, and K. A. Norton, "Prediction of the Cumulative Distribution with Time of Ground Wave and Tropospheric Wave Transmission Loss," NBS Rep. 5582; June 30, 1958.

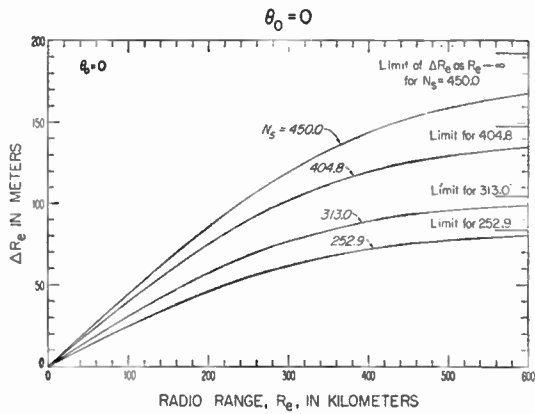


Fig. 13—Radar-range error as a function of the apparent range and N_s , $\theta_0=0$.

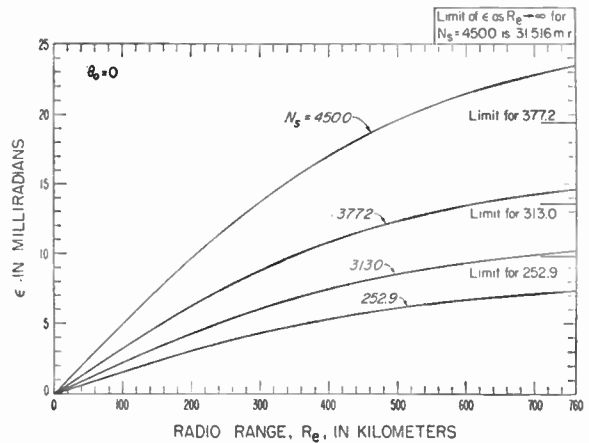


Fig. 15—Radar elevation-angle error as a function of the apparent range and N_s , $\theta_0=0$.

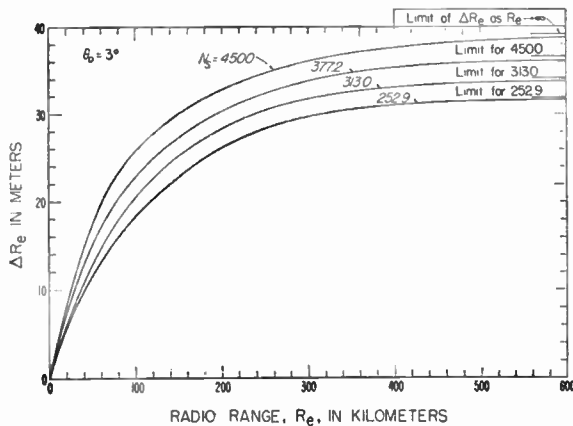


Fig. 14—Radar-range error as a function of the apparent range and N_s , $\theta_0=3^\circ$.

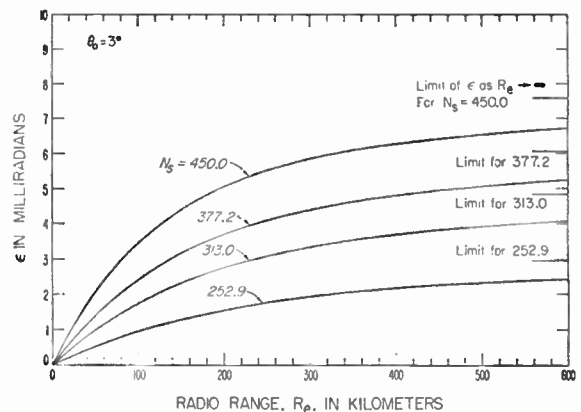


Fig. 16—Radar elevation-angle error as a function of the apparent range and N_s , $\theta_0=3^\circ$.

This may be expressed as

$$\tau = \tau_E + \delta\tau_E, \tag{26}$$

where τ_E is determined from the exponential reference atmosphere for the particular values of N_s , θ_0 and $h-h_s$ desired. The additive correction factor, $\delta\tau_E$, is a function of both the initial elevation angle and $\delta N'$. It was assumed, arbitrarily, that a single correction could be applied to the bending at all heights and that this correction could be determined by assuming that the initial N gradient occurred in a ground-based layer 100-meters thick, regardless of the actual observed thickness.

In practice we may approximate $\delta N' \equiv (dN/dh)_{h=h_s}$ by δN where

$$-\delta N \equiv \frac{N_s - N_h}{h - h_s}. \tag{27}$$

The value of N_h is obtained either from the first level reported by the radiosonde above the surface or from tower or tethered-balloon measurements; $h-h_s$ will range in practice from, say, 30 to 300 meters. We will wish to know what value of N_s the observed value of δN will correspond to, in the exponential reference atmosphere. Since, in the exponential reference atmosphere

$$N(h-h_s) = N_s \exp \{ -(h - h_s)c_e \} \tag{28}$$

where

$$c_e = \ln \frac{N_s}{N_s + \Delta N},$$

then the value of N_s corresponding to δN , designated as N_s^* , may be determined from

$$-\delta N = \frac{N_s^* - N_s^* \exp \{ -(h - h_s)c_e^* \}}{h - h_s}$$

or

$$N_s^* = \frac{-\delta N(h - h_s)}{1 - \exp \{ -(h - h_s)c_e^* \}} = \frac{-0.1\delta N}{1 - \exp \{ -0.1c_e^* \}}. \tag{29}$$

where c_e^* is the decay factor corresponding to N_s^* . In the second expression on the right in (29) we have assumed, quite arbitrarily, for the purpose of our correction process that $h-h_s=0.1$ km, regardless of the actual height over which δN was determined. Thus N_s^* is the value of N_s specified by the exponential reference atmosphere when the N difference between the surface and 100 meters is equivalent to δN . Since (29) is a

transcendental function of N_s^* , it is more convenient, in practice, to use graphical methods, such as those in Fig. 17. The additive correction factor is then the simple difference between the 100-meter layer bendings as determined by N_s and N_s^* which may be obtained from Fig. 18. This may be expressed as

$$\delta\tau_E = \tau_{0.1}(N_s^*) - \tau_{0.1}(N_s), \quad (30)$$

where the subscript 0.1 refers to the bending in the initial 100 meters above the surface.

This method of prediction was tested by comparison with all available bending data, which comprised the previously-enumerated 22 different N profiles typical of 5-year mean conditions as well as 67 individual N profiles typical of superrefraction and elevated layer conditions. The superrefractive profiles, which comprise 32 of the 89 profiles, are those with an initial layer in which N decreases more rapidly than $-100 N$ units/km as compared to the normal decrease of about $-45 N$ units/km. It is important to repeat that the initial gradient is not derived from N_s and the value at 100 meters, but rather from the first reported radiosonde level regardless of its actual height. It could as easily be obtained from tower measurements of the common meteorological elements of pressure, temperature and humidity.

The radio-ray bending calculated for each of the 89 different N profiles, for $\theta_0=0$, and for $h-h_s=70$ km is compared to the exponential reference atmosphere on Fig. 19 and the values of observed τ minus the correction factor are plotted in Fig. 20. It is seen that this method produces a considerable reduction in the scatter of points about the exponential reference atmosphere curve. Indeed, by the addition of the initial gradient correction, the exponential reference atmosphere no longer forms a lower boundary, but rather has the appearance of a least squares fit to the data.

At first glance, one might assume that the improvement of prediction obtained by the initial gradient correction applied only to the superrefractive conditions. This is not true, however, since the application of $\delta\tau_E$ to the 5-year mean data alone reduces the rms deviation of that data about the exponential reference atmosphere from 2.17 to 0.64 milliradians, a reduction of 71 per cent in the uncertainty of prediction.

It is expected that this method will yield equally good results for heights of less than 70 km. Some reduction of the scatter of data points about the exponential reference atmosphere curve is expected for higher initial elevation angles, even though the effects of the surface layer become less pronounced with increasing elevation angle. Preliminary investigation indicates that the $\delta\tau_E$ correction would result in no more than a 50 per cent reduction of rms error at $\theta_0=10$ milliradians ($\sim 0.5^\circ$) and the reduction would be negligible for θ_0 of 3° or more.

The $\delta\tau_E$ correction procedure does not appear to be critically dependent on the thickness assumed for the initial layer so long as this thickness is about 100 meters.

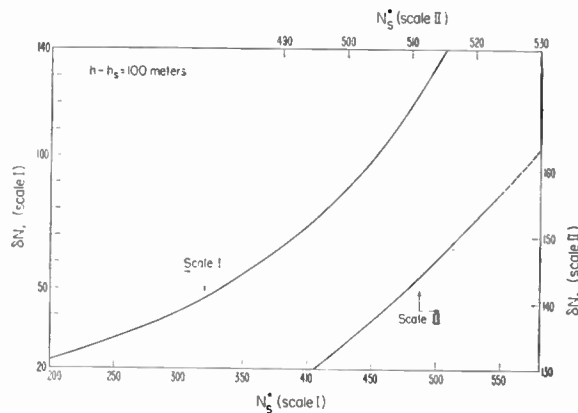


Fig. 17—Effective $N_s = N_s^*$ associated with the observed initial N gradient, δN_s , from the exponential reference atmosphere.

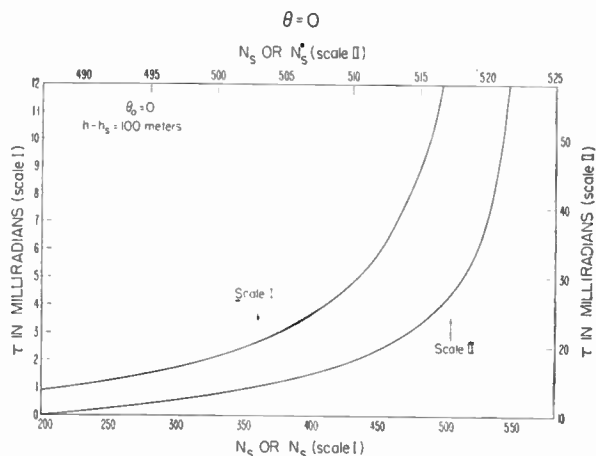


Fig. 18—Bending in the first 100 meters as a function of N_s or N_s^* .

CONCLUSIONS

The most important conclusion of the present study is that the average refractivity structure of the atmosphere is more nearly represented by an exponential decrease with height than by the linear decrease inherent in the effective earth's radius model. This then leads to the conclusion that the effective earth's radius model is a systematically biased estimation of refraction effects, indicating too little refraction between the earth and about 16 km, and too much refraction above that height. The 4/3 earth model indicates nearly twice as much bending at the base of the ionosphere as is obtained from long-term mean N profiles.

Another conclusion is that a rather well-defined exponential relationship exists between N_s and ΔN and that the refractivity is nearly constant at an altitude of 9 km. By utilizing this information, model atmospheres may be constructed. The bending of radio rays derived from these model atmospheres is in good agreement with the bending obtained from mean radiosonde data, although it yields values noticeably lower than those obtained from superrefractive profiles. This latter point led to a critical examination of whether ΔN or N_s is the better predictor. The general conclusion of this examination is that N_s is the better predictor for $\theta_0 \geq$ milli-

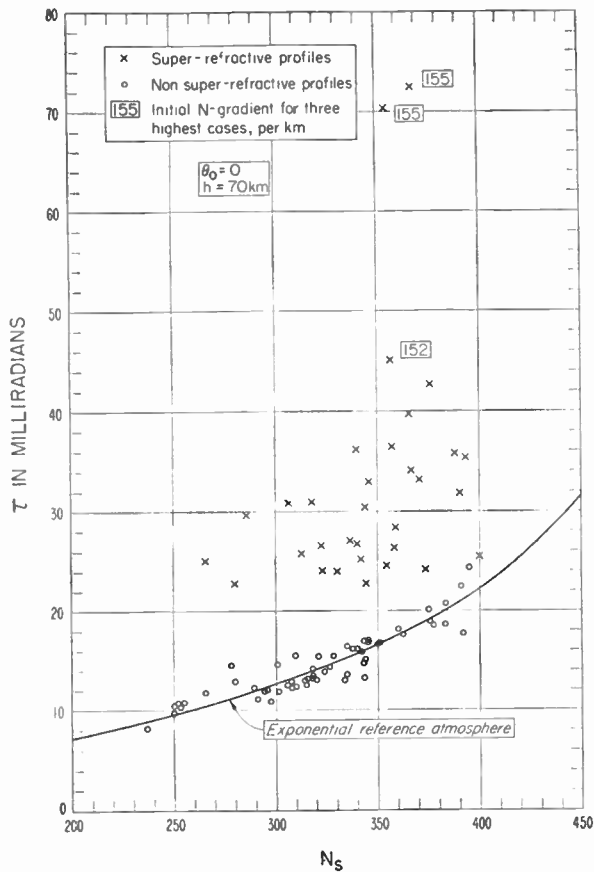


Fig. 19— τ vs N_s from 89 profiles compared with values predicted by the exponential reference atmosphere, $\theta_0=0$ and $h-h_s=70$ km.

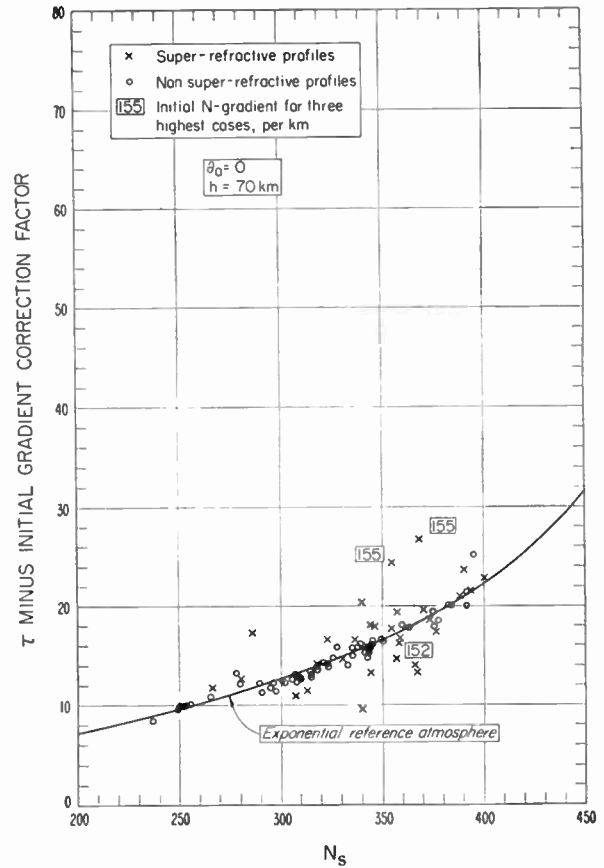


Fig. 20—Corrected τ vs N_s from 89 profiles and the values predicted by the exponential reference atmosphere, $\theta_0=0$ and $h-h_s=70$ km.

radians and that ΔN is the better predictor for $\theta_0 < 10$ milliradians, although neither N_s nor ΔN is in good agreement with the bending from all profiles at $\theta_0=0$. If, however, the value of the initial gradient is available, then one may adequately predict the total bending even for superrefractive profiles and $\theta_0=0$.

In the absence of detailed N profile information, one may estimate the appropriate values of the reference atmospheres from climatic charts such as those in Figs. 21–24. These charts were prepared from average values of N_s determined from individual weather observations taken over an 8-year period at 62 U. S. Weather Bureau stations for day and night, during both summer and winter. Figs. 21–24 are in terms of a reduced-to-sea-level value N_0 rather than N_s , since it has been shown to be four or five times more accurate to estimate N_s from charts of N_0 than from charts of N_s .²² N_s is recovered from these charts by use of

$$N_s = N_0 \exp \{ - 0.1057 h \}$$

where h is the station altitude in kilometers above sea level. World-wide charts of N_0 are being published.²²

The model atmospheres presented above were derived

to be useful in a wide variety of climates. The method could be improved by deriving individual regression lines of ΔN upon N_s for each station or climatic region and then, perhaps, presenting maps of the statistically determined constants of these regression equations. Some recent French reports^{23,24} on N climatology indicate essential agreement with the present N_s - ΔN relationship for France, but some disagreement for North Africa, thus perhaps indicating the need for some adjustment of the regression equation for tropical regions. This need is somewhat obscured, however, by the shortness of the meteorological records for North Africa (in general, one year of record) since this area is noted for its large year-to-year variations of N .

ACKNOWLEDGMENT

The authors wish to acknowledge the technical assistance and aid of B. A. Cahoon and B. J. Weddle. We wish to express particular appreciation to K. A. Norton for many stimulating discussions and his active participation in the derivation of accurate methods of determining the ray path.

²³ F. du Castel and P. Misme, "Elements de radio climatologie," *Onde Electrique*, vol. 37, pp. 1049–1052; November, 1957.

²⁴ F. du Castel and P. Misme, "Etude statistique de parametres radio meteorologiques," *Ministere des P.T.T., Centre National D'Etudes des Telecommunications, Dept., "Transmission," Issy-les-Moulineaux, France, Note Tech. d'Information No. T 150; June 12, 1958.*

²² B. R. Bean and J. D. Horn, "The radio refractive index near the ground," *J. Res. NBS.*, pt. D (Radio propagation), to be published. (A preliminary version of this paper is available from the authors as NBS Rep. No. 5559; March 3, 1958.)

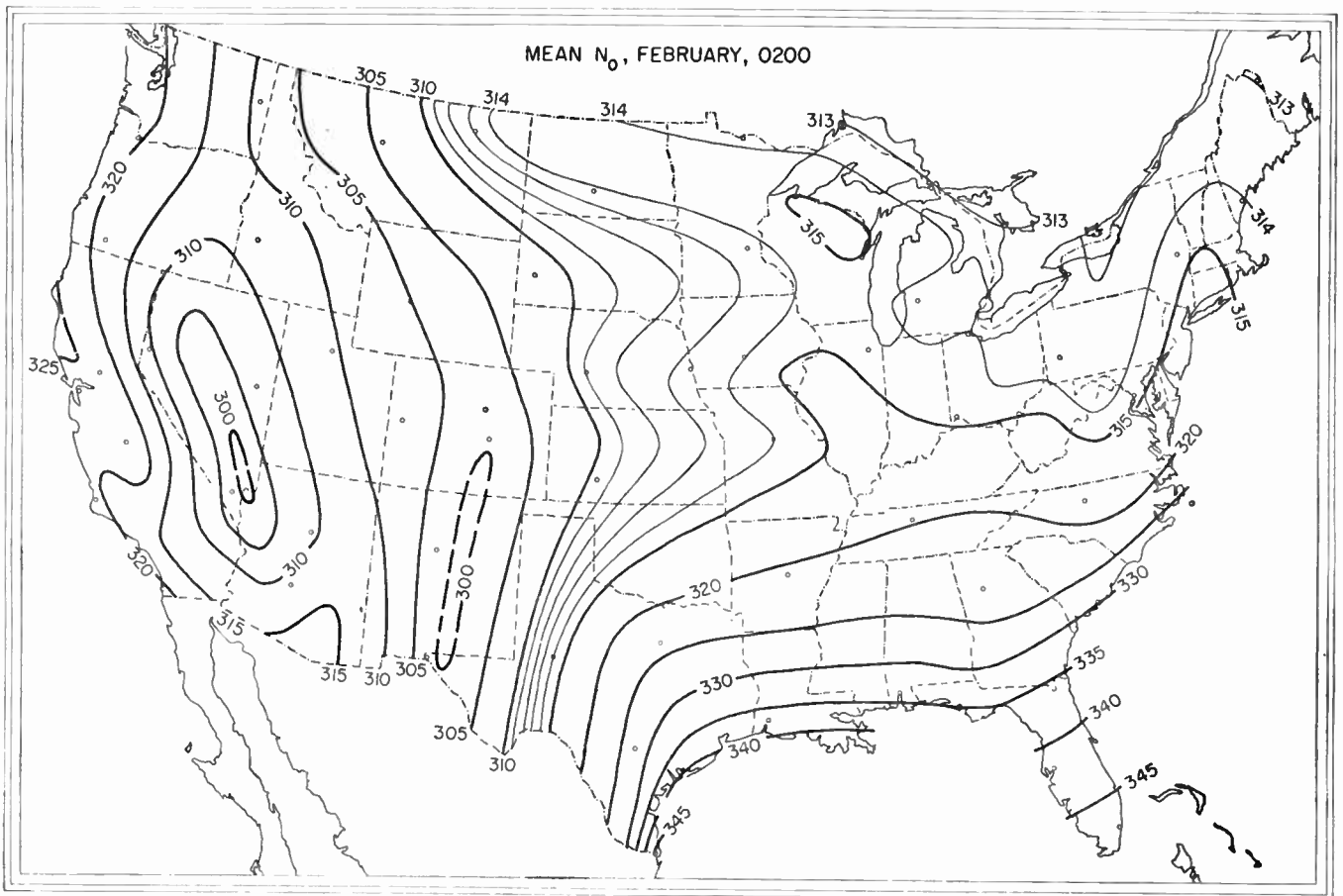


Fig. 21—Reduced-to-sea-level refractivity typical of winter nights as represented by 0200 local time during the month of February.

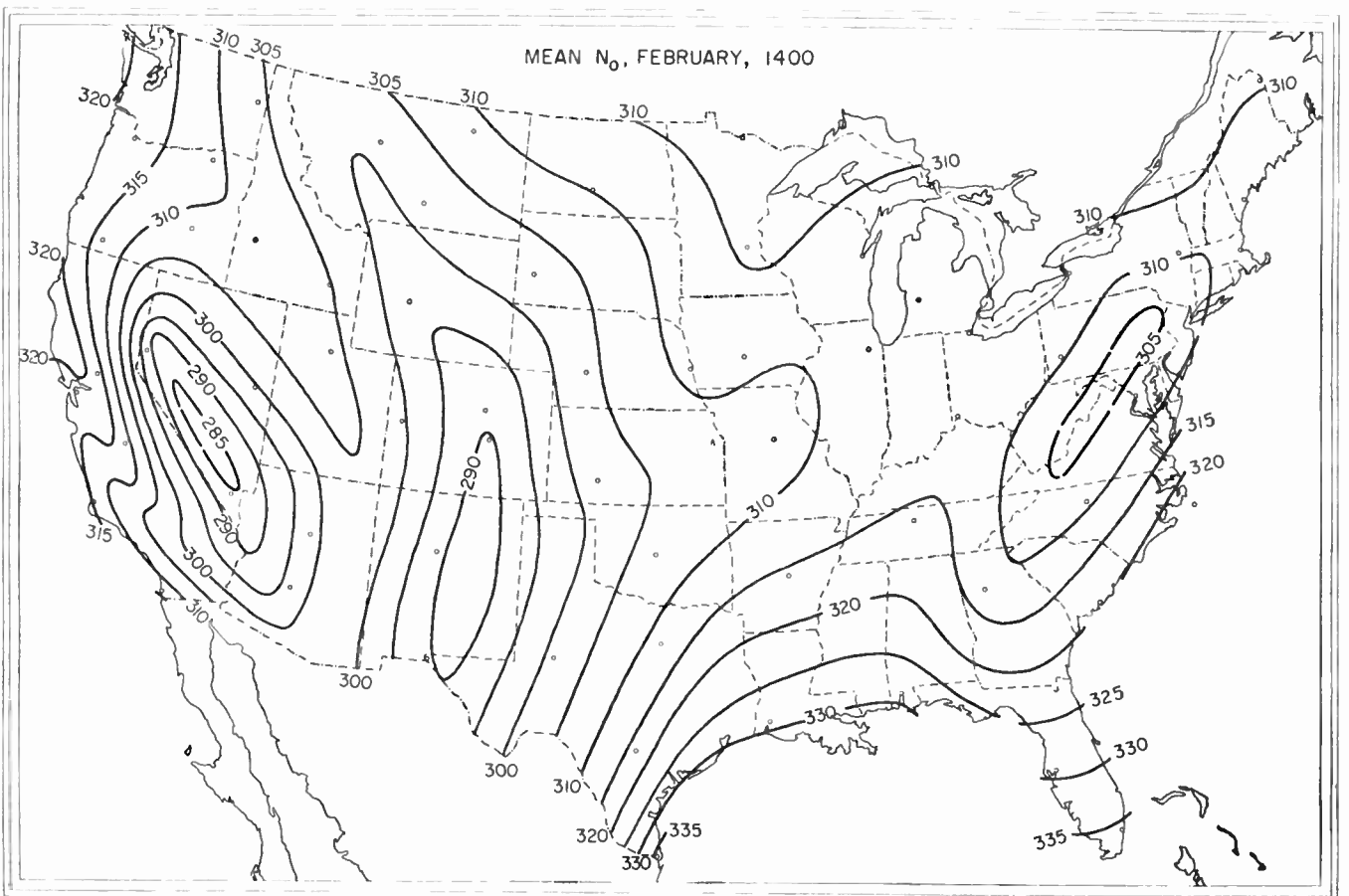


Fig. 22—Reduced-to-sea-level refractivity typical of winter days as represented by 1400 local time during the month of February.

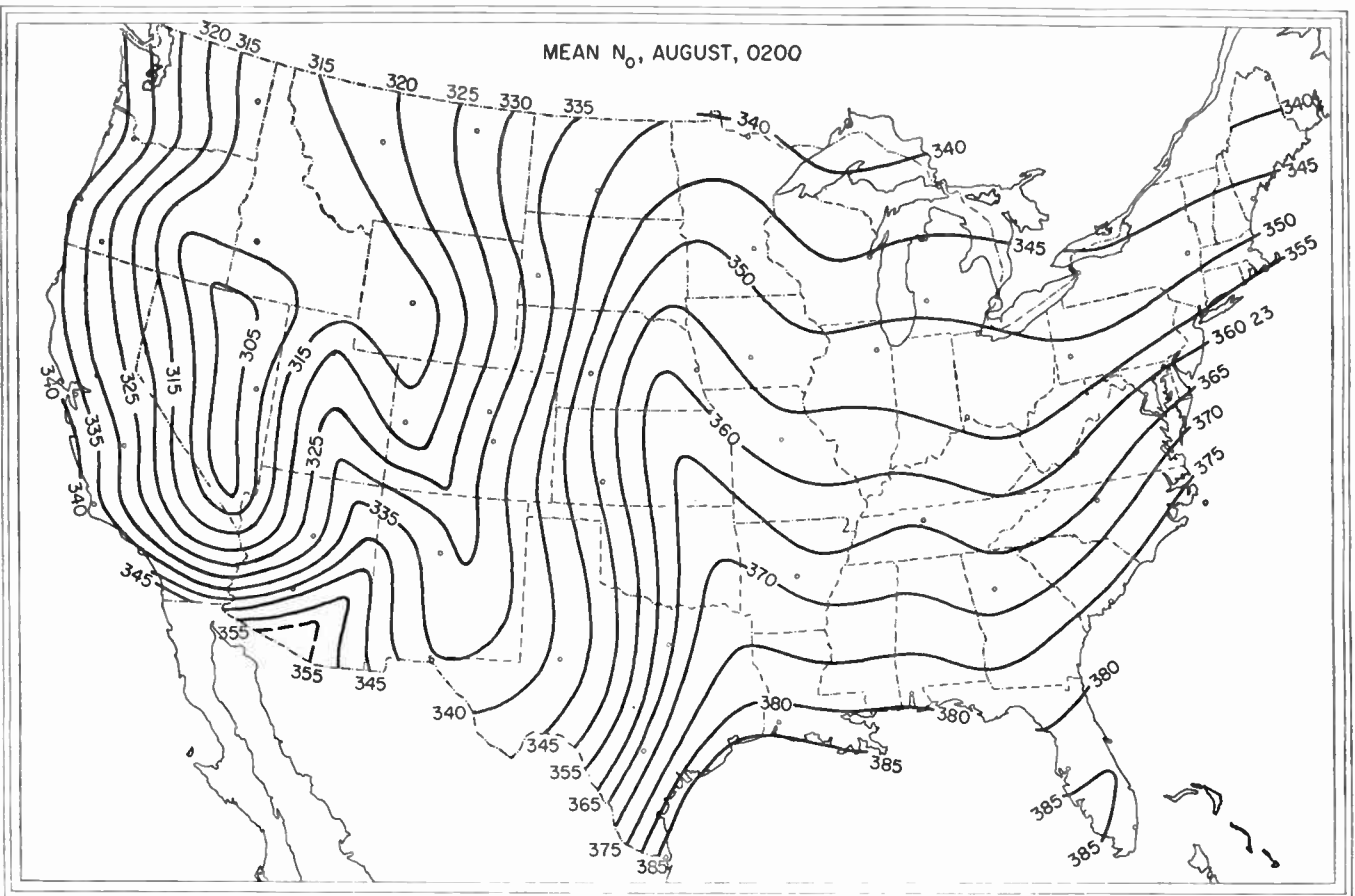


Fig. 23—Reduced-to-sea-level refractivity typical of summer nights as represented by 0200 local time during the month of August.

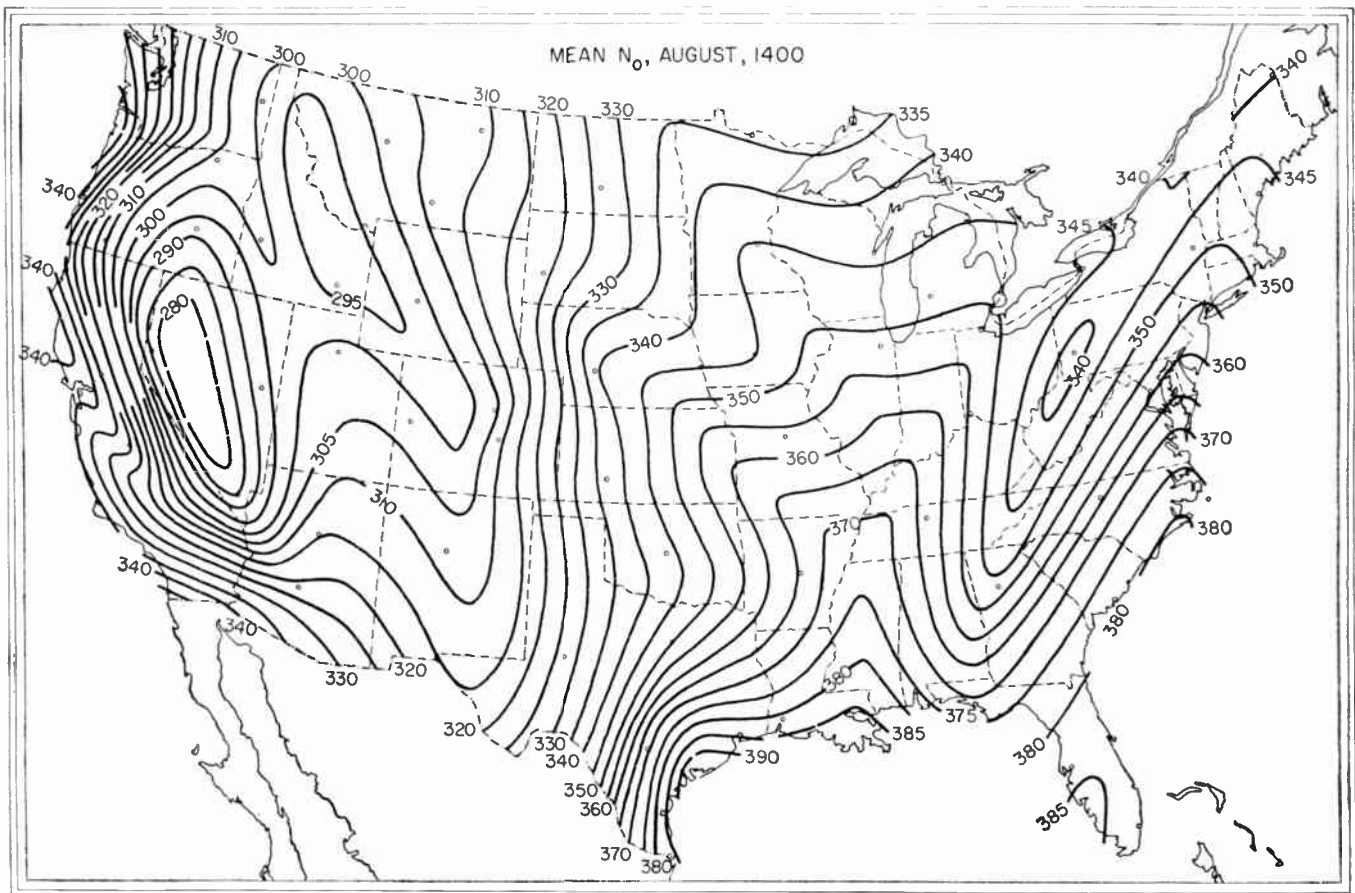


Fig. 24—Reduced-to-sea-level refractivity typical of summer days as represented by 1400 local time during the month of August.

Airborne Radiometeorological Research*

W. S. AMENT†, SENIOR MEMBER, IRE

Summary—This paper deals with the Naval Research Laboratory's airborne research in military radar and propagation problems. The military problems include predictions of radar and propagation conditions at specific places and times, and hence, fruitful research must be directed at establishing the geophysical causes of signal behavior. The flying laboratory system for quantitative measurements of ground and sea-clutter and of microwave field strengths is described, and some of the experimental conclusions given. Philosophy and equipment of a new research aircraft for radar and propagation studies is described, and its optimum use in airborne propagation research is outlined.

INTRODUCTION

THE military services use many kinds of microwave radar and communications systems in ground-to-ground, ground-to-air, and air-to-air geometries. Each system is affected to some degree by the refractive conditions of the atmosphere through which the microwaves are propagated. Thus the ability to determine and predict "propagation weather," the state of the atmosphere as it affects microwave transmission, has obvious military value.

The Naval Research Laboratory conducts research in microwave radar and propagation phenomena and their geophysical causes. The chief tool in this program is an aircraft, heavily instrumented for direct radiometeorological measurements and for microwave radar and radio field measurements. Because of its cost and application, most airborne radiometeorological work is conducted in government laboratories.

It is felt that, for this Government Research Issue of the PROCEEDINGS, an article describing some of the problems, methods, and potentialities of research with a general-purpose aircraft is more relevant than a technical paper in which the airplane appears incidentally as an aid to measurement. Therefore, airborne radiometeorological research is described against an outlined background of the military problem and the state of radiometeorological knowledge. Conclusions as to fruitful airborne research procedures are described and will be incorporated in the instrumentation of a new aircraft and its proposed use.

SCIENTIFIC ASPECTS OF TWO MILITARY PROPAGATION WEATHER PROBLEMS

The following indicates two examples of military needs for improved understanding of the causal connection between synoptic meteorological and propagation conditions, and outlines the primitive state of scientific knowledge applicable to these needs.

Example 1: Airborne picket-radar coverage is limited by radio holes [1] caused by certain atmospheric layers, and targets are obscured on PPI scopes by sea clutter [2]. Can the meteorologist predict layer altitudes, radio hole depths, and the magnitude of the clutter-producing waves? Alternatively, can special instrumentation rapidly determine clutter conditions and layer structure? Given the predicted or measured conditions, can a ray tracer [3], or other methods, determine the tactical geometry to best provide a specified degree of radar coverage?

Instrumentation [4], [5] for determining layer structure is under development, and the surface winds chiefly responsible for clutter levels can be predicted with some accuracy. Thus an *ad hoc* solution to this relatively simple problem can, at least, be pursued intelligently. Airborne measurements of field strength, layer structure, and clutter magnitude have already been made (and will be mentioned later) for delimiting parts of the problem, but airborne research into the problem as a whole has not yet been attempted.

Example 2: The landing operations of the last war relied heavily on scientific predictions of visibility, ocean currents, and surf conditions. In an electronic-age landing, equally reliable forecasts are required for radar, jammer, intercept, and scatter communication ranges for specific hours at a specific shoreline. The meteorologist must predict not only the mean atmospheric refractive stratification but also the statistics of refractive irregularities (blobs) at and between the stratifications, the prediction to include inshore-offshore effects imposed by the land mass in question. This prediction is not possible now but would be demanded were a landing to be scheduled next week.

Atmospheric phenomena, known to constitute or manifest stratifications affecting microwave propagation, and which, therefore, require prediction are: the incidence and characteristics of radiation cooling [6] and trade-wind surface ducts [7], ground fog, subsidence [8], [9], and trade-wind inversions [10], the top of the mixing level, haze and stratus layers [11], [12], Bauer layers [13], the jet stream [12], and probably the tropopause [14].

Ducts were investigated originally with surface equipment. In the first definitive work [7] on the trade-wind surface duct, NRL's transmitters were on a moving ship which made possible unquestionable measurements of the ducting and the first clear-cut discovery of scatter propagation. The propagation effects of the elevated stratifications were originally discovered and investigated with airborne radio and/or refractometer

* Original Manuscript received by the IRE, December 11, 1958; revised manuscript received, February 16, 1959.

† Naval Research Lab., Washington 25, D. C.

flights. The importance of the instrumented aircraft for sheer exploration in this area is obvious.

Meteorologists understand qualitatively the physical origins of stratifications but are unable to predict their horizontally averaged dielectric structure, let alone the local deviations from that average which constitute blobs. Thus, the first task of propagation weather research is to observe particulars and statistics of the atmosphere's dielectric structure on the micro and meso-scale of a propagation path, and to learn how to deduce that structure from the causal synoptic conditions. Then, the research task is to observe in detail how microwave fields are affected by the causal dielectric structure, to learn how to deduce field behavior from the structure, and hence from predictions based on synoptic meteorology. (A ray tracer modification [15] for computing propagation in blobby, stratified atmospheres is under study.) The current problem is still one of acquiring suitably complete observations; the role here of the instrumented aircraft is again obvious.

THE FLYING LABORATORY SYSTEM

Under the energetic and imaginative direction of Martin Katzin, NRL instrumented an aircraft in 1950 for general airborne radio and micrometeorological research.



Fig. 1—The flying laboratory. The nacelles under the wings contain steerable 47" dish radar antennas, three with variable polarization. The 15' retractable mast carries a Texas refractometer, the NRL ultraviolet humidimeter and other meteorological probes of slower response.

The flying laboratory (Fig. 1) provides mobile equipment capable of making quantitative measurements of radio and meteorological phenomena in the actual airborne geometry and geographical circumstance of a military problem [16]. Systems on four frequencies with steerable antennas and variable polarization were incorporated so that measurements could be over a wide range of frequencies simultaneously, with fixed conditions of atmosphere and sea- or ground-surface. All electronics were built to airborne specifications and stabilized as in a laboratory. The basic wiring was constructed along plug-board lines to permit rapid shuffling or substitution of gear as the experimental program

changed. The basic experimental records are photographic. The over-all system includes automatic film readers [17] which feed directly into computing machinery printing out signal statistics in the form of amplitude distributions and auto- and cross-correlations [18].

It is considered important that this film-reading and computer instrumentation accompanies the flying laboratory on its major field trips. With these equipments available in the field, day-to-day flight plans, measurement techniques, and analysis procedures may be altered or prescribed on the basis of reduced data from the flight of a previous day. This approach not only guarantees that the intended data has actually been acquired but also that the program is proceeding in a fruitful direction.

The flying laboratory is at all times available to the scientific personnel and is under their practical control. The pilots and crew have long-term assignments to the aircraft and all become familiar with its mission and idiosyncrasies. These important advantages are unfortunately not always the rule with aircraft assigned to research and development.

CLUTTER MEASUREMENTS WITH THE FLYING LABORATORY

For clutter measurements, the flying laboratory is self-contained, allowing simultaneous measurements at four radar frequencies (400, 1250, 3300, 9375 mc). Transmissions are made on both horizontal and vertical polarization and both polarizations are received simultaneously. A typical flight plan includes pulse-by-pulse measurements with the above frequencies and polarizations, pulse lengths being varied from 0.25 to 5.0 μ sec, and angles of incidence ranging from grazing to normal. It is considered important to vary these experimental parameters frequently during a single flight so that the resulting changes in clutter behavior reflect parameter changes rather than changes in target conditions.

Major full-scale sea clutter measurements have been staged at Bermuda where open-sea conditions occur a few miles off the beach. These programs, aimed at describing sea clutter as a function of radar and geometric variables and of oceanographic conditions, had the following results:

- 1) Especially for near-grazing incidence and for small illuminated area, clutter tends to become "spiky," *i.e.*, resolved into persistent target-like patches [19]. Spikiness presents a fundamental limitation to high resolution radar as a method for enhancing target-to-clutter ratios.

- 2) The polarization measurements indicate that there is little hope of reducing radar sea clutter through the use of polarization techniques (Fig. 2).

- 3) For the first time, correlations between physical measurements on the surface of the sea [20] and radar measurements [21] have been obtained. Fig. 3 shows correlation between wave slope and radar return, and a

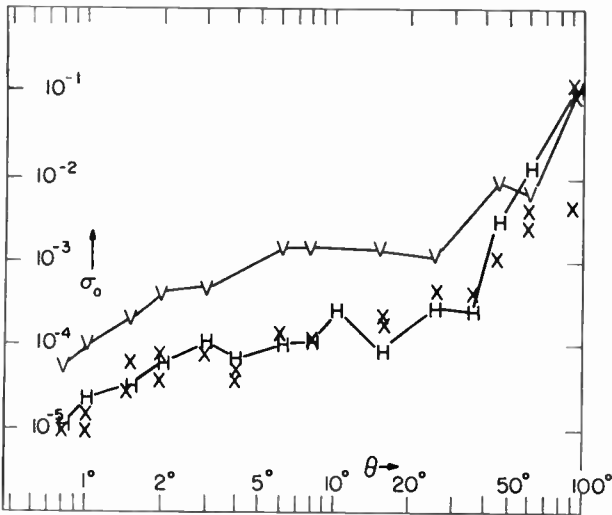


Fig. 2—Flying laboratory measurements of sea clutter echo at 1250 mc, January 20, 1956. σ_0 is radar area per unit horizontal area of sea surface, θ is grazing angle (angle radar ray makes with mean-sea-level plane). Symbols $V(H)$ indicate vertical (horizontal) radar polarization and X , cross-polarized component. At higher radar frequencies, the small- θ tendency of the V curve toward the H , X curve is more marked.

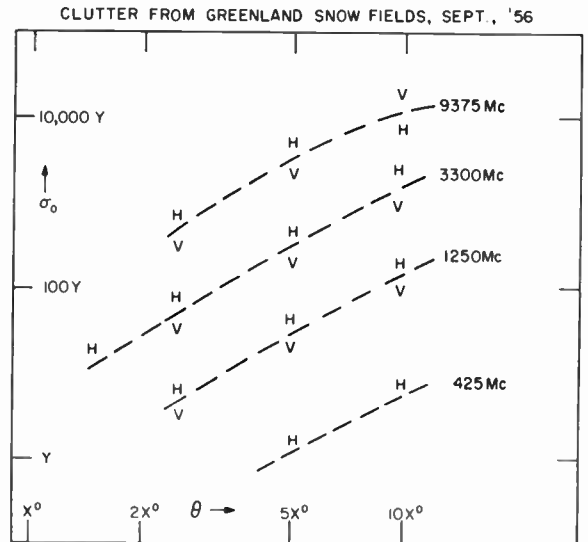


Fig. 4—Flying laboratory measurements of σ_0 , the median radar area per unit horizontal area, vs grazing angle θ , showing frequency and polarization trends. The symbols $V(H)$ denote vertical (horizontal) radar polarization, and symbol size approximates scatter of the data among different runs. The H 's lie systematically above the V 's, contrary to similar observations over other terrains.

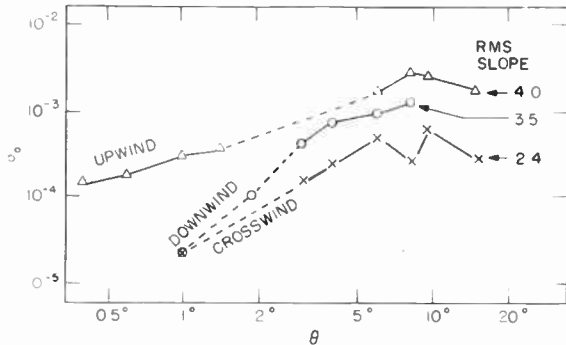


Fig. 3—Flying laboratory measurements of σ_0 vs θ for sea clutter, as function of wave slope. The slope-components, measured by Woods Hole Oceanographic Institution and given in relative units, are those visible to the radar and in the vertical plane through the radar.

directionality in wave shape consistent with the dependence of radar return on azimuth angle.

4) Measurements were obtained for angles of incidence ranging from near grazing to normal with a variety of sea states. These measurements, coupled with the polarization measurements, have been helpful toward establishing a radar/sea model [2], [22] which is required in a full understanding of the mechanism of radar sea clutter. Additional simultaneous airborne radar and oceanographic measurements are required for a full understanding of the sea surface as a clutter producer.

The flying laboratory has made ground clutter measurements over a wide variety of terrain. Fig. 4, from which the absolute scales have been removed for classification reasons, shows radar return from snow fields in Thule, Greenland [23]. Measurements of two non-homogeneous targets (Baltimore and farmland near Annapolis) have been made [24]. Several homogeneous targets (grassland and low trees in North Carolina,

desert and pecan trees in New Mexico, and large trees in New York and New Jersey) are included in a program now under way. This program is intended to catalog a number of targets as to amplitude and short-term statistics for radar map-matching information.

These ground clutter measurements, relatively easy and routine to perform with the flying laboratory, would be programs of considerable extent without such a vehicle.

THE FLYING LABORATORY IN PROPAGATION WEATHER RESEARCH

Under ONR contract, the University of Florida set up ground transmitters at Cape Canaveral and receivers at Nassau, the basic equipment being L-band and (later) X-band pulse radar units. The aim of the first flights of the flying laboratory in this 262-nautical mile link was to determine whether ground-to-ground scatter signal levels would correlate with features of refractometer sections of the first 10,000 feet of atmosphere, as suspected from earlier airborne measurements in trade-wind conditions [25]. It was felt that to realize this aim refractometer coverage would be required around-the-clock to detect immediately changes in either meteorological structure or radio field level and to negate chances of happenstance correlations that might occur with less coverage. Accordingly, cooperating aircraft were borrowed from several service agencies for round-the-clock airborne refractometer flights. To complete the meteorological coverage, arrangements were made for six hourly radiosonde balloon ascents to be made during the test period from nearby military installations. Finally, to assist in the understanding of the data as it was obtained, radio and meteorological

experts from a number of agencies were asked to assist. Thus, a program too large for NRL became instrumented and capably manned for the test period. A rewarding atmosphere of cooperation and idea-sharing was maintained at all times.

The concentrated effort of this week-long experiment in December, 1956, resulted in the following knowledge of propagation weather. All of the 105 full-length refractometer soundings taken during the test period showed a tradewind inversion but otherwise amounted only to random samples of this inversion (Fig. 5). The ground-to-ground hourly-median signal, otherwise steady, rose some 10 db for 18 hours. This rise correlated with nothing in the refractometer sections. At the suggestion of a cooperating physicist, wind-shear across the inversion at midpath was examined. Using available balloon-drift data, evidence of high shear implied high signal with odds three to one against chance. The similarity of all refractometer sections, and advice from cooperating experts, led to revision of the flying laboratory's flight plan from the refractometer routine to making frequent measurements of radio field vs altitude near the receiver terminal. These measurements consistently showed a 10-db height-gain maximum at inversion altitude. Above and below inversion altitude, the signal faded in the Rayleigh distributed manner characteristic of a scatter signal, while at the inversion altitude, this fading was underlain by the slow, deep fading of a duct-propagated signal.

The refractometer sections indicated an irregular elevated duct at altitude 4000–7000 feet at the tradewind inversion, with relatively featureless, smooth air above extending into the "common volume" which began at about 10,000 feet. The indicated mechanism for signal propagation (Fig. 6) was that the mixing wet and dry airs at the inversion formed strong blobs which scattered energy into the duct; the energy then proceeded over the horizon in the duct-propagated manner indicated by the height-gain sections, and other blobs scattered this energy to the ground receiver. The mechanism of shear in creating scatterers or enhancing the ducting is unclear, as is the reason that the pulse shape as received at Nassau and in the flying laboratory indicated two-path multipath transmission [10].

In an effort to learn more about the mechanism of shear and to investigate further the phenomenon of height-gain, a later experiment [26] was carried out in April–May, 1957. Again a number of aircraft and experts were borrowed from several different agencies.

The main aims of this test were negated by heavy thunderstorms during most of the two-week test period. The corresponding scatter signals were erratic, arrived from a wide range of azimuths, and exhibited multipath delays of up to 5 μ sec. Calm areas between storm centers were often observed (on airborne radars) to produce trapping.

The X-band pulse signal was normally found at about 4 db over set noise (as predicted) and was recordable

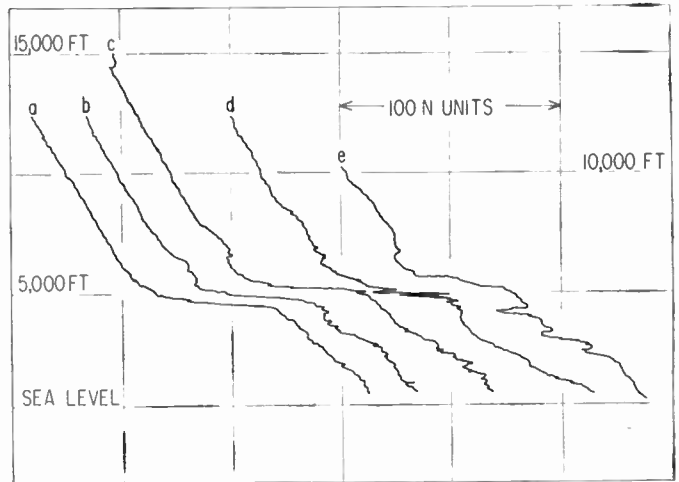


Fig. 5—Refractometer sections for 0128–0423, December 7, 1956. Taken during climbs by a cooperating aircraft from the Naval Air Development Unit. Section a) Patrick Air Force Base to 100 fathom line, b) 100 fathom line to Gulf Stream axis, c) Gulf Stream axis to Grand Bahama Island, d) Grand Bahama Island to Berry Islands, e) Berry Islands to Nassau. The 5-N-unit Bauer layer at 15,000 ft was generally present, as was the increasing complexity of low-level structure toward Nassau. Violent refractive index changes at about 5000 ft indicate "blobs" of mixing wet and dry airs. (Curves are traced from original X-Y recordings.)

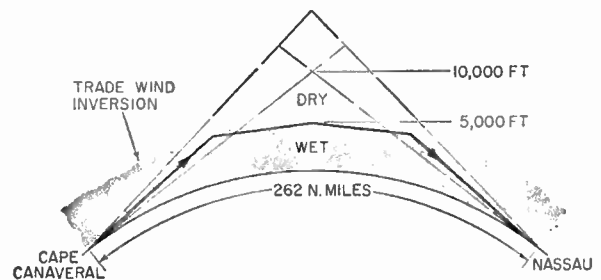


Fig. 6—Indicated propagation mechanism for ground-to-ground scatter propagation in tradewind conditions. Energy is scattered into and out of the elevated duct, formed by the trade-wind inversion, in the manner suggested by the heavily drawn ray.

at Nassau for only about three hours during a condition of elevated rain when the signal level increased some 30 db. Airborne measurements of the X-band signal at this time showed a nearly equal signal level over 50° of azimuth (both transmitting and receiving antennas had 1.25° bandwidths). It is suggested that the elevated rain, whose ragged lower level was about 7000–8000 feet, formed good reflectors which allowed the signal to be propagated by multiple bounces to Nassau and to the airborne receiver.

Lack of surface refractometer and storm radar coverage makes the stormy weather signal behavior difficult to interpret in detail. Nevertheless, the storms provided washed-out atmospheres (at least locally), and the lack of height-gain during the storms provided an interesting comparison with height-gain measurements on the final day, when trade-wind conditions were reasserted [Figs. 7(a), (b)].

Hindsight on these two test periods has led to the following conclusions: In a scientific propagation weather research program aimed at the ability to pre-

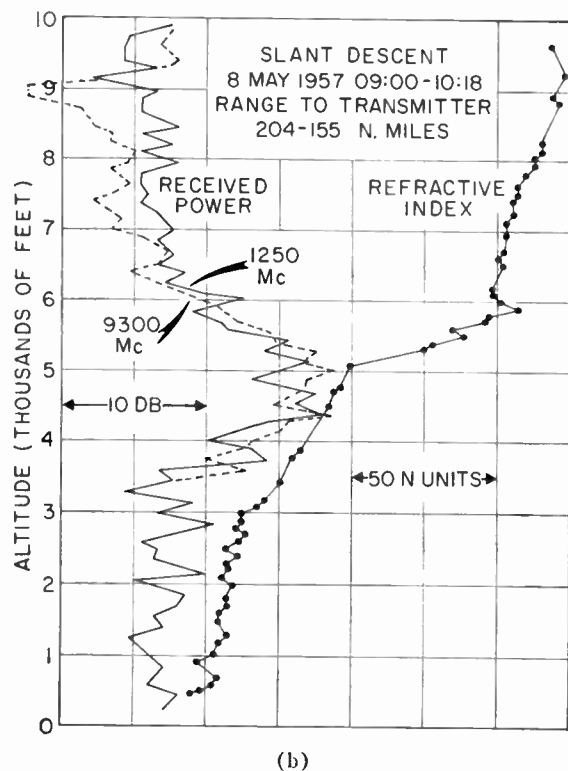
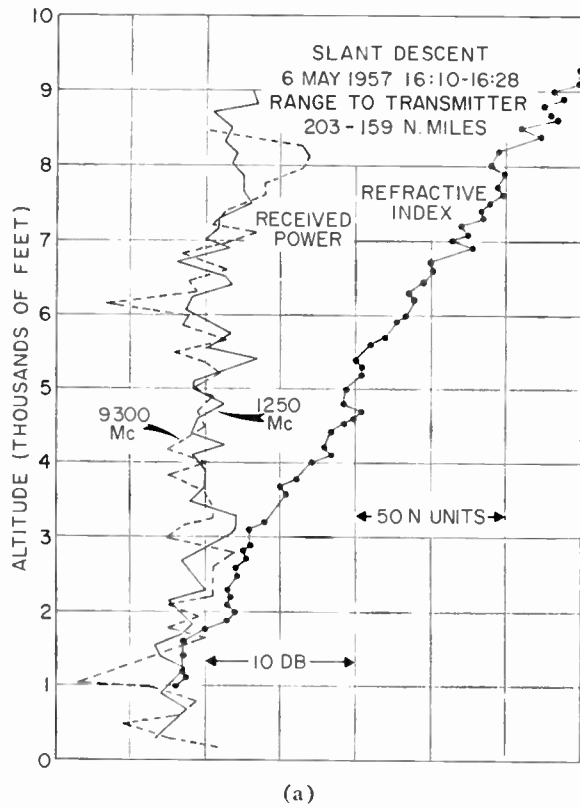


Fig. 7—Height gain and refractometer sections during storms (a) and during tradewind conditions (b). Refractive index decreases to the right. Allow 0.17 db/nautical mile for increase of signal as the transmitter is approached in the slant descent to obtain height gain along a vertical line.

dict, both airborne and ground measurements of signal and refractivity are necessary. Propagation weather is a branch of geophysics, and real research progress will be made by concentrated research on the geophysics, in cooperation with cloud physicists and others looking at the same atmospheric phenomena from ostensibly different viewpoints.

FUTURE RESEARCH PROGRAM

The flying laboratory has limited range and altitude capabilities, and its basic electronics are obsolescent. A new aircraft is being instrumented for continuing radar and propagation weather research.

The new aircraft, a WV-2, will provide a considerable improvement in range, altitude, and speed. All frequencies and time bases will be generated from a super-stable crystal to eliminate tuning difficulties and to make narrow-band measurements possible. Four flexibly modulatable klystron amplifiers will transmit from a stabilized antenna system. Data recordings will be on multi-channel magnetic tape, and a multichannel fast-response galvanometer system will provide immediately visible records as an improvement on the film records. An AN/APN-66 Doppler Navigator will provide local wind (and hence wind-shear) measurements and permit accurate navigation into the line-of-sight field of a narrow-beam antenna for obtaining free space field levels. A refractometer with at least two exposed cavities, a fast vortex thermometer, and an NRL ultraviolet humidimeter [4] will be the initial micro-meteorological probes.

The first research flights of the WV-2 will be to measure refractive structure of a probable elevated duct over the equatorial Atlantic; possibilities of such ducts for long-range microwave communications are indicated by observation of the previously mentioned height-gain maximum at trade wind inversion altitudes. The first field-measuring flights will be against transmissions and meteorological measurements by the Navy Electronics Laboratory at San Diego. The aim will be to investigate the seaward variation of the meteorological conditions (generally, the subsidence inversion), and its consequences on microwave signal level and channel capacity.

Equipping and flying the new aircraft will require all our personnel, and we will not be able to man ground transmitters necessary for exploratory one-way propagation measurements. To realize the potentialities of the new aircraft for propagation weather research, we have, therefore, proposed a facility for conducting radio meteorological research in a progressive, planned fashion and to serve as a permanent base for cooperative meteorological and cloud physics investigations. The proposed site (southeastward from Cape Canaveral) lies in a region having uniform, synoptically understandable weather regimes, well monitored by conventional

ground weather stations. Local flight rules and the presence of airbases will permit frequent, planned research flights. The proposed ground hardware will consist of high-power, flexibly modulatable klystron transmitters at several frequencies with side-by-side or common large antennas, and compatible receiving systems installed initially at two scatter-propagation ranges. To provide uniform low-level meteorology, understandable foregrounds, circumstances for mobile surface measurements, and especially, freedom from restrictive U. S. flight rules, the proposed island-to-island siting is considered a requirement.

Every effort should be made to measure micro-meteorological conditions from the surface with tower and balloon refractometry, storm radars, sky cameras, etc., abetted by studies of synoptic conditions and by a high level of airborne refractometry. The aim in the routine round-the-clock operation of the ground equipment should be to gain a growing insight into the measured causal meteorological structure, its synoptic origins, and its effects on microwave signal level, bandwidth, etc.

The basic equipment and the growing insight should serve as a foundation for special cooperative airborne meteorological and field measuring efforts, flexibly designed for pinning down specific causes of specific propagation phenomena. To this end, the significance of all data as they are being acquired should be understood. It would be important to attempt forecasts of micro-scale weather and its effects on signal levels. As intensive research identifies specific atmospheric causes of signal behavior, it will be possible to proceed from merely empirical to scientific, cause-and-effect, predictions, and hence, to the ability to predict for other geometries and weathers.

BIBLIOGRAPHY

Neither the text nor the following reference list is intended as a complete summary of the current state of radiometeorological knowledge. The monograph "Radiometeorology," by J. S. Marshall and W. E. Gordon, (*Meteorological Monographs*, vol. 3, No. 14, 1957), is recommended as an excellent general summary, particularly of the radar aspects.

- [1] L. H. Doherty, "Geometrical Optics and the Field at a Caustic With Applications to Radio Wave Propagation Between Aircraft," Cornell Univ. Res. Rep. EE 138; September 10, 1952.
- [2] M. Katzin, "On the mechanisms of radar sea clutter," *Proc. IRE*, vol. 45, pp. 44-54; January, 1957.
- [3] M. S. Wong, "Ray-Tracing Picture of Radio Wave Propagation in Arbitrary Atmosphere," Aircraft Radiation Laboratory, AF Tech. Rep. No. 6631; July, 1951.
- [4] J. M. Bologna, O. K. Larison, D. L. Randall, and D. L. Ringwalt, "An Airborne Lyman-Alpha Humidiometer," Naval Research Lab., Rep. No. 5180; August 12, 1958.
- [5] A. P. Deam, "Status Report on the Development of an Expendable Atmospheric Radio Refractometer," Elec. Eng. Res. Lab., University of Texas, Rep. No. 5-32; April 15, 1958.
- [6] L. J. Anderson, E. E. Gossard, "Prediction of the nocturnal duct and its effect on UHF," *Proc. IRE*, vol. 41, pp. 135-139; January, 1953.
- [7] M. Katzin, R. Bauchman and W. Binnian, "3- and 9-centimeter propagation in low ocean ducts," *Proc. IRE*, vol. 35, pp. 891-906; September, 1947.
- [8] L. J. Anderson, E. E. Gossard, "The effect of the oceanic duct on microwave propagation," *Trans. Amer. Geophys. Union*, vol. 34, pp. 695-700; October, 1953.
- [9] L. J. Anderson, E. E. Gossard, "Prediction of oceanic duct propagation from climatological data," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-3, pp. 163-167; October, 1955.
- [10] D. L. Ringwalt, W. S. Ament, and F. C. Macdonald, "Measurement of 1250-mc scatter propagation as function of meteorology," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 208-209; April, 1958.
- [11] S. Q. Duntley, A. R. Boileau, "Image transmission by the troposphere I," *J. Opt. Soc. Amer.*, vol. 47, pp. 499-506; June, 1957.
- [12] Unpublished reports from picket-aircraft radar operators.
- [13] J. R. Bauer, "The Suggested Role of Stratified Elevated Layers in Transhorizon Short-Wave Radio Propagation," MIT Lincoln Lab., Lexington, Mass., Tech. Rep. No. 124, Contract AF 19(122)-458; September 24, 1956.
- [14] B. J. Starkey, "Some aircraft measurements of beyond-the-horizon propagation phenomena at 91.3 mc/s," *Proc. IEE*, vol. 103-B, pp. 761-763; November, 1956.
- [15] W. S. Ament, "Modification of a Ray-Tracer for Monte Carlo Prediction of Multiply Scattered Radio Fields," Symp. Stat. Problems in Radio Wave Propagation, University of California, Los Angeles; June, 1958. To be published.
- [16] D. L. Ringwalt, "An airborne radar and wave propagation laboratory," 1955 IRE CONVENTION RECORD, pt. 1, pp. 82-85; 1955.
- [17] A. Shapiro, "Film reader measures radar echoes," *Electronics*; January, 1957.
- [18] F. C. Macdonald, A. Shapiro, "Statistical analysis equipment for propagation research," *Proc. Symp. Stat. Problems in Radio Wave Propagation*, University of California, Los Angeles; June, 1958.
- [19] F. C. Macdonald, "Characteristics of radar sea clutter. Part I. Persistent target-like echoes in sea clutter," NRL Rep. 4902; March 19, 1957.
- [20] H. G. Farmer, "Some Recent Observations of Sea Surface Elevation and Slope," Ref. No. 56-37, Woods Hole Oceanographic Inst.; June, 1956.
- [21] F. C. Macdonald, "The correlation of radar sea clutter on vertical and horizontal polarization with wave height and slope," 1956 IRE CONVENTION RECORD, pt. 1, pp. 29-32.
- [22] F. C. Macdonald, "Sea clutter at X and-L bands," *Symp. Rec., Advanced Technique in Radar Components and Systems for Battlefield Surveillance*, Eng. Res. Inst., University of Michigan, Ann Arbor; February, 1956 (Secret).
- [23] F. C. Macdonald and D. L. Ringwalt "Terrain clutter measurements in the far north," presented at the Third Annual Radar Symp. at the University of Michigan; February, 1957.
- [24] F. C. Macdonald, W. S. Ament, and D. L. Ringwalt, "Terrain Clutter Measurements," NRL Rep. 5057; January 21, 1958 (Confidential).
- [25] W. S. Ament and M. Katzin, "Signal fluctuations in long-range overwater propagation," *IRE TRANS. ON COMMUNICATIONS SYSTEMS*, vol. CS-4, pp. 118-122; March, 1956.
- [26] D. L. Ringwalt, W. S. Ament, and F. C. Macdonald, "Scatter Propagation in Thunderstorm Conditions," NRL Rep. of Progress, pp. 17-25; August, 1957.

Sporadic E at VHF in the USA*

R. M. DAVIS, JR.†, E. K. SMITH†, SENIOR MEMBER, IRE, AND C. D. ELLYETT†

Summary—An analysis is made of sporadic-E propagation observed on the Cedar Rapids to Sterling path during the four years, 1952–1955. VHF transmissions over this path at frequencies of 27.775 mc and 49.8 mc provided the Es data.

At this latitude, the summer months receive the preponderance of sporadic E. Diurnally, sporadic-E occurrence tends to favor the daytime and evening hours with peaks of incidence about 1000 and 1800. No variation with sunspot number has been discerned. As a rule, higher Es signal intensities are recorded at 28 than at 50 mc. Cumulative distributions of signal intensity are presented for the two frequencies.

A relationship is found between the frequency dependence of Es signal intensities ≥ -70 db, relative to inverse distance, and the distribution of fEs values at Washington, D. C. The relationship promises to be useful in the prediction of Es signal intensities on a worldwide basis.

An inverse correspondence is shown between sporadic-E occurrence and geomagnetic activity. The correspondence holds only over selected time intervals.

The frequency dependence of received power under sporadic-E conditions is different from that during normal scatter. The median frequency exponent is two or more times as large for sporadic E, and the exponents cover a much wider range of values. This is tentatively explained by considering the Es region to be composed of patches of intense nonuniform ionization, an hypothesis previously used to explain vertical-incidence data.

INTRODUCTION

ONE of the most distinctive features of VHF radio transmission is the occasional occurrence of sporadic-E propagation. When this takes place, the received power undergoes an increase of many times its normal value and the usual conditions of VHF radio propagation are greatly altered.

A substantial number of oblique-incidence sporadic-E data have been acquired by the National Bureau of Standards in carrying out a related research project. Early in 1951, this agency began an extensive experiment in the transmission of VHF radio waves by the mechanism of forward scatter.¹ Horizontally polarized CW transmissions were inaugurated over the path from Cedar Rapids, Iowa, to Sterling, Va., a distance of 1243 km. From the beginning of the program until June 30, 1958, transmissions at one or more frequencies were recorded continuously.

Early in the program it was recognized that the recordings incorporate valuable information on the incidence and characteristics of sporadic-E propagation. It is the aim of this paper to present an analysis of the Es

data recorded from 1951 through 1955.

Sporadic E gives rise to a type of propagation in which radio waves are returned from the E region of the ionosphere with relatively high reflection coefficients for periods ranging from a few minutes to several hours. At oblique incidence, the presence of a sporadic-E region is revealed by a large increase in the signal intensity of a transmitted wave. At vertical incidence, sporadic E is observed on ionograms as a trace at the approximate height of the E layer, but normally extending above the E-layer critical frequency. The top frequency attained by the Es trace is designated fEs.

The incidence of sporadic E on the Cedar Rapids to Sterling path is the first subject of investigation in the following sections. Next, an analysis is made of the signal intensities that characterize Es propagation. A relationship is sought between the signal intensities observed on the oblique path and the values of fEs recorded in the same geographical area. The possibility that sporadic-E incidence is related to geomagnetic activity is then explored. Finally, a comparison is made of the power received at two frequencies, and an explanation is offered for the frequency dependence that is found to exist during periods of sporadic E.

RECORDING OF SPORADIC-E DURATION AT VHF

The systematic recording of sporadic E was carried out under a system in which rhombic antennas with 25-wavelength legs and plane-wave gains of 18 db were employed at each terminal. The half-power beamwidth of the antennas was 6° in the horizontal plane, and the first lobe of the vertical antenna pattern was directed at the great circle midpoint of the path at a height of 110 km. The frequencies recorded most consistently on the Cedar Rapids to Sterling path were 27.775 mc and 49.8 mc. All the observations reported here were taken at these two frequencies.

Fig. 1 shows the typical appearance of the Esterline-Angus records during a period of sporadic E. The two left-hand records illustrate the effect of Es propagation on the signal received at 27.775 mc, while the corresponding effect at 49.8 mc is shown on the right. The top two records represent the type of recording in which the occurrence of sporadic E usually produced intensities above the limits of the recorder. In the lower charts, the same event is registered on a wide dynamic range recorder where the entire range of Es intensities falls within the chart limits. Wide dynamic range records were made only during the last ten months of the period covered by this report.

To study the incidence of sporadic E it was necessary to choose criteria that would distinguish it from other

* Original manuscript received by the IRE, December 9, 1958. This work was sponsored jointly by the U. S. Signal Corps (MIPR No. R-54-55-SC-91) and the Dept. of the Navy (Govt. Order NAonr-206-55).

† Central Radio Propagation Lab., Nat'l. Bur. Standards, Boulder, Colo.

¹ D. K. Bailey, R. Bateman, and R. C. Kirby, "Radio transmission at VHF by scattering and other processes in the lower ionosphere," Proc. IRE, vol. 43, pp. 1181–1230; October, 1955.

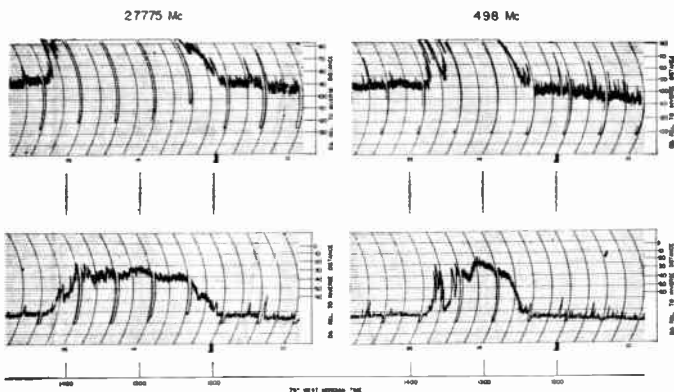


Fig. 1—An occurrence of sporadic E on records of the Cedar Rapids to Sterling transmission, July 2, 1955.

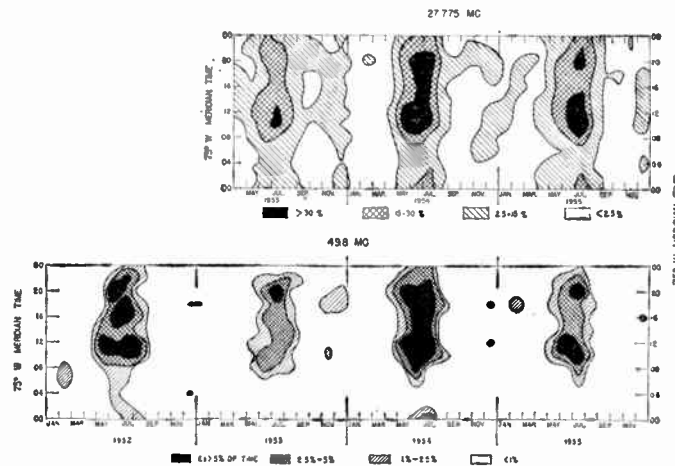


Fig. 2—Temporal variation of occurrence of sporadic E at 27.775 mc and 49.8 mc, Cedar Rapids to Sterling.

signal enhancements, principally those due to long-duration meteor echoes and atmospheric interference. The criteria used were:

- 1) The enhancement must last at least 12 minutes, during which time the signal level must be clearly distinguishable from the background scatter signal.
- 2) For at least one minute, the signal must maintain a level of -60 db or higher, relative to the inverse-distance value.

All the values of signal intensity mentioned in this paper refer to decibels relative to the inverse-distance value. This value is based on the assumption that the ionosphere and the ground at both path terminals act as perfectly reflecting planes and that the sky reflection takes place at 110 km. The received power of a signal subject to inverse-distance attenuation can be calculated from (7) of Bailey, Bateman, and Kirby.¹

On the basis of the above criteria, a list was made of all the occurrences of sporadic E between March 1951 and December 1955. The temporal distribution of Es incidence at both 49.8 mc and 27.775 mc is portrayed in Fig. 2. In this figure the occurrence of sporadic E at the two frequencies is presented by contours representing the percentage of time that Es was observed, as a function of month of the year and hour of the day. It will be noted that the 28-mc contours represent far higher values of probability than the 50-mc. For example, the highest contour for 28 mc represents a 30 per cent Es incidence compared to a maximum value of 5 per cent for 50 mc.

The outstanding feature of this illustration is the pronounced summer maximum in each year. At both frequencies, most of the total sporadic E of the year falls in the months May through August. In other months of the year, a certain amount of sporadic E, as much as a 15 per cent incidence, occurs at 27.775 mc. But at 49.8 mc, there is only a very light occurrence of Es in winter months and none at all around the equinoxes. Diurnally, the heaviest Es incidence occurs between the hours 0800 and 2200, 75°W Meridian time. (Local time at the path midpoint may be found by subtracting 0.6 hour from 75°W Meridian time.)

The varying numbers of observations incorporated in Fig. 2 imply a degree of statistical confidence that varies with hour and month. By making reasonable statistical assumptions it is possible to estimate confidence intervals for various observed percentages of time of Es. Rough but conservative estimates of the 95 per cent confidence limits are listed in Table 1. They apply to both transmission frequencies and are given for the percentage values of the contour lines of Fig. 2.

TABLE I
95 PER CENT CONFIDENCE INTERVALS OF PERCENTAGE OCCURRENCE OF SPORADIC E (VALUES APPLY TO BOTH 27.775 MC AND 49.8 MC)

Observed Percentage	95 Per Cent Confidence Interval
1 Per Cent	0-12 Per Cent
2.5 Per Cent	0-15 Per Cent
5 Per Cent	1-19 Per Cent
15 Per Cent	5-29 Per Cent
30 Per Cent	17-45 Per Cent

Average diurnal and seasonal curves of Es incidence are shown in the next two figures. In Fig. 3 the 49.8-mc curve is the diurnal average of four years of observations; the 27.775-mc curve is the average of two years. At both frequencies there are two diurnal peaks of Es incidence, one between 0900 and 1200 hours, the other, perhaps smaller, between 1700 and 2000, local time at path midpoint. Only a small amount of sporadic E occurs in this geographical area from 0000 to 0800 hours.

An average curve of seasonal variation of Es occurrence throughout the observing program is given in Fig. 4. The data for each month were normalized to take account of the number of hours of equipment failure on the circuit. Again, the summer maximum appears prominently in the curves. There is also some evidence of a secondary, relatively minor annual maximum in the winter months with the greatest effect in December.

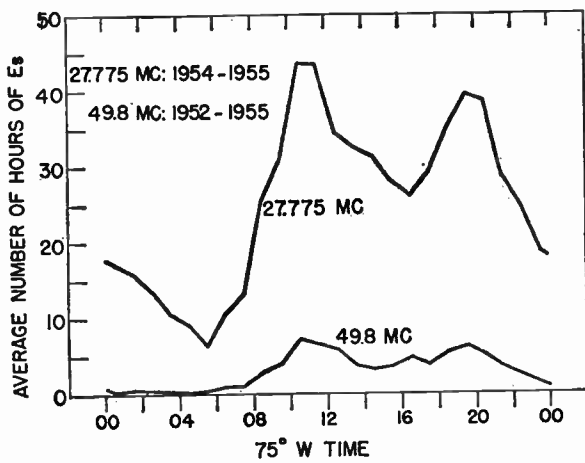


Fig. 3—Average diurnal variation of sporadic-E occurrence at 27.775 mc and 49.8 mc, Cedar Rapids to Sterling.

The data contained in the last two figures are summed over whole years in the following table:

TABLE II
TOTAL HOURS OF Es

	1952	1953	1954	1955
49.8 mc	86.1	53.7	104.3	57.4
27.775 mc			616.9	604.0

This table bears out the conclusions of Bailey, Bateman and Kirby,¹ Dyce,² and Smith³, that incidence of sporadic E does not show any solar cycle dependence. The same authors have presented essentially the same diurnal and seasonal variations of Es occurrence as are shown in Figs. 3 and 4.

SPORADIC-E SIGNAL INTENSITIES AT VHF

A complete description of the effect of sporadic-E propagation at VHF should include a study of the signal intensities associated with sporadic E. During the interval May, 1955 through February, 1956, for which wide dynamic range recordings are available, it has been possible to make such an analysis.

In this ten-month period, Es data were extracted from the records in a special way. All occurrences of sporadic-E enhancement were divided into five-minute intervals. In each interval the median value of signal intensity was scaled from the record. To insure that all the intervals used were governed by sporadic E, only those intervals were retained in the study for which the median intensity was equal to or greater than -70 db relative to the inverse-distance value. This criterion had the effect of retaining some of the periods designated as Es under the old requirements and eliminating others. As a result, the total durations of Es determined by the

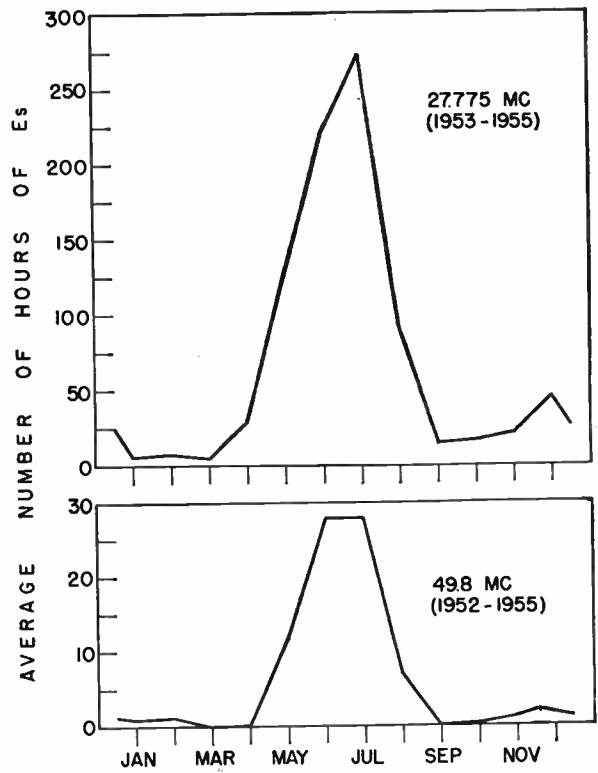


Fig. 4—Average seasonal variation of sporadic-E incidence, Cedar Rapids to Sterling.

two methods are not identical. In the five-minute period analysis, the scaled Es values of signal intensity range from -70 db to -20 db, so that they fall into eleven groups each 5 db wide.

The signal intensities read from the records in each five-minute period of sporadic E are treated statistically in Fig. 5. Here are shown the cumulative distributions of signal intensity for the two frequencies 27.775 mc and 49.8 mc. The figure indicates the distribution for all hours of the day (solid lines), the daylight hours 06-18 (broken lines), and the nighttime hours 18-00 and 00-06 (dotted lines), 75°W time.

Analysis of the cumulative distributions of signal intensities for the periods when sporadic E propagation is controlling indicates that these distributions are approximately log-normal. The largest departure from the log-normal form is exhibited by the daytime curve for 28 mc. This may be due to the effects of absorption.

Several important characteristics of sporadic-E occurrence are seen in Fig. 5. First, the enhanced signal intensities due to sporadic E occur more often at 28 mc than at 50 mc, though this effect is far less noticeable at higher intensities. Second, sporadic E produces enhanced signal intensities a greater proportion of the time in the daylight hours than at night. Nevertheless signal intensities as high as -20 db relative to inverse distance were recorded only in the nighttime period.

It should be emphasized that the values of Es duration and signal intensity relative to inverse distance given in this report apply to narrow-beam, high-gain

² R. Dyce, "VHF auroral and sporadic-E propagation from Cedar Rapids, Iowa, to Ithaca, N. Y.," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-3, pp. 76-80; October, 1955.

³ E. K. Smith, Jr., "Worldwide Occurrence of Sporadic E," Nat. Bur. Standards, Boulder, Colo., NBS Circular 582; March 15, 1957.

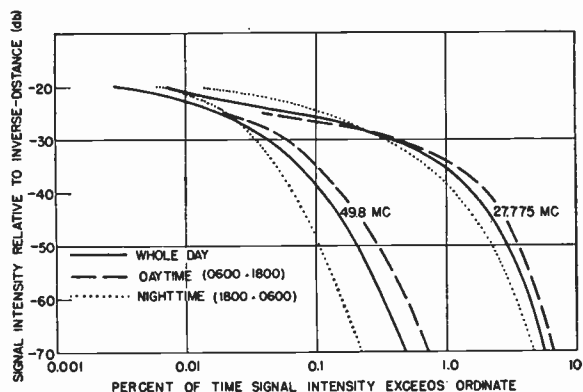


Fig. 5—Cumulative distribution of signal intensities enhanced by sporadic E, Cedar Rapids to Sterling, May, 1955 through February, 1956.

antennas. For broad-beam antennas, Es duration for a given reflection coefficient and intensity relative to inverse distance would both be greater.

THE RELATIONSHIP OF ES SIGNAL INTENSITIES TO fES

An area of vital interest in the field of VHF communications is the problem of estimating the signal intensities to be expected from sporadic-E propagation. The possibility of applying fEs data to the prediction of Es signal intensities at VHF will now be explored.

If fEs is to be used as the basis for predicting Es signal intensities, it must be shown that its variation in time is similar to the variation of oblique-incidence sporadic-E enhancements. To make such a test it may be assumed that sporadic-E propagation obeys the secant law. Experimental evidence for this has been reported by Wright.⁴ A value of 4.53 is found for the secant ϕ factor applicable to the Cedar Rapids to Sterling path, on the assumption of a sporadic-E layer height of 110 km. When the transmission frequencies 49.8 mc and 27.775 mc are divided by this figure, the values 11.0 and 6.1 mc, respectively, are obtained. These may be termed the "equivalent vertical-incidence" frequencies corresponding to the two transmission frequencies.

In Fig. 6, the temporal distribution of sporadic E observed at 27.775 mc on the Cedar Rapids to Sterling path is compared with the distribution of fEs ≥ 6.1 mc at Washington, D. C., near the eastern end of the path. Three of the contours represent the same incidences on the two maps, namely 30 per cent, 15 per cent, and 5 per cent. The outer contour defines a 1 per cent probability at oblique incidence against a probability of 2.5 per cent at vertical incidence. (The same confidence intervals given in Table I apply to the oblique-incidence contours of Fig. 6.) Although the distance of Washington, D. C., from the path midpoint is more than 400 miles, the figure indicates that there is a reasonably good correlation between the two sets of data. When sporadic-E occurrence at 49.8 mc is compared with vertical-incidence Es at values of 11.0 mc and above, the statisti-

⁴ J. W. Wright, private communication, January, 1958.

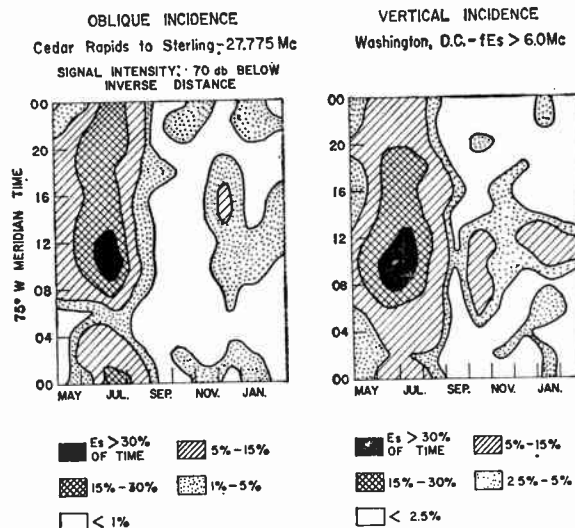


Fig. 6—Comparison of the temporal variation of sporadic E observed at 27.775 mc on the Cedar Rapids to Sterling path with the occurrence of fEs ≥ 6.1 mc at Washington.

cal association is weaker. This might be expected because there is only a limited sample of fEs data at values ≥ 11.0 mc.

If oblique and vertical-incidence sporadic E are assumed to be related through the secant ϕ factor, the cumulative distribution of fEs at Washington can be compared with the cumulative distribution of sporadic E on the Cedar Rapids to Sterling path. This is done in Fig. 7 for seven levels of signal intensity from -70 db to -20 db relative to inverse distance. At each signal level, two experimental points are available from the oblique Es results, after conversion to equivalent vertical-incidence frequencies of 6.1 and 11.0 mc. Straight lines have been drawn between the known points, in accordance with the relationship introduced by Phillips, $\log(\text{probability of fEs}) = a + bf$, where a and b are constants.⁵

It is clear from Fig. 7 that occurrence of signal intensities ≥ -70 db varies with frequency in the same way as the occurrence of vertical-incidence Es at Washington. From this correspondence it should be possible to make use of the distribution of fEs values at Washington to predict the probability of receiving Cedar Rapids to Sterling transmissions at intensities of -70 db and above. The relationship applies only to paths of roughly the same length in the same geographical area.

When separate graphs of the type of Fig. 7 are constructed for daytime and nighttime hours, it is found that the distribution of fEs is still a close parallel to the distribution of intensities ≥ -70 db.

The slopes of the curves of Fig. 7 indicate that the signal intensity in db is a fairly linear function of $f \cos \phi$ for signal intensities up to -40 db, but not for higher

⁵ M. L. Phillips, "Variations in sporadic-E ionization observed at Washington, D. C.," *Trans. Am. Geophys. Union*, vol. 28, pp. 71-78; February, 1947.

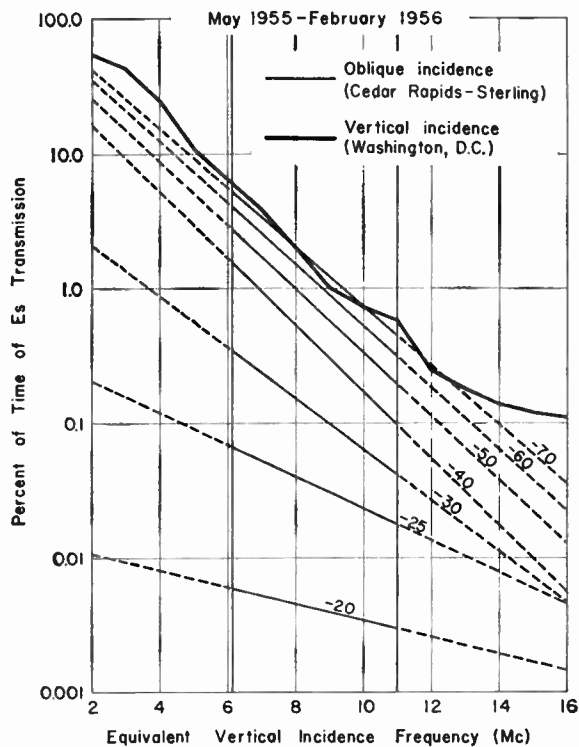


Fig. 7—Comparison of the cumulative distribution of fE_s at Washington with the distribution of $f \cos \phi$ at various signal intensities on the Cedar Rapids to Sterling path.

intensities. Fig. 7 displays a striking difference from the frequency dependence exhibited by Japanese data.⁶ On the basis of the Yamagawa data from this experiment Smith⁷ concluded that the relation $\log(\text{probability of } E_s) = a + bf$ could be used for oblique-incidence E_s predictions with the constant a set equal to zero. This relation is at variance with the data of Fig. 7. The slopes of the curves in this figure decrease with rising signal intensity, which is opposite to the tendency shown by the Japanese data. Caution should therefore be exercised in applying the formula $\log P_1 / \log P_2 = f_1 / f_2$ to the prediction of E_s probabilities.

It would seem that a method of the kind suggested here for relating oblique and vertical-incidence E_s would be useful in predicting E_s signal intensities on a worldwide basis. For such a project it would be necessary to obtain appropriate oblique-incidence E_s data, which could then be used in conjunction with existing fE_s observations.

DEPENDENCE OF SPORADIC E ON MAGNETIC ACTIVITY

In a study of sporadic E at both vertical and oblique incidence, Smith found an inverse correspondence between magnetic activity and the occurrence of sporadic

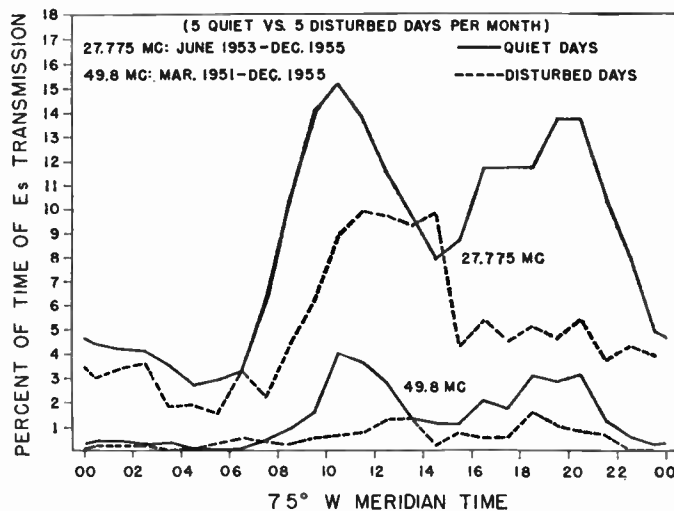


Fig. 8—Probability of sporadic-E occurrence on magnetically quiet and disturbed days, Cedar Rapids to Sterling.

E at temperate latitudes.³ The oblique-incidence data from the Cedar Rapids to Sterling transmission have been examined for a similar relationship.

Fig. 8 is a graph of the diurnal variation of all the sporadic E recorded throughout the observing program on quiet days (solid lines) and disturbed days (broken lines). By "quiet" days are meant the five days per month selected for their low level of magnetic activity; the "disturbed" days are the five most magnetically active days per month. In Fig. 8 the 28-mc curves incorporate $2\frac{1}{2}$ years of observations, the 50-mc curves nearly five years. The figure shows that there is more E_s on quiet than on disturbed days at most of the hours of the day.

In Fig. 9 the influence of magnetic activity on sporadic E is shown for individual years and is also portrayed for individual months. These histograms represent per cent occurrence of E_s on quiet days, intermediate days, and disturbed days, in this order from left to right. All the sporadic E that occurred in a given year is lumped together according to magnetic classification in the right-hand panels. On the left, the distribution of this E_s occurrence among the months of the year is displayed.

When data for a whole year are considered, sporadic E shows a clear preference for quiet days in 1952, 1953, and 1954. In 1951 and 1955, E_s incidence at 49.8 mc did not decrease progressively from quiet to disturbed days. The tendency is insignificant at 27.775 mc in 1955. In the monthly histograms the relationship of magnetic activity to sporadic E is not at all dependable. In February, 1952, June, 1954, and July, 1955, the tendency was reversed, more E_s falling on disturbed days than on quiet days. Thus it is evident that the inverse correspondence of E_s incidence with magnetic activity holds only over a long term. Due to the random nature of sporadic E, this behavior may not be apparent over a short period such as one month. Whether the solar

⁶ T. Kono, Y. Uesugi, M. Hirai, and G. Abe, "Study of long distance propagation of VHF waves by sporadic E ionization," *J. Radio Res. Labs.*, vol. 1, pp. 1-10; March, 1954.

⁷ Smith, *op. cit.*, p. 219.

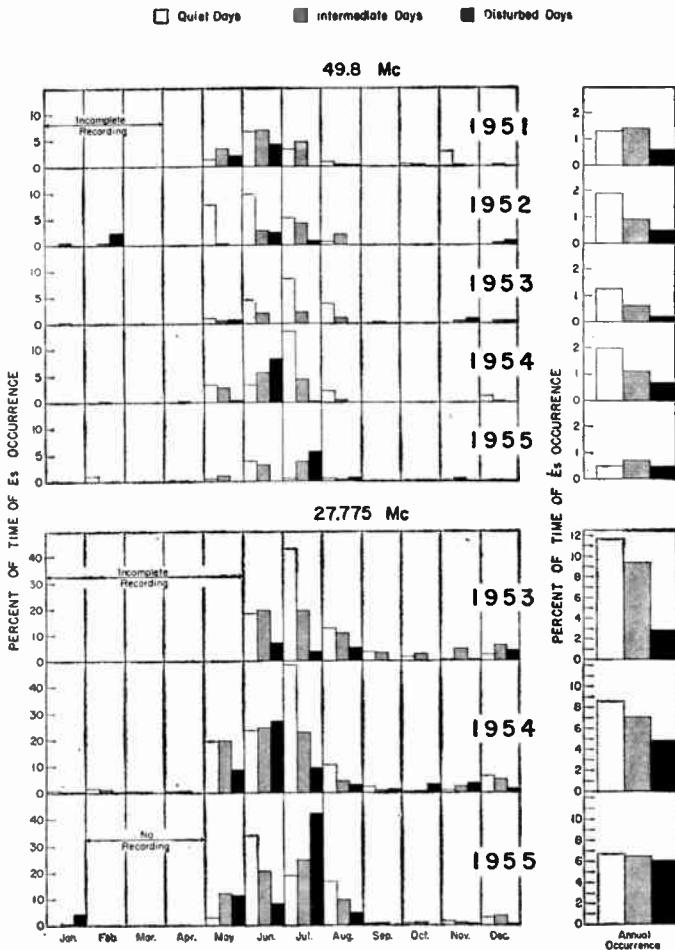


Fig. 9—Magnetic activity dependence of monthly and annual probabilities of Es occurrence, Cedar Rapids to Sterling.

cycle, which reached a minimum in 1954, exerts any influence on the magnetic activity dependence of sporadic E cannot be determined from the sample of data in Fig. 9.

FREQUENCY DEPENDENCE OF RECEIVED POWER DURING Es PROPAGATION

According to theories of scattering from ionospheric irregularities⁸⁻¹⁰ received power P_r is proportional to the scattering cross section σ , which in turn is inversely proportional to some power of the transmission frequency. When signal intensities received simultaneously at 27.775 mc and 49.8 mc are compared during periods of normal scatter propagation, the median value of the frequency exponent is about 5. This value is obtained after adjustment for similar system conditions at the two frequencies and applies to equal apertures at the

⁸ D. K. Bailey, R. Bateman, L. V. Berkner, H. G. Booker, G. F. Montgomery, E. M. Purcell, W. W. Salisbury, and J. B. Wiesner, "A new kind of radio propagation at very high frequencies observable over long distances," *Phys. Rev.*, vol. 86, pp. 141-145; April 15, 1952.

⁹ F. Villars and V. F. Weisskopf, "On the scattering of radio waves by turbulent fluctuations of the atmosphere," *Proc. IRE*, vol. 43, pp. 1232-1239; October, 1955.

¹⁰ A. D. Wheelon, "Radio frequency and scattering angle dependence of ionospheric scatter propagation at VHF," *J. Geophys. Res.*, vol. 62, pp. 93-112; March, 1957.

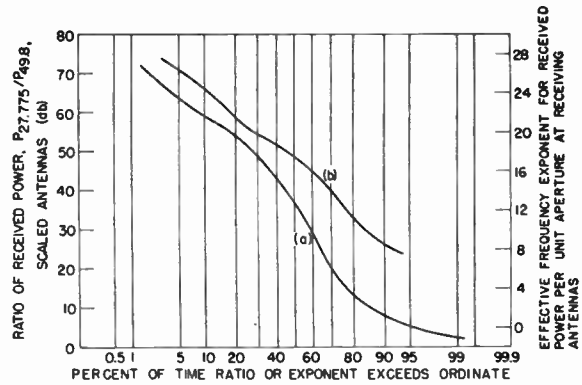


Fig. 10—Frequency dependence of received power during periods of sporadic E at 27.775 mc and 49.8 mc, Cedar Rapids to Sterling; (a) 134 five-minute periods, June, 1955; (b) 34 five-minute periods, December, 1955 and January, 1956.

receiving antennas. (A more complete explanation of frequency dependence will be found in Bailey, Bateman, and Kirby.¹¹)

If the same type of comparison of signal intensities is made when propagation by sporadic E is clearly predominant at both frequencies, a very much higher value of frequency exponent is usually encountered. Fig. 10 presents the cumulative distribution of Es power ratios during two observing periods. The first period, in June, 1955, contained 134 five-minute intervals of simultaneous sporadic E. The second period, in December, 1955, and January, 1956, yielded only 34 five-minute intervals. In June the ratios of signal intensity, calculated for scaled antennas at transmitter and receiver, range from 2 to 72 db. Corresponding frequency exponents, expressed for unit aperture at the receiving antenna, range from a slightly minus value to about 26. The median value of the exponent is 12. For the smaller wintertime sample, the range of frequency exponents is 8 to 27 and the median is 17.

Because of the greatly increased frequency exponents found during sporadic E and because of the wide range of values covered by these exponents, it seems likely that the primary mechanism responsible for Es is basically different from that involved in scatter propagation. If this is true, then a direct comparison of the frequency exponents under the two conditions is fruitless, at least in the present state of the theory of sporadic-E propagation.

The wide variation in the values of frequency exponent when sporadic E becomes the dominant mode has been an intriguing feature of this analysis. A simple qualitative explanation of the variability is offered without any present certainty of its correctness. The explanation makes use of a model of the sporadic-E layer similar to that employed by Rawer¹² and Thomas.¹³

¹¹ Bailey, Bateman, and Kirby, *op. cit.*, pp. 1196-1199.

¹² K. Rawer, "Die Ionosphaere," Noordhoff, Groningen, The Netherlands; 1953.

¹³ J. A. Thomas, "Sporadic E at Brisbane," *Aust. J. Phys.*, vol. 9, pp. 228-246; June, 1956.

Let it first be assumed that the sporadic-E region contains at least small patches of sufficient electron concentration to refract at oblique incidence the frequencies 27.775 mc and 49.8 mc. The geometry of the path is pictured in Fig. 11. The vertical-incidence frequencies corresponding to 27.775 mc and 49.8 mc, namely 6.1 mc and 11.0 mc, are of reasonable occurrence on h'f ionosonde records. This strongly indicates that the oblique transmissions on this path are indeed subject to sporadic-E propagation from time to time.

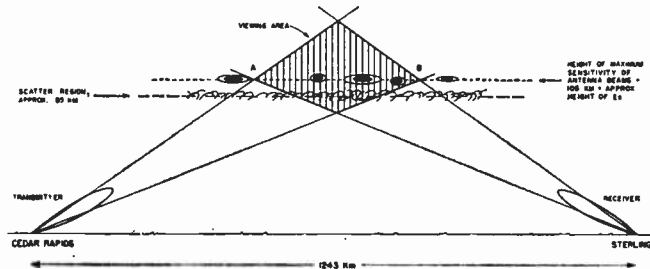


Fig. 11—Path geometry used in explanation of frequency dependence of received power.

It should be mentioned, parenthetically, that the antennas on the Cedar Rapids to Sterling path are not directed at an angle to receive maximum signal from the scatter region at approximately 85 km, but are beamed to center in the Es region at 110 km. This should increase the extent of time during which Es controls the signal. Strictly, during such periods, 85-km scatter will still be contributing to the received signal, but it is so weak by comparison that it can be ignored.

The sporadic E may be visualized either as a single "cloud" with uneven internal electron structure or as a number of discrete patches. Such a sporadic-E region would be of the constant-height type (Esc).¹³ In either case it is required that the refracting areas be small compared with the first Fresnel zone to account for a received power 20 db below that of a continuous refracting region.

It seems logical to suppose that the electron concentration rises to a maximum at some interior location in the blob of ionization. Moreover, the patches in the area *AB* of Fig. 11 must change with time, either by growth or decay, or by a drift of ionization across the region. Clear evidence for the existence of both phenomena has been found.¹⁴ Either process could give rise to the variations observed in the present results.

As the cloud, with its electron concentration increasing toward the center, materializes in the region *AB*, first the 28- and then the 50-mc signal can be returned. As the cloud leaves the region first the 50-mc signal penetrates and is lost, followed later by the 28-mc signal. Thus the 50-mc signal should be wholly contained

in time within the occurrence of the 28-mc signal. Such proves to be the case. Also the 28-mc signal should commonly occur without the occurrence of any 50-mc signal. This too is frequently observed, and it merely implies that the central electron concentration is insufficient to refract the higher frequency.

Under conditions of Es propagation at 28 mc it must be supposed that there is a certain portion of *AB* that is able to refract 28 mc but not 50 mc. However, this portion can vary widely. Consequently, the mean 28-mc signal intensity can be greater than the 50-mc signal in any ratio greater than unity. The frequency exponent, therefore, can assume any value greater than zero, and occasionally it should rise to high positive values.

The small negative exponents that have been observed may be accounted for as an effect of nondeviative absorption, which decreases with rising frequency. Though the effect is small at these frequencies, it may be sufficient on occasion to cause the 50-mc received power to exceed the 28-mc power.

It is possible that the frequency dependence results can also be explained on the basis of a more general theory based on turbulence at E-region heights.¹⁵ This theory is not available in sufficient detail, however, to allow comparison to be made with the present data.

CONCLUSIONS

From data taken during an extensive experiment in ionospheric forward scatter over the Cedar Rapids to Sterling path, it is found that the occurrence of VHF sporadic E varies diurnally, seasonally, and with frequency. No variation with sunspot number has been established. Sporadic E occurs more often at 28 mc than at 50 mc. At both frequencies there are diurnal maxima between 09 and 12 hours and between 17 and 20 hours, local time at path midpoint. Seasonally, the months May through August receive the heaviest incidence of sporadic E. A peak occurs in June-July.

The cumulative distributions of signal intensities due to Es propagation are different at the two frequencies 27.775 mc and 49.8 mc. Sporadic E is characterized by higher signal intensities at 28 mc than at 50 mc, but the effect is almost lost at the highest signal intensities.

VHF signal intensities during sporadic E can be predicted by the use of vertical-incidence Es data. For this purpose the two transmission frequencies are converted to equivalent vertical-incidence frequencies, $f \cos \phi$, of 6.1 and 11.0 mc. From the known probabilities of each signal intensity at these two frequencies, a family of curves of probabilities vs $f \cos \phi$ can be drawn. It is found that the cumulative distribution of fEs at Washington closely matches the distribution of $f \cos \phi$ for intensities of -70 db and above, relative to inverse distance.

¹⁴ C. Clark, "Motion of Sporadic-E Patches Determined from High-Frequency Backscatter Records," Radio Propagation Lab., Stanford University, Stanford, Calif., Tech. Rep. no. 24; 1957.

¹⁵ R. M. Gallet, "Aerodynamical mechanisms producing electronic density fluctuations in turbulent ionized layers," Proc. IRE, vol. 43, pp. 1240-1252; October, 1955.

There is evidence of an inverse correspondence between occurrence of VHF sporadic E and magnetic activity, but the relationship is not dependable for short time intervals. Over a period of two to four years, more sporadic E has occurred on quiet than on disturbed days in local time intervals 07–13 and 15–23 hours. Data taken over the whole year indicate that the relationship holds in most but not all years. A single month's observations frequently violate the quiet-day preference of sporadic E.

The frequency dependence of received power during periods of Es propagation is found to be quite different from that during periods of scatter propagation. This suggests that different propagation mechanisms are in effect under the two conditions. The median exponent n in the relation $P_r \propto 1/f^n$ is about twice as large for

sporadic E as for normal ionospheric scatter. Ratios of signal intensity at the two frequencies, 27.775 mc and 49.8 mc, show a wide range of values during periods of Es. This is in accord with a conception of the sporadic-E layer as composed of patches of varying electron concentration, such that 28-mc waves would be returned from a larger portion of the layer than would 50-mc waves.

ACKNOWLEDGMENT

The authors wish to express their thanks to J. W. Finney and D. H. Zacharisen for their assistance in preparing various sections of this paper, and to D. K. Bailey, R. Bateman and R. C. Kirby, who designed and supervised the experiment from which the data were obtained.

Some Characteristics of Persistent VHF Radiowave Field Strengths Far Beyond the Radio Horizon*

L. A. AMES†, E. J. MARTIN‡, AND T. F. ROGERS‡, SENIOR MEMBER, IRE

Summary—The most recent results of a continuing research program devoted to the study of long-distance VHF radiowave propagation are presented. Data obtained from measurements made at 50 and 220 mc on the surface and in the air are given which extend our previous results and clarify some important aspects of tropospheric and ionospheric propagation at extreme distances and heights.

INTRODUCTION

BY 1952–1953 the results of VHF, UHF, and SHF long-distance radiowave measurements by many workers in the United States and on the Continent were beginning to be appreciated as indicating a marked and consistent departure of experimental results from diffraction theory at distances deep within the diffraction zone.¹

* Original manuscript received by the IRE, December 12, 1958; revised manuscript received, February 16, 1959.

† Air Force Cambridge Res. Center, Bedford, Mass.

‡ Radio Phys. Div., Lincoln Lab., Mass. Inst. Tech., Lexington, Mass. Formerly Air Force Cambridge Res. Center, Bedford, Mass.

¹ See, for instance: a) K. Bullington, "Characteristics of beyond-the-horizon radio transmission," *Proc. IRE*, vol. 43, pp. 1175–1180; October, 1955. b) J. H. Chisholm, P. A. Portmann, J. T. deBettencourt, and J. F. Roche, "Investigations of angular scattering and multipath properties of tropospheric propagation of short radio waves beyond the horizon," *Proc. IRE*, vol. 43, pp. 1317–1335; October, 1955. c) K. A. Norton, P. L. Rice, and L. E. Vogler, "The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain," *Proc. IRE*, vol. 43, pp. 1488–1526; October, 1955.

Early in 1953, Bullington² suggested that the tropospheric field strengths were of sufficient strength and persistence such as to indicate their usefulness for communication for perhaps 100–200 statute miles beyond the radio horizon. By the middle of that year plans were laid for a propagation investigation in Newfoundland, the results of which were basic to the design of the first UHF tropospheric "beyond the horizon" communication system.³

In view of the experimental facilities available at the Air Force Cambridge Research Center, and the Air Force communications requirements which might hopefully be expected to be met by exploitation of this relatively newly-discovered radiowave propagation phenomenon, it was determined that we would concentrate initially upon the measurement of 220-mc field strengths on overwater paths between surface points separated by 200–400 statute miles. We would then proceed to extend this distance both on the ground and in the air as rapidly as circumstances might permit.

² K. Bullington, "Radio transmission beyond the horizon in the 40-to-4000-mc band," *Proc. IRE*, vol. 41, pp. 132–135; January, 1953.

³ K. Bullington, W. J. Inkster, and A. L. Durkee, "Results of propagation tests at 505 mc and 4090 mc on beyond-horizon paths," *Proc. IRE*, vol. 43, pp. 1306–1316; October, 1955.

Also, we would make a careful study of height-gain characteristics with the use of aircraft and attempt to discover characteristics which might distinguish 220-mc propagation from those exhibited at UHF and SHF.

Several of the results of the continuing Air Force Cambridge Research Center tropospheric propagation research program have been published to date.⁴⁻¹⁰ The new results given here are for measurements made beyond 500 statute miles and above 30,000 feet and include a simple summary of our best estimates of 100-1000-mc path loss over a smooth earth as a function of distance, frequency, altitude, and statistical occurrence.

By 1956 it was suspected that the rate of path loss experienced in VHF/UHF tropospheric propagation might increase rather rapidly in the vicinity of 600 miles and, while plans were laid for an investigation of field strength behavior at and beyond this distance, our attention was directed toward studies of the VHF ionospheric "scatter" propagation mode, since this mode is capable of providing a useful signal level at distances of some 1200 miles. In particular, a better appreciation of the important role of the antenna in making measurements of and utilizing the ionospheric scatter mode had by then been gained, and the initial airborne measurements by Abel¹¹ and Gerks¹² indicated that a carefully planned airborne experimental program should be capable of yielding valuable information concerning the general path loss distance dependence at great distances, height-gain, and some short-time signal characteristics. An experimental program was conducted jointly with the Lincoln Laboratory of the Massachusetts Institute of Technology; the characteristics of 50-mc propagation were explored in an aircraft to distances approaching 2000 miles. The results of these experiments have been

⁴ T. F. Rogers, "VHF field strength far beyond the radio horizon," *Proc. IRE*, vol. 43, p. 623; May, 1955.

⁵ L. A. Ames, P. Newman, and T. F. Rogers, "VHF tropospheric overwater measurements far beyond the radio horizon," *Proc. IRE*, vol. 43, pp. 1369-1373; October, 1955.

⁶ L. A. Ames and T. F. Rogers, "Available bandwidth in 200-mile UHF tropospheric propagation," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-3, pp. 217-218; October, 1955.

⁷ L. A. Ames, E. J. Martin, and T. F. Rogers, "Long distance VHF-UHF tropospheric field strengths and certain of their implications for radio communications," *IRE TRANS. ON COMMUNICATIONS SYSTEMS*, vol. CS-4, pp. 102-103; March, 1956.

⁸ L. A. Ames, N. L. Conger, J. W. Frazier, E. J. Martin, and T. F. Rogers, "High Altitude UHF Field Strengths Measured to Great Distances Beyond the Radio Horizon," *URSI Spring Meeting*, Washington, D. C.; May, 1957. Also, 12th General Assembly *URSI*, Boulder, Colo.; August, 1957.

⁹ T. F. Rogers, L. A. Ames, and E. J. Martin, "The possibility of extending air-ground UHF voice communication to distances far beyond the radio horizon," *IRE TRANS. ON COMMUNICATIONS SYSTEMS*, vol. CS-5, pp. 106-121; March, 1957.

¹⁰ E. J. Martin, "The Dependence of VHF/UHF Extra-Diffraction Tropospheric Field Strengths on Distance, Antenna Height and Horizon Angles," *AFCRC*, Bedford, Mass., Rep. TR-58-149; May, 1958.

¹¹ W. G. Abel, "Measurements in an Aircraft of 50 MC Field Strengths Along a 1500 Mile Path," *URSI Fall Meeting*, Berkeley, Calif.; October, 1956.

¹² I. H. Gerks, "Flight Measurements of Transmission Loss at 38.5 MC," *URSI Spring Meeting*, Washington, D. C.; May, 1957.

reported to *URSI*¹³ and formed the basis for a thesis by Orange.¹⁴

A paper reporting the findings in detail is in preparation and only certain features of the measurements will be given here.

In attempting to extend our airborne measurements of 220-mc tropospheric field strengths to distances well beyond 700 miles, an unexpected propagation feature was encountered: at about 800 miles the fields began to exhibit certain characteristics of ionospheric scatter.¹⁵ Further results of longer-term surface measurements are stated here which give additional evidence that 220-mc fields measured over a path length of 900 statute miles are at least predominantly ionospheric scatter in nature.

MEASUREMENT TECHNIQUES

Surface Measurements

Measurement of VHF field strengths at distances hundreds of miles beyond the average radio horizon requires powerful transmitters, very sensitive receivers, and high-gain antennas. All equipment must be rugged and stable over long measurement periods. Great care must be taken in the selection of habitable monitoring sites characterized by low ambient RF noise and the equipment must be carefully calibrated at the site.

Our early 220-mc measurements were made with a pulse transmitter and relatively wide-band receivers whose sensitivities were substantially increased as time progressed. (The characteristics of this early equipment are described by Ames, *et al.*⁵) Eventually, the pulse transmitter was replaced with a 40,000-watt average CW transmitter possessing very good frequency stability—a very few parts in 10⁸ for several hours—thereby permitting a substantial increase in receiver sensitivity through use of predetection bandwidths of 100 and 10 cps. We have continued to use the large "Billboard" dipole transmitting array⁵ and have used various Yagi and stacked dipole receiving antenna arrays possessing plane wave gains of up to 18 db above isotropic in free space.

In our most recent 220-mc experiments, we have had sufficient over-all sensitivity to be able to measure a path loss of some 280 db. An appreciation of this great measurement sensitivity may be gained by noting that it would permit *free-space plane-wave* measurements over a distance of some 5×10⁹ statute miles, *i.e.*, over a path length in excess of that separating the Earth from the planet Pluto, the most remote known body in our solar system.

¹³ W. G. Abel, A. S. Orange, and T. F. Rogers, "VHF Signal Level Measurements Along a 2000 Mile Path," *URSI Spring Meeting*, Washington, D. C.; April, 1958.

¹⁴ A. S. Orange, "VHF ionospheric scatter fields at great distances," submitted in partial fulfillment of the requirements for the M.S. degree, M.I.T., Cambridge, Mass.; June, 1958.

¹⁵ L. A. Ames and T. F. Rogers, "220-mc radiowave reception at 700-1000 miles," *Proc. IRE*, vol. 47, p. 86; January, 1959.

Airborne Measurements

Airborne measurements of very weak fields are inherently more difficult to make than comparable surface measurements. The equipment is subject to physical vibration. Accurate supply voltage and frequency regulation can be a problem. Electrical noise from the many electrical and electronic devices contained within the aircraft must be eliminated, or at least reduced to acceptable proportions. The design, construction, installation, and testing of airborne VHF antennas possessing relatively large gains (10–15 db) and having acceptable mechanical and aerodynamic characteristics is as much an art as it is a science. The oftentimes conflicting demands of the scientist intent upon achieving the best possible electrical performance and those concerned with the airworthiness—and even safety—of the aircraft have led to some interesting studies in the search for acceptable compromise. Fig. 1 is a photograph of a 10-db gain 220-mc Yagi antenna installed on the rear section of a B-47 used for measurements near 40,000 feet and ground speeds of up to 700 mph. Fig. 2 is a photograph of a Yagi array with a calculated gain of some 18 db installed on a C-54 (DC-4). This aircraft was used for relatively low-altitude (500–10,000 feet) flights at slow speeds but out to ranges of 1000 statute miles.

Since at great altitudes our very sensitive receivers are easily subject to interference from signal sources within line-of-sight of the aircraft, even though as much as 200–300 miles away, it was necessary that most of our measurements be made over water well beyond the coastline. The lack of desirably accurate navigation equipment, coupled with the large and generally unpredictable wind shifts often experienced in the ocean area east of the northern New England-Nova Scotia-Newfoundland regions, prohibited the simple use of narrow antenna beamwidths for our 220-mc measurements. Recourse was had to electrical beamswinging of our large ground transmitting antenna until the aircraft was beyond some 500 miles from the transmitter. (See Ames, *et al.*⁵ for details of this measurement technique.)

Our flight path for the 50-mc measurements was dictated by: 1) the location of the Lincoln Laboratory 500-kw average power ground transmitter in southern Texas, 2) the fact that the broad transmitting antenna beamwidth pointed toward the Boston area at an angle of 62° East of North, and 3) a desire to avoid both the intense man-made electrical noise found within line-of-sight of large urban areas and the at times intense atmospheric noise which is generally more severe in lower latitudes along the East coast. As a consequence, the chosen measurement path intersected the United States coast in northern North Carolina and continued out over the sea to a total distance of some 2000 miles. Measurements were usually completed by the time the aircraft was some 500–600 miles off shore and roughly

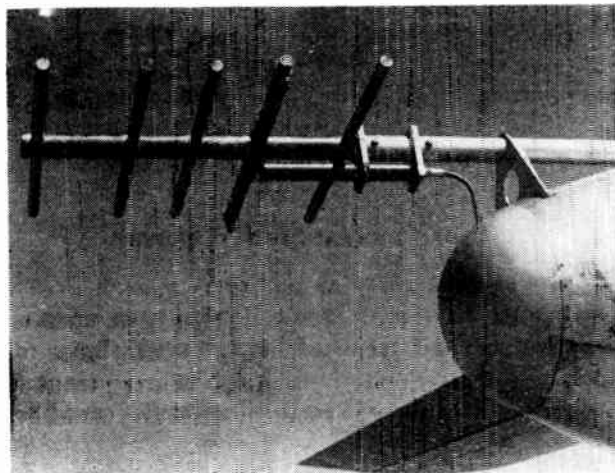


Fig. 1—220-mc Yagi antenna on B-47 aircraft.

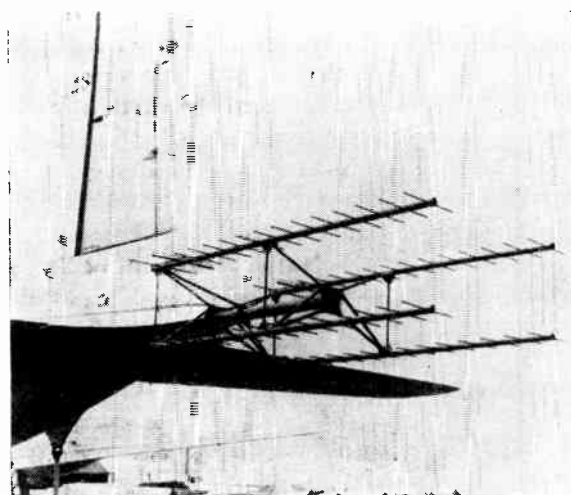


Fig. 2—4-Bay 220-mc Yagi antenna on C-54 (DC-4) aircraft.

halfway between Bermuda and Boston. Since wide angle ($\approx 60^\circ$) antennas were used both on the ground and aircraft, no severe difficulty with aircraft drift was experienced.

In order to encompass the range of altitudes desired to be measured at 220 mc, 500–40,000 feet, three different aircraft were used at various times: a C-54, B-29, and a B-47. All of the 50-mc experiments were conducted with a B-50 aircraft.

As the 50-mc airborne measurements were made with a receiver predetection bandwidth of 200 cps, little difficulty was experienced in keeping the received signal within the pass band. At 220 mc our latest measurements were made with a bandwidth of 10 cps. In this case, continual monitoring of the received signal in our specially constructed receiver was necessary to compensate for both small equipment drifts and the Doppler shifts associated with varying aircraft velocities.

Very stable electrical and electronic equipment is needed for airborne measurements. In addition, how-

ever, competence, imagination, and perseverance are required of scientists, engineers, and flight crews while the experiments are being planned and conducted so that the inherent complexities of this type of experiment may be overcome and accurate results obtained.

TROPOSPHERIC MEASUREMENTS

It was evident early in our investigation of tropospheric radiowave path loss that its dependence upon both distance and terminal height would have to be clearly understood before quantitative assessments and comparisons of various experimental results could be attempted.

LaGrone¹⁶ had originally introduced the concept that the most significant contribution to the power received at distances deep within the diffraction region arrives from the minimum angle, θ_{min} , formed by intersecting rays drawn from the transmitter and receiver terminals. Over a smooth earth, the rays forming a minimum angle are those tangent to the earth at the radio horizons in front of these terminals. (See Fig. 3.)

Thus, the path loss in excess of the free space loss to be expected over the surface distance should be related to the "effective distance" d' (or "angular distance" as Norton's group¹⁷ at the National Bureau of Standards has titled it) which involves the terminal heights directly and not simply to the surface distance alone.

Data from our airborne measurements have been used to test the validity of this "effective distance" concept. First, all of the experimental data points gathered at various altitudes between 500 and 40,000 feet for a given distance were compared to the height-gain curve which this concept predicts.

A comparison at 400 miles is presented in Fig. 4 and, in view of the serious experimental difficulties encountered in such measurements, this agreement is considered good.

A more direct comparison is one in which the fields are measured simultaneously at two widely separated altitudes. Such an experiment was carried out at 3000 and 30,000 feet over the same smooth path at as nearly the same time as the different flight characteristics of the two different types of aircraft would permit. These different heights provide an effective distance separation of about 170 statute miles.

Fig. 5(a) graphs the path loss vs surface distance for the two altitudes. In Fig. 5(b) the individual 10-mile median signal levels for the two altitudes are normalized to the same scale: path loss in decibels below free space vs the effective distance in statute miles. The data

¹⁶ A. H. LaGrone, "Cross polarization of scattered radio waves," *Proc. IRE*, vol. 40, pp. 1120-1123; September, 1952. Also, Rep. No. 48, Electrical Engineering Res. Lab., University of Texas, Austin; April, 1951.

¹⁷ Norton, *et al.*, *loc. cit.*

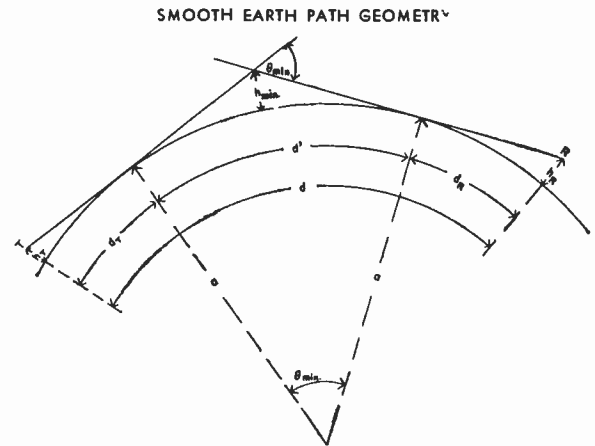


Fig. 3—"Effective distance" geometry.

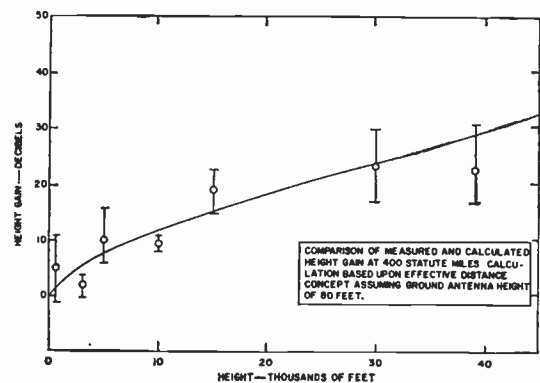


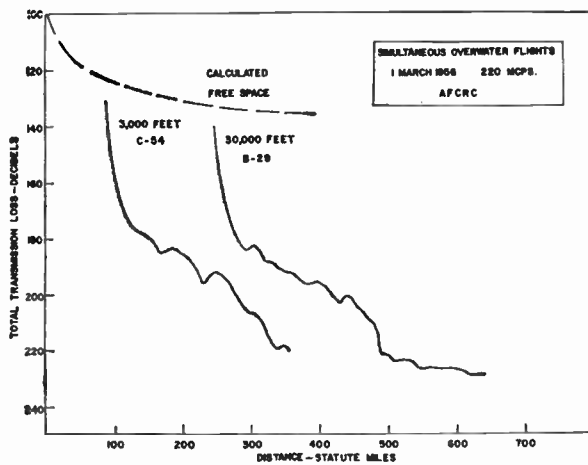
Fig. 4—Comparison of 400-mile data from 500 to 40,000 feet with height-gain curve predicted by effective distance concept.

points for the two altitudes are found to be in excellent agreement when normalized in this manner. Not only are the absolute values of path loss for the two altitudes in close agreement, but the slopes of the two decay curves and the smaller scale spatial variations are seen to be very closely related also.

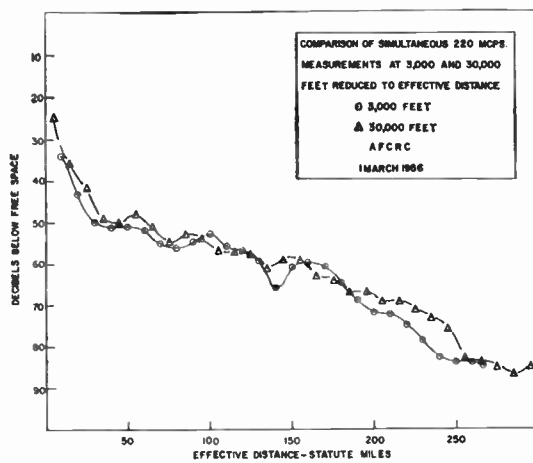
We may conclude, therefore, that the concept of effective or angular distance is a very valuable and accurate one—at least to 40,000 feet—when applied to the determination of path loss well beyond this radio horizon.

Starkey¹⁸ has reported some airborne measurements on 91 mc to a distance of 450 statute miles at 10,000 feet. He attributes certain of the propagation characteristics of the field strengths which he obtained in his measurements to specular reflections from elevated refractive index layers. Certainly this mechanism, or scattering from turbulent layers, does play an important part much of the time in the propagation of radio energy to distances of a few hundred miles beyond the radio horizon. This is brought out not only in the

¹⁸ B. J. Starkey, "Aircraft measurements of beyond-the-horizon propagation phenomena at 91.3 mc," *Proc. IEE*, vol. 103, pt. B, pp. 761-763; 1956.



(a)



(b)

Fig. 5—(a) Path loss results of simultaneous measurements at 3000 and 30,000 feet. (b) Comparison of simultaneous measurements normalized to decibels below free space vs effective distance.

measurements by Megaw,¹⁹ and their interpretation by Gordon,²⁰ but also in our overwater airborne measurements. On more than half of our flights (see Fig. 6) there is a characteristic bulge in the field strength decay with distance in the region of 150 miles corresponding to a rather persistent layer structure near 2500–3000 feet.

It is difficult to understand, however, how prominent layer structures could contribute to the regular tropospheric fields measured at very great distances beyond the radio horizon. Dinger, *et al.*,²¹ and the authors (see below) have measured persistent tropospheric fields at

¹⁹ E. C. S. Megaw, "Scattering of electromagnetic waves by atmospheric turbulence," *Nature*, vol. 166, pp. 1100–1104; December, 1950.

²⁰ W. E. Gordon, "Radio scattering in the troposphere," *Proc. IRE*, vol. 43, pp. 23–28; January, 1955.

²¹ H. E. Dinger, W. E. Garner, D. H. Hamilton, Jr., and A. E. Teachman, "Investigation of long-distance overwater tropospheric propagation at 500 mc," *Proc. IRE*, vol. 46, pp. 1401–1410; July, 1958.

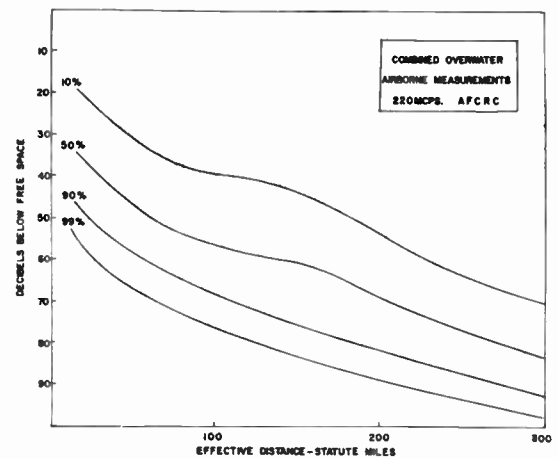


Fig. 6—Overwater 220-mc statistics illustrating the influence of the layer-reflection propagation mechanism on field strengths 100–200 miles beyond the radio horizon.

an effective distance of 700 miles. In order that a mechanism of reflection or scattering from elevated layers be responsible for such long-distance propagation, the continued presence of such layers would be required at an altitude of some 60,000 feet; to our knowledge, there is no strong evidence to suggest the existence of such a continuous structure at such a great height.^{18,20}

The early comparisons of VHF and UHF data by Bullington,² and the comparison which Megaw drew between his 10-cm results and those obtained by Rocco and Smyth²² on 3.2 cm, gave some indication of a frequency dependence in the long-distance fields. This indication appeared to be substantiated when our early 220-mc measurements were compared with Megaw's 3300-mc data.¹⁹ The recent comparison which Dinger, *et al.*,²¹ have made between their 400-mc long-distance measurements and those of Megaw adds further evidence that there is a small but noticeable increase in path loss with frequency. In his study of data from various VHF-UHF paths, Martin¹⁰ concluded that a path loss frequency dependence of between the square root and first power of the frequency fits the data well between 100–1000 mc.

Martin¹⁰ has been able to make a good general estimate of the expected path loss with data available through the end of 1956 from AFCRC surface and airborne measurements and the published surface measurements of the Lincoln Laboratory of M.I.T., the Bell Telephone Laboratories, The National Bureau of Standards, and the Collins Radio Company. His estimate was restricted to the frequency interval 100–1000 mc and distances out to 400 effective miles, with usable estimates out to 600 effective miles. An accurate esti-

²² M. D. Rocco and J. B. Smyth, "Diffraction of high-frequency radio waves around the earth," *Proc. IRE*, vol. 37, pp. 1195–1203; October, 1949.

mate for the 400–600 mile region was not possible at that time because of the paucity of data: only the few NBS measurements at 100 mc were available,^{10,23} and there was a chance that the ionospheric scatter propagation mode might have influenced these data.

The later results of the Lincoln measurements at 400 mc and 600 statute miles²⁴ indicated a median path loss some 6 db greater than Martin's estimates, suggesting that the rate of path loss might be increasing beyond 500 miles. The 400-mc measurements of Dinger, *et al.*,²¹ (see the extreme distance runs presented in their Fig. 6) give rather conclusive evidence that the rate of attenuation is increasing beyond approximately 500–550 statute miles.

During the Spring of 1958 we were able to make aircraft measurements of 220-mc tropospheric fields beyond an effective distance of 700 statute miles (see Fig. 7) and to supplement these data with surface measurements over a 250-hour period at Stephenville in Newfoundland in February. (See Fig. 8.) This latter path has a surface distance of 735 miles, with $d' = 690$ miles. Both our airborne and surface measurements are in very good agreement with the data of Dinger, *et al.*,²¹ at 700 miles when normalized with respect to effective distance and frequency. As a result of these various recent experiments, the earlier median estimations of Martin can be adjusted and extended to approximately 700 effective miles with considerable confidence, and this has been done in Fig. 9(a). Fig. 9(b) presents such an estimate in terms of surface distance and terminal height. It must be kept in mind that ionospheric propagation modes cannot be neglected at the greater distances for frequencies below 300 mc.

It should be noted in passing that the median value of path loss which we have measured near 700 effective miles is in substantial agreement with the recent prediction of Booker and Gordon²⁵ which is based upon a stratospheric scattering model.

IONOSPHERIC MEASUREMENTS

Our interest in ionospheric propagation measurements in the lower VHF region arose from our concern with attempting to make accurate estimates of both the absolute value of the path loss and the dependence of the loss with distance for distances of approximately 1200 statute miles and beyond. Our concern with absolute path loss estimates was connected with a growing

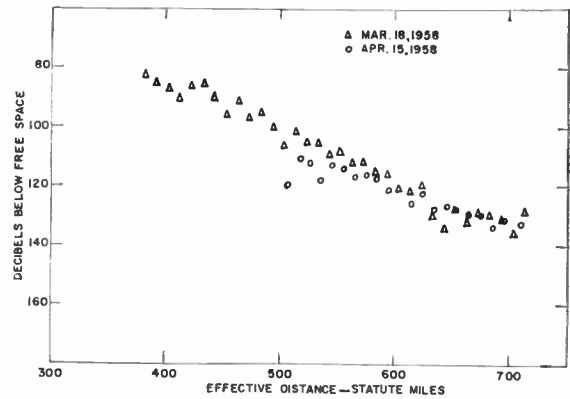


Fig. 7—Airborne measurements of path loss in 400–700 mile d' region.

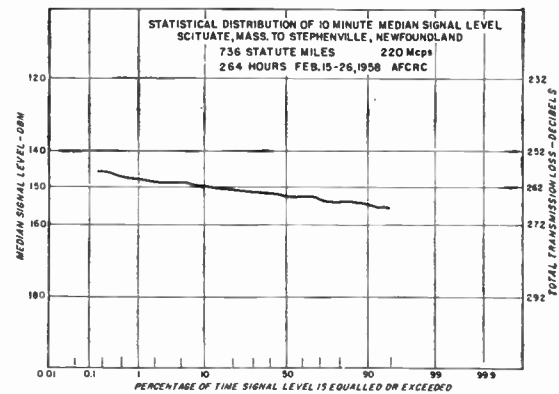


Fig. 8—Cumulative statistics of Stephenville 200-mc ($d' = 690$ miles) measurements.

appreciation of the fact that because high-gain narrow-beam antennas do not maintain their plane wave gain (even to a good first approximation) with the ionospheric scatter mode, values of the basic path loss deduced from measurements utilizing such antennas could be subject to misinterpretation.²⁶ (Rogers²⁷ has discussed this matter in detail in a paper to be published, the results of which have been presented to URSL.) Also, measurements reported by Bailey, *et al.*,²⁸ have shown (see their Fig. 55) that the slope of the cumulative distribution of median received signal levels was much less when measured with 65° antennas than with 6° antennas, and this fact, which is of importance in regard to weak signal reception and utilization, should be substantiated by other similar measurements.

Fig. 10 presents the results of a characteristic flight measurement made at 50 mc with wide angle ($\approx 60^\circ$) antennas. The line-of-sight, diffraction, tropospheric

²³ A. P. Barsis and R. E. McGavin, "Report on comparative 100-mc measurements for three transmitting heights," IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-4, pp. 168–174; 1956. Also, A. P. Barsis, private communication.

²¹ J. H. Chisholm, "Recherches expérimentales sur la diffusion angulaire et sur les possibilités d'utilisation de la propagation troposphérique pour les communications bien au delà de l'horizon," *L'Onde Elect.*, vol. 37, pp. 427–434; May, 1957.

²⁵ H. G. Booker and W. E. Gordon, "The role of stratospheric scattering in radio communications," Proc. IRE, vol. 45, pp. 1223–1227; September, 1957.

²⁶ R. C. Kirby, "VHF propagation by ionospheric scattering—a survey of experimental results," IRE TRANS. ON COMMUNICATIONS SYSTEMS, vol. CS-4, pp. 17–27; March, 1956. (See Fig. 15.)

²⁷ T. F. Rogers, "The Influence of Continuous Meteoric Scattering on Long Distance VHF Communication Circuit Design," URSL Spring Meeting, Washington, D. C.; April, 1958.

²⁸ D. K. Bailey, R. Bateman, and R. C. Kirby, "Radio transmission at VHF by scattering and other processes in the lower ionosphere," Proc. IRE, vol. 43, pp. 1181–1230; October, 1955.

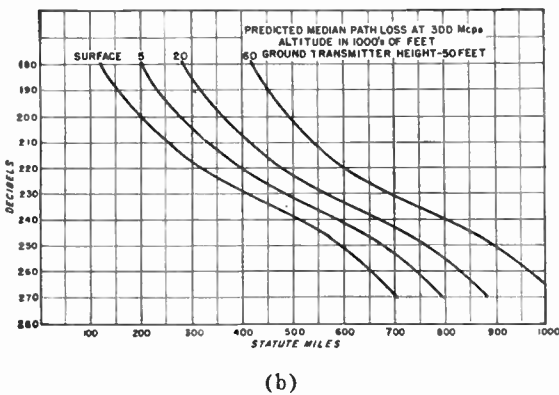
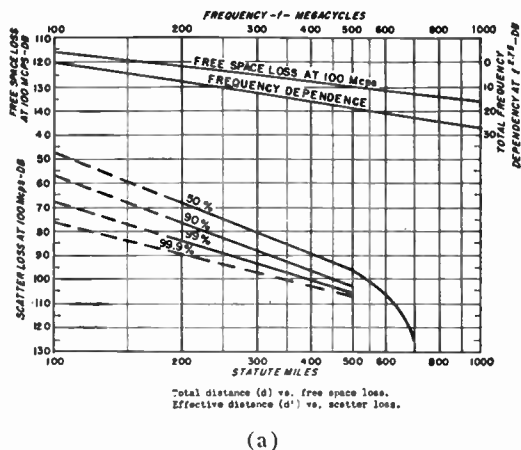


Fig. 9—(a) General estimate of 100–1000-mc tropospheric path loss. (b) Estimate of 300-mc path loss vs surface distance from the surface to 60,000 feet.

extra-diffraction, ionospheric scatter, and quasi-diffraction ionospheric scatter regions are all successively displayed. An analysis of the results of the many such flights made jointly with the Lincoln Laboratory yields an estimate of the statistical variation of path loss with distance as given in Fig. 11. With the assumption that our antenna gains are maintained essentially as well as those of dipole antennas would be—an assumption believed to be a reasonable one based upon our present knowledge of the off-path angles over which the meteoric component may be expected to be important²⁹—these values of path loss can be stated with confidence in absolute terms relative to dipoles or isotropic radiators. As such, they supplement the long-term point-to-point surface statistics gathered by other workers.^{26,28,30,31}

²⁹ W. R. Vincent, R. T. Wolfram, B. M. Sifford, W. E. Jaye, and A. M. Peterson, "Analysis of oblique path meteor—propagation data from the communications viewpoint," *Proc. IRE*, vol. 45, pp. 1701–1707; December, 1957. Also, W. R. Vincent, private communication.

³⁰ W. G. Abel, J. T. deBettencourt, J. H. Chisholm, and J. F. Roche, "Investigations of scattering and multipath properties of ionospheric propagation at radio frequencies exceeding the MUF," *Proc. IRE*, vol. 43, pp. 1255–1268; October, 1955.

³¹ W. J. Bray, J. A. Saxton, R. W. White, and G. W. Luscomb, "VHF propagation by ionospheric scattering and its application to long-distance communication," *Proc. IEE*, vol. 103, pt. B, pp. 236–257; 1956.

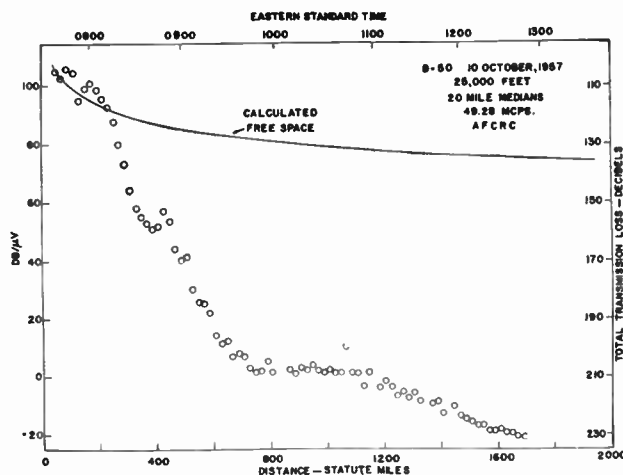


Fig. 10—Characteristic path loss measurement at 50 mc showing line-of-sight, diffraction, tropospheric extra-diffraction, ionospheric scatter, and ionospheric scatter quasi-diffraction regions.

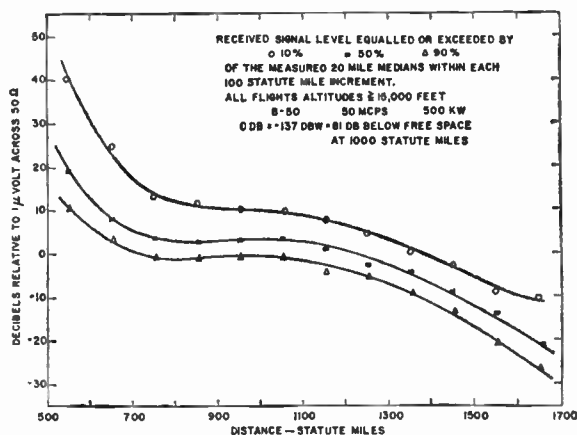


Fig. 11—Cumulative statistics of path loss vs distance at 50 mc obtained from airborne measurements.

By examining all of the flight data available for one distance increment of, say, 50–100 miles, and with an appropriate correction for aircraft altitude, a rough comparison can be made between our wide angle results and similar results obtained by others at the surface with narrow-beam antennas. Fig. 12, for instance, compares the distribution of signal levels which was obtained in the air near 800 statute miles to that obtained by the NBS with their narrow-beam experiments on paths of similar length.²⁸ Our data have been normalized only with respect to transmitter power—the antenna gains have been ignored in each case. With reservation prompted by the differences in experimental conditions—location, time, altitude, and number of samples being the most important—it still appears that the difference in slope is sufficiently marked to be significant, and we take this to be corroborative evidence of the Bailey, *et al.*,²⁸ experience.

We were not able to draw significant conclusions re-

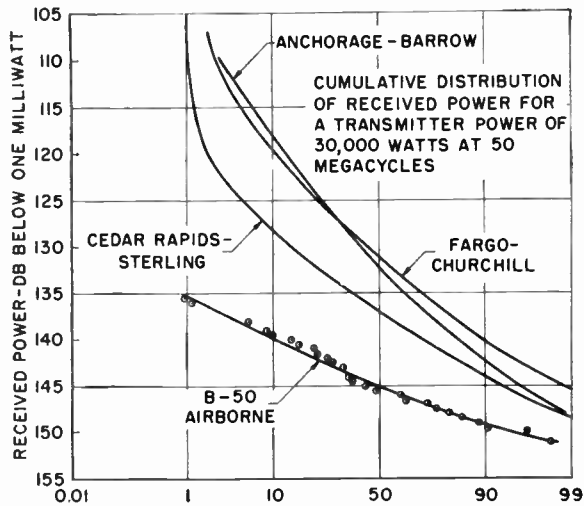


Fig. 12—Comparison of 800-mile region (effective distance) cumulative statistics of received power obtained from airborne measurements with wide-angle antennas to the NBS surface data obtained with narrow-angle antennas. (See Bailey, *et al.*²⁸)

garding the expected distance dependence of the remote quasi-diffraction ionospheric scatter region on terminal height. Circumstances surrounding the experiment prevented us from making measurements at widely separated heights within a short span of time, and the temporal variations in received signal levels equalled or exceeded those expected to be caused by height variations alone.

There did not appear to be any significant difference in short-time signal behavior in the 1400–1800 mile region when compared with that of the 800–1200 mile region.

As reported recently,¹⁵ we were surprised to find that the signal behavior during our 1958 airborne 220-mc measurements beyond 700–800 miles appeared to exhibit the characteristics of ionospheric scatter. In order to supplement our airborne experience in the 800–1000 mile region, a long term measurement was made over a path length of 900 statute miles during June of 1958. A 14-db gain 4-bay Yagi array having an azimuth beamwidth of some 25° was installed in Argentia, Newfoundland, 892 statute miles from our transmitter south of Boston. Signals were measured in a 10-cps pass band in order to provide a sufficient signal-to-noise ratio during the weaker signal periods; valid data were obtained for some 650 hours during the month.

A continuous background signal was present throughout the entire measurement interval. Superimposed upon this continuous signal, however, was a great deal of meteoric activity. When recorded with a 10-cps bandpass and a high-speed recorder, the greater majority of the bursts exhibited durations of much less than a second, measured at points 3–6 db down from their maximum, although a few instances were observed in which (a presumed) meteor-induced field strength

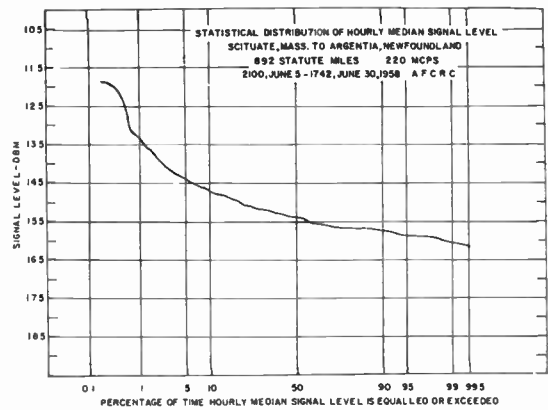


Fig. 13—Cumulative statistics of Argentia ($d=892$ miles) 220-mc ionospheric scatter measurements.

increase lasted over 10 seconds. Qualitatively, our meteor signal observation appeared to be similar to the 200-mc meteor echo observations recently reported by Heritage, *et al.*³²

Because of the relatively high frequency components in the fading power spectrum of the very weak 220-mc continuous scatter signal, the measurement accuracy cannot be claimed to be as high as that associated with our tropospheric measurements. Fig. 13 presents a cumulative distribution of the hourly median signal levels recorded during the measurement period, and we judge that these data are probably accurate to within ± 5 db. In view of the relatively narrow antenna beamwidths used (4° and 25°), we are unable to relate the ratio of received to transmitted power quantitatively to the transmission loss, since the plane-wave gains of our antennas would not be expected to be generally realized.

A comparison with the tropospheric signal distribution obtained at Stephenville (see Fig. 8) is interesting. The tropospheric signal level distribution is essentially a decibel-Gaussian distribution with a very small standard deviation. The Argentia signal level statistical distribution, when similarly graphed, is seen to have a greater slope, is not decibel-Gaussian, and exhibits relatively strong received power levels for short periods of time. These are characteristics of other lower frequency ionospheric scatter cumulative distributions. The Argentia ionospheric scatter median received power is seen to be essentially the same as the Stephenville tropospheric level, even though the path length to Argentia is 155 miles greater. The strongest signals over the 900-mile path, however, were some 20 db greater than at 700 miles.

The characteristics of very rapid short-time fading, the slope of the hourly median signal level distribution,

³² J. L. Heritage, S. Weisbrod, and W. Fay, "An Experimental Study of Meteor Echoes at 200 MC," URSI Spring Meeting, Washington, D. C.; April, 1958.

and the median signal distance dependence in the 800–1000 statute mile region strongly suggest that these extreme distance 220-mc fields are propagated via an ionospheric scattering mechanism.

CONCLUSIONS

Tropospheric field strengths have been measured at 220 mc to distances of as much as 700 statute miles beyond the radio horizon and at various altitudes between 500 and 40,000 feet. At distances well into the extra-diffraction region it has been demonstrated that the “effective” or “angular” distance concept is valid. At distances of 100–200 miles beyond the radio horizon, reflection or scattering from elevated layers is an important contributing propagation mode but does not appear to be the sole agent responsible for the measured field strengths. The influence of such layers on field strength characteristics decreases with distance, perhaps becoming negligible at distances in excess of some 300–400 miles beyond the radio horizon. There are sufficient data available for a reasonably accurate prediction of median tropospheric path loss in the 100–1000-mc region out to 700 miles beyond the radio horizon and up to heights of 40,000 feet.

Field strengths of 50 mc have been measured in the air to a distance of nearly 2000 miles, and the long-distance tropospheric-ionospheric scatter and ionospheric scatter quasi-diffraction regions observed. More accurate absolute estimates of the 50-mc path loss and its variation can now be made for distances in the 1000–1700 mile region. Very weak 220-mc fields have been measured at the surface and in the air in the 800–1000 statute mile region and have been identified as ionospheric in origin.

The combination of fairly long-term surface measurements made with portable equipment and shorter-term

airborne measurements at various altitudes and distances have yielded a great amount of valuable VHF radiowave propagation data over the past five years. These data have enabled the scientist to obtain a deeper insight into variegated VHF propagation phenomena; to the radio engineer, they supply information vitally needed in the exploitation of this part of the radio spectrum for new and reliable long-distance communication methods. This research program presents a good example of how a long-term continuous effort by a government laboratory, oftentimes utilizing equipment and measurement techniques readily available perhaps only to a laboratory of its kind, can yield results of both important scientific interest and valuable engineering utility.

ACKNOWLEDGMENT

The authors wish to express their appreciation to those from AFCRC who have contributed to this research program. R. Barrett, E. Bosma, J. Frazier, M. Hill, and T. Willson have been associated with the program since its inception some five years ago. H. Chapman, N. Conger, P. Newman, A. Orange, and J. Reegan have each made valuable contributions.

We keenly appreciate the contribution which the AFCRC airborne installation groups and flight and ground crews have made to the airborne measurement program.

Our association with the Lincoln Laboratory, M.I.T., and especially with J. Chisholm, W. Abel, and W. Morrow has been particularly close and helpful over the years. We would like to thank them and various individuals at the Bell Telephone Laboratories, Collins Radio Company, The National Bureau of Standards, the Naval Research Laboratory, Cornell University, and the University of Texas whose interest, advice, and assistance have been important and gratifying.

A Lightweight and Self-Contained Airborne Navigational System*

R. K. BROWN†, N. F. MOODY†, P. M. THOMPSON†, R. J. BIBBY†,
C. A. FRANKLIN†, J. H. GANTON†, AND J. MITCHELL‡

Summary—The paper discusses the design of a self-contained navigational aid for aircraft. There are many techniques and variations upon which such a system may be based. The one chosen for discussion is a hybrid system which combines an inertial north reference, a Doppler navigational radar, and an airspeed indicator. A velocity triangle computer relates these three sources of input data, so that the airspeed indicator may fill in when there is no radar signal. The output of these instruments, track angle and distance, is fed to a positional computer and thence to a steering computer, so that both positional and steering information may be supplied to the pilot.

Part I introduces the general problem of a self-contained navi-

gational aid, the possible sources of input data, the usual frames of reference within which aircraft position may be stated, and the forms of route to a destination which may be derived. Part II describes an FM Doppler radar as part of an over-all navigational system, and Part III the circuit implementation of the radar, based on transistors. Part IV describes a converter for connecting the radar to an analog positional computer, and some velocity vector triangle computers. Part V discusses a digital form of positional computer and some of the necessary digital processes. The paper concludes with Part VI, which presents the problem of steering computation and some ideas on the implementation of a complete system.

Part I—An Introduction to Aircraft Navigation

P. M. THOMPSON, N. F. MOODY, AND R. K. BROWN

THROUGHOUT the history of navigation, it has been the usual practice for navigators to carry their equipment with them. Such is still the case at sea, but in the air there has been the opposite tendency, so that today most airborne navigation is by means of radio aids which depend upon surface stations. The cause of this trend is that the earlier aircraft were unable to carry anything but the simplest radio equipment and were forced to rely upon these surface stations to measure their position for them. Historically this has led to the many sophisticated land-based radio navigational aids today. However, there are large parts of the surface of the earth, such as the Canadian Arctic, still not covered by these aids. Furthermore, in time of war their use might well be denied the navigator since they would be subject to jamming, or could be of equal use to an enemy.

There is a requirement, then, for a self-contained airborne navigational aid, capable of operating anywhere over the surface of the earth at any time and under any conditions. Such an aid must be light in weight as, although it is now possible for an aircraft to carry heavy loads, any unnecessary weight represents a reduction in payload and may thus reduce the efficiency of the aircraft by a prohibitive amount.

The process of navigation of an aircraft consists essentially of discovering its position, and then determining a route to its destination. It is impossible to define a position without referring to its surroundings, since no point can be defined by reference to itself; so an external

frame of reference suitable for the determination of the relative positions of aircraft and destination is necessary. The function of an aircraft navigation system, then, is to collect available data on the aircraft's progress, to place it within some frame of reference, and finally to compute a route to its destination.¹

There are several sources of data which might be suitable for such a navigational aid. Magnetic devices or inertial devices could be used as a directional reference or compass. Position could be given by terrestrial or astronomical map reading, or could be computed from direction and speed (or distance) given by radar or airspeed indicators. Also, speed may be given indirectly by inertial devices which measure accelerations.

Fuller discussion follows of these several sources of data. In assessing their relative suitability as navigational aids, it is necessary to introduce a criterion which becomes more important as aircraft speed, and also accuracy of navigation, increases. In general, positional information, within a frame of reference suitable for navigation, is obtained as a result of some interpretative or conversion process on the input data. The process involves some time delay, so that the resultant information is available only in retrospect. Present position, therefore, is based upon a prediction. Hence accurate navigation requires that the aids should yield data promptly in a form suitable for rapid conversion processes.

1) Magnetic and inertial heading references: Magnetic compasses will provide a heading reference useful at low geomagnetic latitudes. However their indications are disturbed in the vicinity of magnetic bodies, particu-

* Original manuscript received by the IRE, February 12, 1959. This project was done under the authority of the Defence Res. Board of Canada.

† Defence Res. Telecommun. Est., Ottawa, Ont., Can.

‡ Link Aviation, Binghamton, N. Y.; formerly with Defense Res. Telecommun. Est., Ottawa, Ont., Can.

¹ W. R. Fried, "Principles and performance analysis of Doppler navigation system," IRE TRANS. ON AERONAUTICAL AND NAVIGATION ELECTRONICS, vol. ANE-4, pp. 176-196; December, 1957.

larly near the poles, where it is usual to rely upon inertial devices.

2) Terrestrial and astronomical map reading: Navigation by map reading consists in identifying points on the ground, or in the sky, with points on a map; its success depends equally on being able to see and possessing an adequate map. Seeing is not limited to the visual spectrum, and radio wavelengths may be useful. Terrestrial map reading may be assisted by such radars as H₂S, which remove the limitations of normal vision since the aircraft carries its own source of illumination which can penetrate weather and darkness. Astronavigation by radio stars is still limited by the present state of development of radio receivers. Visual methods are limited by speed of interpretation, weather and time of day. In the Arctic regions the twilight period, when it is impossible to see either the sun or the stars, may last for many hours and often for a whole flight. When it is possible to see the stars, there are many suitable maps, but there are still parts of the earth not adequately mapped, and many places such as the sea or the Arctic where there are few, if any, identifiable points.

3) Doppler navigational radar¹⁻⁸: Radar, using the Doppler shift in signals scattered back from the ground, gives ground speed (or distance) directly and appears capable of providing a useful output anywhere over the surface of the earth.

4) Airspeed indicator: An aircraft may be navigated with respect to the air mass through which it flies, using true airspeed and a directional reference. Information about the movement of this air mass can be provided by ground stations and used to compute the progress of the aircraft with respect to the ground.

5) Inertial devices: Acceleration sensitive devices alone will provide sufficient information for an aircraft's progress to be measured. However such completely inertial systems are very expensive, and it is economical to use inertial devices only to provide a directional reference.

A system based on a combination of these sources of input will permit an aircraft's progress to be measured continuously with an accuracy and stability high enough for most purposes. However, a navigator must know his position within a frame of reference. The four types of reference most commonly used are:

1) Latitude and longitude: A latitude and longitude reference is universal and unique and finds its major application where aircraft must fly long distances, and where a pilot or navigator wishes to communicate his position with widely spaced centers, perhaps under different administrations. However, it becomes difficult to use near the poles, where special "square" grids are in use.

2) "Square" grid: Reference grids, which are close spherical approximations to the rectangle, are in use over most of the populated regions of the earth, as well as at the poles. They constitute a fairly simple method of representing position, whenever the aircraft remains in the area covered by a single grid.

3) Track and bearing: Giving position in relation to a desired track is a very simple system. Position, stated as a distance along and across this track, finds application in flying along air lanes. The bearing system is often used when the position of an aircraft operating from one station is stated as a range and a bearing from that station.

4) System references: Many ground-based radio navigation systems such as Decca, Dectra, Loran, Shoran, and Gee indicate position on their own system of curved coordinates. For an important traffic lane, these curved coordinates are sometimes made to approximate a great circle by careful siting of stations.

Once present position and destination are known in the desired system of reference, steering information may be computed. The steering information required usually is the bearing on which to fly, and the approximate distance so that the fuel requirements may be computed for the flight. The bearing is usually required as a great circle or close approximation, a rhumb line,⁹ or with reference to an air lane.

A discussion of a complete navigation system follows. First the sources of input data are described, followed by techniques for computing position, and thence by steering information. Only automatic computing techniques are discussed, since in modern aircraft the navigator has insufficient time to make the computations unaided.

A. AN OUTLINE OF THE SYSTEM DESIGN

A convenient and economical design (Fig. 1) uses input data consisting of direction and speed of flight. The heading¹⁰ of the aircraft is provided by an inertial heading reference, and when heading is added to the drift angle¹¹ this gives track.¹² Information from which speed and drift angle may be computed is available from two sources: a Doppler navigational radar and an airspeed indicator. The computation of ground speed from

⁹ Rhumb line, a line between two places on the surface of the earth, which maintains a constant bearing with true north.

¹⁰ Heading, angle between datum line of aircraft and a reference direction (north on a latitude-longitude system).

¹¹ Drift, angle between aircraft datum and actual direction of flight.

¹² Track, angle between direction of flight and reference direction.

¹ E. G. Walker, "Factors in the design of airborne Doppler navigation equipment," *J. Brit. IRE*, vol. 18, pp. 425-442; July, 1958.

² F. B. Berger, "The nature of Doppler velocity," *IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS*, vol. ANE-4, pp. 103-112; September, 1957.

³ F. B. Berger, "The design of airborne Doppler velocity measuring systems," *IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS*, vol. ANE-4, pp. 157-175; December, 1957.

⁴ M. W. McKay, "The AN/APN-96 Doppler radar set," 1958 *IRE NATIONAL CONVENTION RECORD*, pt. 5, pp. 71-77.

⁵ T. Gray and J. Moran, "Decca Doppler and airborne navigation," *Brit. Commun. and Electronics*, vol. 5, pp. 764-771; October, 1958.

⁶ F. A. McMahon, "The AN/APN-81 Doppler navigation system," *IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS*, vol. ANE-4, pp. 202-211; December, 1957.

⁷ M. A. Condie, "Basic design considerations—Automatic navigator AN/APN-67," *IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS*, vol. ANE-4, pp. 197-201; December, 1957.

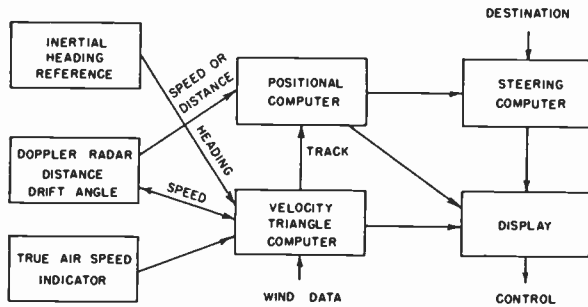


Fig. 1—The over-all navigational system consisting of the following. 1) Input organs: north reference, Doppler navigational radar, and airspeed indicator. 2) Computers: velocity triangle computer, positional computer, and steering computer.

true airspeed, and wind speed and direction, is performed by a velocity triangle computer, which also computes track from drift and heading. Ground speed and drift are also given directly by the radar. Note, at this stage, that the radar provides speed as an output waveform whose frequency is such that one cycle represents unit increment of distance. While the radar may therefore yield direct distance information, the system may be understood more readily if the output of the radar is considered as representing ground speed.

This part of the system is capable of operating in either of two modes: the Doppler mode, when the radar provides ground speed and drift angle directly, and the velocity triangle computer provides track from heading and drift angle; and the memory mode, when the velocity triangle computer uses stored wind data, true airspeed, and heading to compute track and ground

speed. When the system is switched on with the aircraft standing on the ground, there will be no output from the radar, and the system will be in its memory mode (with the wind data set by hand). As soon as the radar obtains a usable signal, usually just on leaving the runway, the system switches to the Doppler mode and thereafter calculates wind continuously. Should the Doppler signal fail due to aircraft attitude, or for any other reason, the system will switch to memory and use the last computed wind.

The north reference, radar, airspeed indicator and triangle are connected in such a way that the ground speed (or distance) output is from the radar, and the track angle output is from the velocity triangle computer. A positional computer uses these two outputs, together with a preset starting position, to compute position continuously during flight.

A steering computer, in turn, uses this calculated position to determine a route to a preset destination. The information from all three computers may be displayed as required, or used in connection with such devices as an automatic pilot to control the aircraft.

Certain of the following sections of this paper describe in detail the Doppler radar, which is the heart of the system. Many of the circuits are discussed, and, since weight is at a premium and reliability is of paramount importance, transistors are a natural choice for the active devices. The sections on computational techniques do not present any single system, since the requirements will vary with application. The heading reference and the airspeed indicator are outside the scope of this paper, and are not described.

Part II—The Principles and Performance of the Doppler Radar

R. K. BROWN

The purpose of the navigational radar within the over-all navigational system (Fig. 1) is to measure aircraft ground speed (or distance traveled over the surface of the earth) and drift angle. This part of the paper treats the radar as part of a system, describing the principles by which it operates and the measured performance. The separate circuits will be treated in Part III.

The principle by which the radar operates is based upon the Doppler shift of a microwave signal transmitted from the aircraft and scattered back from the ground at a point either ahead of or behind the aircraft. Fig. 2 shows an aircraft traveling over the surface of the earth. First consider a single microwave beam, directed forward from the aircraft to strike the ground at A [Fig. 2(a)]. A receiver located at A on the ground would receive a signal shifted in frequency from that of the transmitter in the aircraft due to the Doppler effect. The frequency shift, f_d' , is given by

$$f_d' = \frac{F V_g \cos \phi}{c} \quad (1)$$

where

f_d' = Doppler frequency,

F = transmitter frequency,

c = propagation velocity of electromagnetic radiation,

V_g = ground speed of aircraft,

ϕ = angle between the aircraft's direction of flight and the direction of the transmitted signal.

If some of the energy striking the ground at A is scattered back to the aircraft, the signal, as received at the aircraft, suffers a second Doppler shift exactly equal to the first. The total change in frequency as observed at the aircraft and called the Doppler frequency, f_d , is given by

$$f_d = 2f_d' = \frac{2F V_g \cos \phi}{c} \quad (2)$$

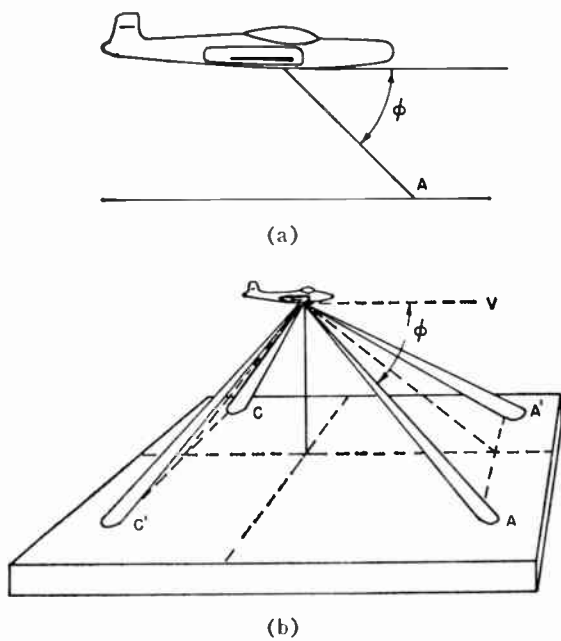


Fig. 2—Doppler navigational beam structure. (a) A single forward beam. (b) The four beams of the practical radar.

In practice, because of the finite width of the beam and the uneven nature of the ground, the Doppler return is not a single frequency but a noise spectrum approximately Gaussian in shape, whose width depends primarily upon the beamwidth. A typical spectrum is shown in Fig. 3(b).

If (2) is rearranged to make V_{θ} the independent variable and if, further, both sides are integrated, we obtain the distance R which the aircraft has traveled. Thus

$$R = \frac{n\lambda}{2 \cos \phi}, \quad (2a)$$

where n = number of Doppler cycles and λ = wavelength of the transmitted signal.

It is seen that the distance flown may be determined by simply counting the total number of Doppler cycles. In practice, the frequency counted in such a system is that of an oscillator in the tracker whose function is to provide a highly accurate representation of the mean Doppler frequency. (Reference will be made later to this element, which for clarity, will be called the tracker oscillator.)

In order to determine the aircraft drift angle, the two forward beams A and A' of Fig. 2(b) are substituted for the single beam. If the aircraft velocity vector is not along heading, *i.e.*, drift exists, the Doppler shifts for the two beams will not be the same. This information is used via a suitable servoloop to rotate the antenna system in azimuth until the Doppler frequencies obtained from the two beams become equal. When this has been done, the axis of symmetry of the antenna system has been aligned with the ground track of the aircraft, and the angle between this axis and the aircraft heading defines the drift angle.

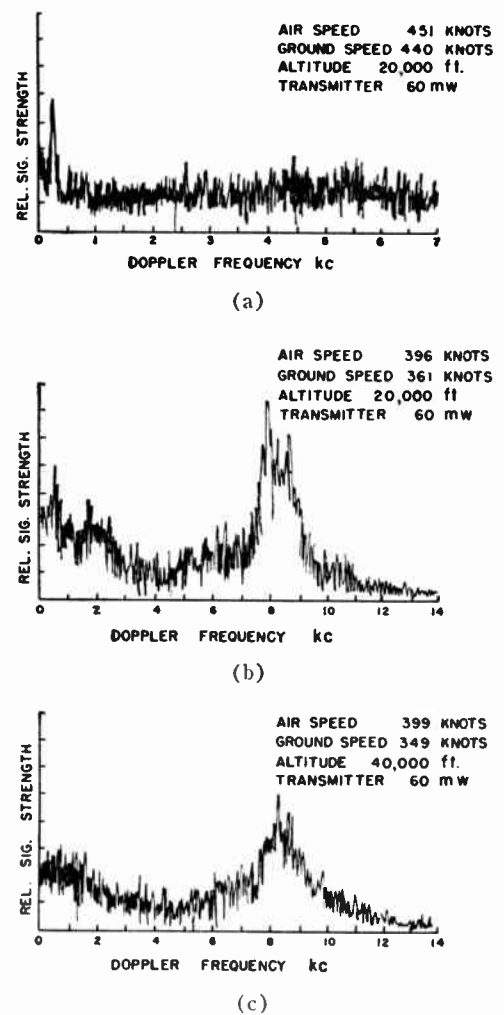


Fig. 3—Measured spectra of Doppler radar return signals. (a) An unusable spectrum obtained when using a single beam. (b) A spectrum obtained under similar conditions to (a) when using both fore and aft beams. [The peak at approximately 8 kc would correspond with a peak at 4 kc in (a).] (c) A spectrum which illustrates that a 10-dB SNR is obtainable when using a 60-mw klystron at 40,000 feet.

Fig. 2(b) also shows a rearward beam corresponding to each forward beam (C with A and C' with A'). The two pairs of beams are time-shared, C and A being on together for about 0.5 second, followed by C' and A' for 0.5 second. The instantaneous sums of the Doppler shifts on beams C and A and C' and A' are now available. It will be shown that the substitution of two paired beams for two single beams results in considerable improvement in accuracy.

The use of such paired beams produces a system referred to as a Janus system, after the Roman god who looked forward and backward at the same time. Three important advantages result from the use of such a technique:

- 1) The error introduced by any vertical component of aircraft velocity is completely eliminated.
- 2) The ground speed error due to variations in the angle caused by pitch movement of the antenna is greatly reduced. From (2) it is seen that f_d is proportional to $\cos \phi$. With the values of ϕ , used in most

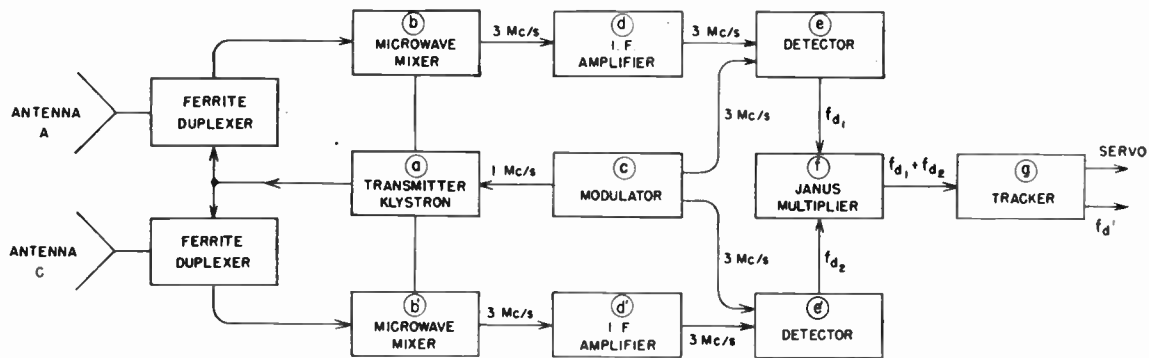


Fig. 4—A simplified block diagram of the Doppler radar system. (Several auxiliary parts of the system are omitted in the interests of clarity.)

Doppler navigation radars, a small change in ϕ produces a fairly large change in f_d . For example, for the system to be described a single beam develops a 1 per cent error for an 11 minute change in ϕ . However when two beams are used, ϕ can change by nearly 10° before a 1 per cent error develops. Antenna pitch stabilization requirements are therefore much less critical with Janus beams, and for some applications pitch stabilization can be omitted.

3) There is a first-order correction of Doppler spectrum spread due to short-term frequency variation of the transmitter itself. If a single beam is used and the Doppler spectrum obtained from a comparison of the return signal from the ground with a signal directly from the transmitter, noise FM of the transmitter causes a spread of the spectrum. This blurring of the spectrum increases with altitude and is very serious in some aircraft flying above 30,000 feet. Fig. 3 illustrates the effect. Fig. 3(a) shows an unusable spectrum obtained when using a single beam. The acceptable spectrum of Fig. 3(b) is obtained, under similar conditions, when both fore and aft beams are used.

It will be noted that this technique doubles the audio frequency at which a Doppler signal is observed. Thus the peak at approximately 8 kc in Fig. 3(b) would correspond with a peak at 4 kc in Fig. 3(a). This is treated in greater detail in Section D of Part III.

A. THE FM TECHNIQUE

The Doppler radar developed for this project uses an FM transmitter. It is instructive to review briefly the history of Doppler development, which reveals the reason for this choice. Early systems used search radar techniques with high peak power (kw) and short pulses ($\frac{1}{2}$ μ sec with a duty cycle of 1:1000).⁷ These were all very heavy systems of 300–400 pounds weight. The quest for lightweight systems led to more efficient use of the transmitted power. CW systems,⁸ where both transmitter and receiver are on the whole time, theoretically offer the most efficient use of transmitted power. However, such a system reveals serious practical problems. It is difficult to prevent energy from leaking directly into the receiver and swamping the ground return at high altitudes. The magnitude of the problem may be

appreciated when it is noted that 150 db of receiver isolation are required for operation at 50,000 feet. It is more convenient to use a system in which a direct leak of transmitter energy into the receiver is unimportant, and this is possible when (as here) the signal returns are delayed.

Two solutions to this problem which exploit such a time delay and do not involve a great reduction of efficiency are the long pulse system,^{5,6} where the transmitter is on for about one half of the total time, and the FM technique.² The latter method was adopted for the Canadian development because of the ease with which a klystron could be modulated and because it required no increase in the number of microwave operations by comparison with the CW system. Fig. 4 shows a block diagram of the radar, which is complete except for the servo-tuned "tracker oscillator" which produces a highly accurate representation of the Doppler frequency.

The system makes use of duplicate antenna channels A and C, which perform the Janus operation. The signals reaching each microwave crystal mixer consist of a local oscillator signal, which is a sample of the transmitted energy, and the Doppler shifted return from the ground. Both signals carry a sideband spectrum typical of FM at 1 mc (Fig. 5). The receiver channel has a pass band at 3 mc, and the transmitter modulation index is so chosen that the output of the microwave mixer has maximum possible energy concentration in this band. If the received signal has suffered a time delay which is negligible compared with the modulation period, as would be the case with transmitter leak via the duplexer or radome, there is no output from the microwave mixer at 3 mc, and the system is thus insensitive to such noise. However, if the received signal has traveled to the ground and back and, therefore, has been delayed, there will be a 3-mc output from the crystal mixers. Thus the system provides receiver isolation which proves to be equivalent to more than 150 db for a duplexer isolation of only 25 db.

That the system operates as described above can be seen by examining the result of mixing the transmitted and received signals in a multiplier circuit. Let

$$e_1 = E_0 \sin(\omega t + M \sin pt) \quad (3)$$

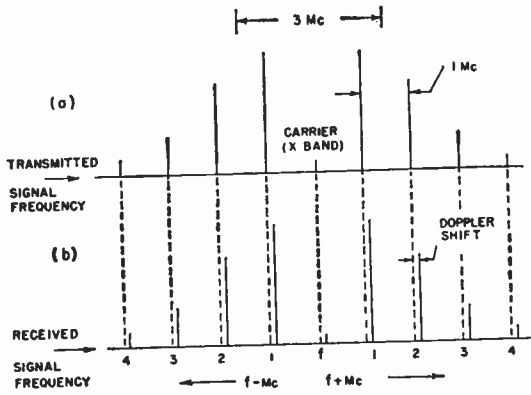


Fig. 5—Diagram of a microwave carrier, frequency f , frequency modulated at 1 mc, and its sidebands (above). Below is a diagram of the same signal after having received a Doppler shift.

be the transmitted signal and the form of the local oscillator signal and let

$$e_2 = E_0 \sin [\omega_1(t - t_1) + M \sin p(t - t_1)] \quad (4)$$

be the received signal, where

ω = carrier frequency (radians/second),

p = FM (radians/second),

M = modulation index $\left[\frac{\text{frequency deviation}}{\text{FM}} \right]$,

ω_1 = received frequency (radians/second),

t_1 = delay time $\left[\frac{2 \text{ range to ground}}{c} \right]$.

Then the difference term in the product ($e_1 \times e_2$) which appears at the receiver is

$$\frac{E_0^2}{2} \cos [(\omega t + M \sin pt) - \omega_1(t - t_1) - M \sin p(t - t_1)] \quad (5)$$

$$= \frac{E_0^2}{2} \cos [\omega t - \omega_1(t - t_1) + M [\sin pt - \sin p(t - t_1)]] \quad (6)$$

which, after expansion and rearrangement, becomes

$$\frac{E_0^2}{2} \cos \left[\omega t - \omega_1(t - t_1) + 2M \sin \frac{pt_1}{2} \left[\cos \left(pt - \frac{pt_1}{2} \right) \right] \right] \quad (7)$$

Eq. (7) represents an FM signal with a carrier frequency $(\omega - \omega_1)/2\pi$, the same FM as in the transmitted signal, and a modulation index given by $2M \sin pt_1/2$.

It might appear that the system would lose signal at altitudes corresponding to a time delay equal to the FM period. This has not been detected during flights over land, but recent overwater flights have demonstrated a very high signal attenuation at altitudes corresponding to the two lowest predicted "holes." Apparently, at higher altitudes the variation in range over the beamwidth masks such effects.

B. FLIGHT TRIALS

A series of flight trials has illustrated that the radar has sufficient accuracy to be the major input organ for a useful navigation system. This has been done using the system in Fig. 1, with the exception of a steering computer. The Doppler radar used a 60-mw transmitter, and the positional computer was based on simple analog techniques.

The trials have demonstrated that two of the objectives have been attained: 1) distance can be measured reliably to an altitude of at least 40,000 feet and 2) accuracies of 0.5 per cent in distance and 0.5° in drift can be attained easily.

Fig. 3(c), typical of a number of spectra, is taken with a transmitter power of 60 mw and shows that a 10-db SNR is possible at 40,000 feet. The system under construction uses a 600-mw transmitter (without any increase in size or weight) and this is expected to extend the operational altitude well above 60,000 feet.

The accuracy of the radar has been measured over various surfaces. Extensive measurements have been made along a marked course to obtain a reliable measure of the accuracy with which the system measures distance over the ground. The equipment has been flown over sections of electric power transmission line, and the distance as measured by the radar compared with the surveyed distance. The results of 100 measurements show a standard deviation of 0.2 per cent in the distance flown. It has not been possible to carry out a reliable determination of drift angle accuracy since there has been no sufficiently accurate reference for comparison. However, correlation to $\pm 0.5^\circ$ has been demonstrated by comparing average drift as measured by Doppler with the average of a number of drift sight readings taken by a navigator. It is probable that most of the error was contributed by the navigator and that the radar is considerably more accurate. These measurements were all made over land.

Accuracy determination over water has been done using the same general technique, but with a Decca navigation system to provide a measured line 80 miles long. Approximately 50 measurements have been made over the Atlantic Ocean southeast of Halifax, N. S. All these measurements have shown a significant difference between the distance measured when flying up wind and that measured flying down wind (reference is to winds at the water surface). The standard deviation of these measurements was 1.5 per cent. Attempts to correct the results for movement of the water due to winds and tides reduced the standard deviation only to 1.0 per cent. This is not surprising, since our knowledge of the surface winds existing at the time of the runs, and our understanding of the effect of wind on water movement, are both sketchy.

Two effects which may cause navigation errors when flying over water are known.⁴ One occurs over relatively smooth water and is due to an increase in specular reflection of the incident beam. This reduces the strength

of the signal received at the aircraft and possibly favors the lower half of the beam. The effective angle of depression of the beam (angle ϕ , Fig. 2) may thus be increased, producing a ground speed error. An error of approximately 0.5 per cent has been predicted theoretically for a beamwidth of 3° . The existence of this phenomenon could not be verified during these trials, because at no time was the sea smooth enough, and because the second effect, described below, was large enough to mask it.

The second effect is due to actual movement of the water in mass caused by tides or, at the surface, by winds. Any navigation carried out over such a medium,

using Doppler measured ground speed, will contain an error determined by the net movement of the water during flight.

It is proposed to examine this phenomenon further by correlating more detailed water movement information with the results of Doppler radar distance measurements.

From this series of flight trials, it may be concluded that the radar measures distance over land with an over-all accuracy of 0.2 per cent and probably measures drift with the same accuracy. However, larger errors are to be expected over surfaces which may be either smooth or moving, such as water.

Part III—The Circuit Implementation of the Radar

P. M. THOMPSON, N. F. MOODY, C. A. FRANKLIN, AND R. J. BIBBY

It will be recalled from Section A of Part II that the navigational radar is based upon the FM principle according to the block diagram of Fig. 4. The circuits are now to be described, with the exception of some auxiliaries not considered of sufficient importance to warrant inclusion.

Apart from the transmitter klystron, no vacuum tubes are used in the instrument to be described: all circuits are based on transistors. The design has been directed to make the performance insensitive to transistor parameter variations, so that the accuracy depends primarily on passive elements. Every attempt has been made to attain the highest reliability, and to this end the design is directed to eliminate certain components of questionable life, such as electrolytic capacitors.

A. THE MODULATOR

The modulator c in Fig. 4 performs two functions. It generates a 1-mc sinusoidal signal to frequency modulate the klystron \textcircled{a} , and it provides a 3-mc carrier which is combined with the IF signal at the detectors \textcircled{b} and \textcircled{c} . A circuit fulfilling these requirements consists of a 1-mc crystal controlled square-wave generator driving a high-pass and a low-pass filter in parallel. The two filters operate in such a way that the low-pass filter selects the 1-mc fundamental component of the square wave to modulate the klystron, while the high-pass filter passes the remainder of the frequency components, the strongest of which is at 3 mc, to feed the detectors.

The chief requirements of the modulator are these: 1) the amplitude of the 1-mc modulating signal shall be accurately defined, for upon this level depends the FM sideband distribution; 2) there shall be a fixed harmonic and phase relationship between the 1-mc and 3-mc signals; and 3) the frequency stability must be such that the received signals fall properly within the pass band of the IF receiver. A conventional LC resonant circuit would probably provide sufficient stability, but the use

of a quartz crystal permits an extremely simple and lightweight circuit of excellent performance and built-in accuracy.

The requirements for accurate amplitude control and constant harmonic relationship are conveniently met by generating a 1-mc square wave, passed through filters to provide the necessary output signals. The amplitude of the 1-mc fundamental component is defined accurately when the square wave amplitude is defined. Harmonic relationship between the 1-mc and the 3-mc output is a direct result of their common source.

A circuit diagram of the square wave generator is shown in Fig. 6, and operates in the following manner. The emitter resistors R_2 and R_3 are equal, and so set equal bias currents in transistors J_1 and J_2 . The value of the emitter current is given by

$$I_e \sim \frac{20 - V_0}{R_2} \quad (8)$$

Because $(20 - V_0)$ is so much greater than the emitter-base voltage drops of J_1 and J_2 , or the forward voltages of diodes D_1 and D_2 , the average current of each transistor is still defined by (8) when the circuit is oscillating.

A positive feedback loop exists via the frequency controlling crystal X , which connects the collector of J_1 to the base of J_2 , and by means of the condenser C_1 which effectively provides a short circuit between emitters at the frequency of oscillation. At this fundamental frequency the crystal exhibits series resonance, and is of low impedance, while the circuit LC , tuned near this fundamental, is of high impedance. If the diodes D_1 and D_2 are ignored for the moment, it is seen that the loop gain is high and oscillation will occur. Spurious oscillations, on the other hand, are not possible; for at frequencies above or below the fundamental, the reactance of the crystal rises, while the tuned circuit imposes simultaneously a shunt from the feedback loop to ground.

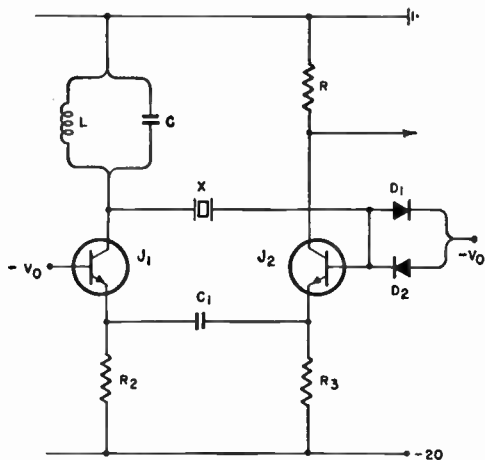


Fig. 6—The circuit diagram of the prime generator of the modulator. (A crystal controlled square wave oscillator.)

The resulting 1-mc oscillation causes the total transistor currents $[2I_e$ of (8)] to commute between J_1 and J_2 . There results a 1-mc square wave of current at the collector of J_2 of peak-to-peak amplitude

$$I_s = \frac{2(20 - V_0)}{R_2}, \tag{9}$$

which is accurately defined. Hence an accurately controlled square wave of current will flow in the resistor R , terminating the filters. It follows that the outputs of the two filters also have accurately defined amplitudes.

The diodes D_1 and D_2 are silicon junction diodes, whose threshold of forward conduction (~ 0.6 v) is sufficiently large for the transistor commutating action to remain undisturbed by their presence. The diodes serve merely to limit the voltage excursions at the base of J_2 , and also to supply a low-impedance path for any crystal current in excess of its requirements.

The filters are complementary high- and low-pass units, sharing the common source load, R . Their design is illustrated in Fig. 7(a) and 7(b), respectively. The amplitudes of the respective outputs from these filters, the fundamental and third harmonic, are readily computed from the Fourier expansion of an ideal square wave. Thus the RMS voltage of the fundamental is

$$V = \frac{2 \times 2}{\pi} \frac{R(20 - V_0)}{R_2} \times 0.707, \tag{10}$$

which depends only on passive components.

The performance of the modulator easily exceeds the specification demanded by the radar. Within the temperature range -55°C to $+70^\circ\text{C}$ the frequency is stable to within 30 cycles, and the level of the 1-mc output is constant to within ± 1.5 per cent. Third harmonic, which must be rigorously excluded from the 1-mc output, is 60 db down. Little attempt is made to remove harmonics from the 3-mc signal to the mixer, for these are not harmful.

Since the harmonic content of the oscillator is influ-

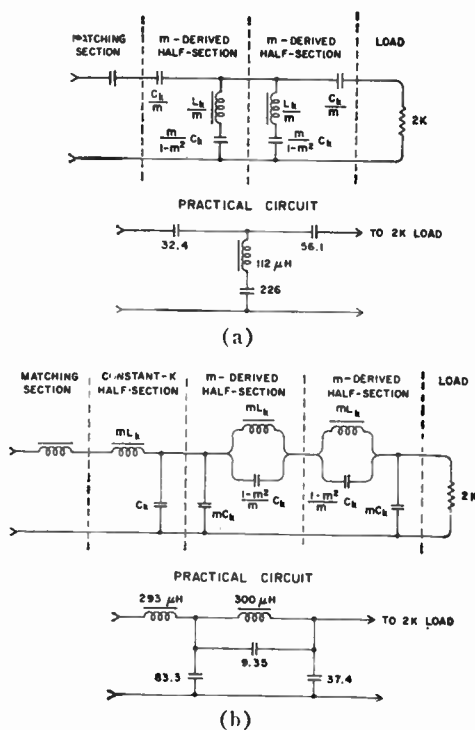


Fig. 7—The high-pass (a) and low-pass (b) filters of the modulator.

enced by temperature, this output is less stable in amplitude, but adequate for the purpose required.

B. THE IF AMPLIFIER DESIGN

Three designs for the 3-mc IF amplifier exist. One model uses germanium alloy transistors whose f_{avo} is approximately 10 mc, another is based on surface barrier transistors, and the third uses drift transistors. The designs, using both surface barrier transistors and drift transistors, have undergone the full series of environmental tests demanded of the over-all equipment. However, the design using drift transistors has met the target noise-figure specification of 2 db, so this design is described.

The specifications for the IF amplifier call for an over-all amplification of approximately 100 db, and a bandwidth in the order of 100 kc. It will tolerate a temperature range from -55°C to $+70^\circ\text{C}$, and contain provision for automatic gain control (AGC). The actual controlling circuit, part of the Janus unit, is called upon to hold the signal level, at the appropriate point in its circuit, constant within ± 5 per cent for variations at the input of the IF amplifier of 40 db.

The amplifier has five stages capable of a total unilateralized gain of 220 db. However, this is reduced to 100 db by stabilizing and coupling losses. Fig. 8 (left) illustrates a typical stage of the amplifier. The transformer (T_1) is single-tuned because there are no stringent requirements on the shape of the pass-band response, thus simplifying the method of neutralization (R_5C_5). The ac and dc stabilization of the circuit are controlled by R_1 , R_{1a} and R_4 , respectively. The collector voltage is

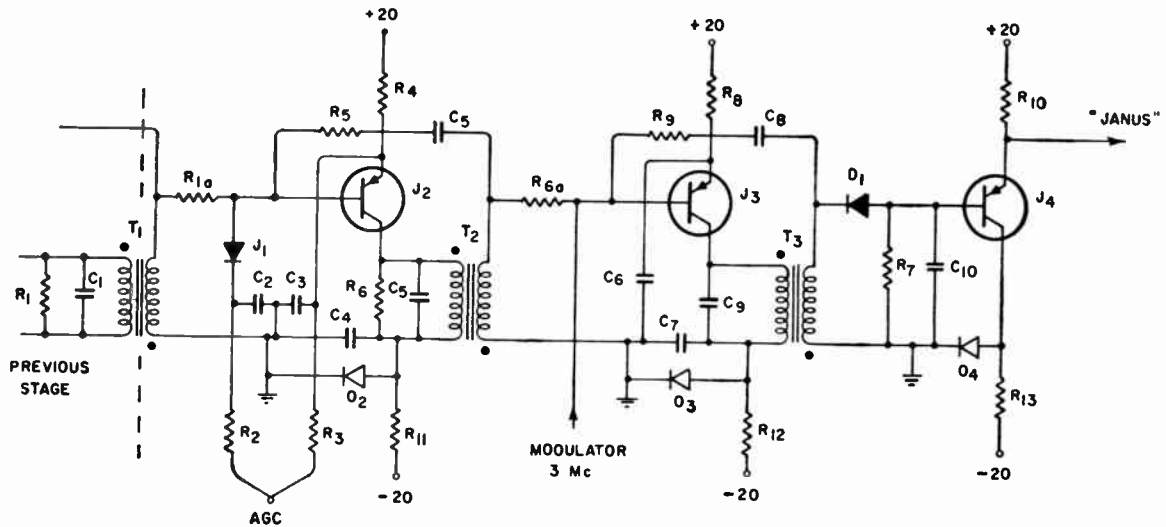


Fig. 8—The last two stages of the IF amplifier, the detector, and output stage.

set at -12 by R_{11} and avalanche diode D_2 . The AGC,¹³ applied only to the center three stages, is arranged to cause negligible detuning of the receiver. The circuit consists of the network $R_2R_3C_2C_3J_1$, and operates as follows. If a negative potential is applied to the AGC terminal, currents flow in R_2 and R_3 . The current in R_2 causes a forward current to flow in the diode J_1 , thus decreasing its incremental impedance, and the current in R_3 subtracts from the emitter current of transistor J_2 , increasing its base input impedance. For the signal frequency, J_1 is connected in parallel with the input terminals of J_2 via C_2 and C_3 . Thus, as the AGC voltage is increased, the signal is diverted from the base of J_2 to J_1 , reducing the stage gain. The values of R_2 and R_3 are chosen so that the decrease of impedance of J_1 approximates the increase of input impedance of J_2 . Thus there is little change in input impedance of the stage, hence little detuning of the IF due to AGC.

C. THE 3-MC DETECTOR

The output of the IF amplifier consists of a 3-mc signal carrying a Doppler shift of $\pm f_d$. The actual Doppler spectrum is extracted by injecting a 3-mc carrier from the modulator and rectifying the combined signal in a conventional diode detector. Fig. 8 (right) illustrates the last stage of the IF amplifier J_3 into which the 3-mc carrier, the envelope detector D_1 and the common collector output amplifier J_4 are injected. The last stage of the IF amplifier has no provision for AGC, because the AGC circuit holds the audio signal at the Janus multiplier constant. If there were AGC on this stage, the proportion of carrier energy to sideband energy would not remain constant and there would be conditions where either the IF amplifier stage would overload, or there would be insufficient carrier for undistorted envelope modulation.

¹³ C. R. Hurtig, "Constant resistance AGC attenuator for transistor amplifiers," IRE TRANS. ON CIRCUIT THEORY, vol. CT-2, pp. 191-196; June, 1955.

The detection is performed at a fairly high impedance, so the common collector amplifier, J_4 , is employed to provide an output impedance low enough to drive the audio frequency amplifier in the Janus unit to be described next.

D. THE JANUS MULTIPLIER

It will be remembered that the ground speed is determined from a Doppler spectrum derived in the Janus system from both the fore and aft beam system of the antenna. In normal operation, the axis of the antenna system will have set itself in line with the ground track (Fig. 2). Then the forward beams AA' may be treated as a single beam, since each bears identical information. The rear beams CC' are treated in the same manner. This pair of single beams, each handled by a separate receiver, is shown as A and C of Fig. 4. The function of the Janus unit is to perform instantaneous multiplication of the two receiver output waveforms, so generating a composite Doppler spectrum.

1. General Considerations in the Janus System

Consider the instantaneous product of the two sine waves $A \cos \omega_1 t$, and $B \cos \omega_2 t$

$$(A \cos \omega_1 t)(B \cos \omega_2 t)$$

$$= \frac{AB}{2} [\cos (\omega_1 - \omega_2)t + \cos (\omega_1 + \omega_2)t]. \quad (11)$$

The product is seen to contain only the sum and the difference frequencies. The result may be extended to cover the multiplication of two input spectra such as the two detector outputs of Fig. 4. Then the output spectrum contains all the sum and difference components of the frequencies existing in the detector outputs. This output feeds the tracker which measures the center frequency of the sum components and this center frequency represents the aircraft ground speed by (2).

It is convenient to include the audio frequency amplifiers and AGC circuits as part of the Janus unit, so that

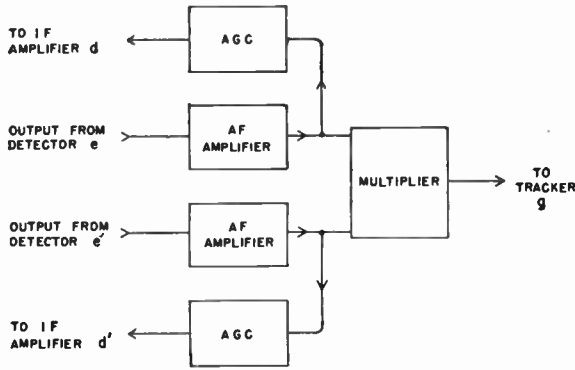


Fig. 9—Block diagram of the complete Janus unit, including audio amplifiers and AGC.

the connection between the second IF detectors and the multiplier represents the division between the high-frequency and the audio-frequency sections of the radar. The AGC voltage is derived within the multiplier because the function of the AGC system is to control the signal at the input of the multiplier proper, thus holding the input to the tracker at a constant level. Placing the AGC here eases problems of gain stability of the audio amplifiers because it places them within the control loop. Fig. 9 shows a block diagram of the complete Janus unit. First the audio-frequency amplifier and AGC are described, and then the multiplier.

2. The Audio Amplifier

The audio-frequency amplifier must supply a 2-ma peak signal from a high impedance to the multiplier, whose input is of low impedance. The accuracy of the ring modulators, the heart of the multiplier, is dependent on the amplifier impedance being several orders higher than that of the modulators. An amplifier output impedance of 15 k Ω or higher is preferable. A further requirement is that this amplifier shall have a fairly stable gain, independent of transistor changes, so that the signal level at the detector will remain fairly constant.

The circuit of the audio-frequency amplifier [Fig. 10(a)] shows how the problem of obtaining a high output impedance is solved. The output is taken from the collectors of a *p-n-p* and an *n-p-n* transistor, J_2 and J_1 , which are connected in series, thus eliminating the need for a collector load resistance, which would itself make the output impedance undesirably low. As the collector impedances are much higher than R_3 and R_4 , these resistances, in fact, set the output impedance of the circuit. Capacitors C_1 and C_2 are of low impedance compared with R_3 and R_4 at the frequency of operation and connect the base of J_1 to the detector output, an emitter follower, and the base of J_2 to the +20-v rail, both of which are also of low impedance. It follows that, at the frequency of operation, there is little feedback to the bases of J_1 and J_2 via R_3 and R_4 , so that the output impedance of the amplifier approximates the parallel values of R_3 and R_4 . The output current is determined by the input voltage and by the value of the resistor R_1 . It is

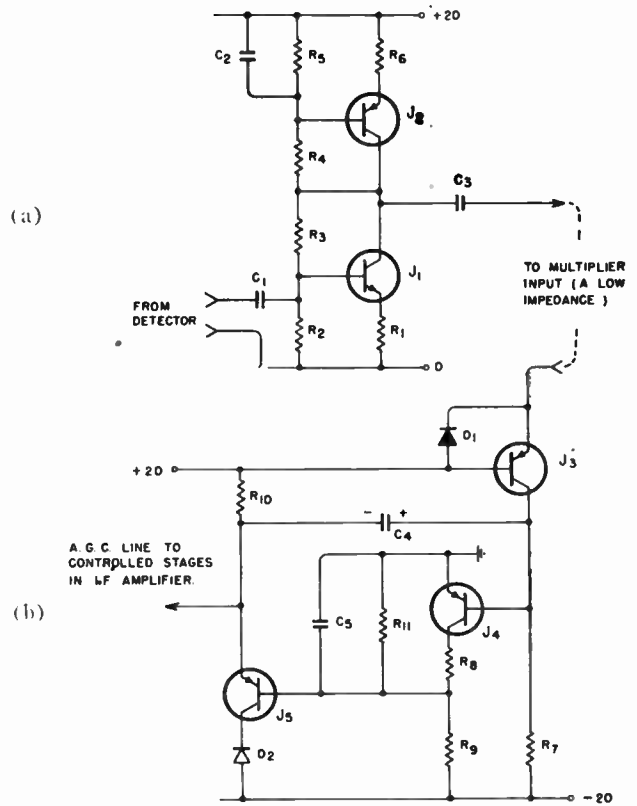


Fig. 10—Circuit of (a) one of the audio-frequency amplifiers and (b) one of the AGC units.

thus essentially independent of variations in the parameters of the transistors J_1 and J_2 . The resistive networks stabilize the dc operating currents of the transistors, which they set to $I_c = 3$ ma. The values are also chosen to permit variations in total transistor base current ($I_b - I_{co}$) of $\pm 100 \mu a$.

A design feature of this circuit lies in the high impedances at both the input and output terminals. Although relatively low audio frequencies must be transmitted, the use of electrolytic coupling condensers is thereby avoided.

3. The AGC System

The AGC circuit controls the mean level of the signal current at the output of the audio-frequency amplifier as follows. The return current from the multiplier is rectified at the emitter of J_3 [Fig. 10(b)], and the rectified current at the collector is compared with the current in R_7 which constitutes the AGC delay. An integrating amplifier J_4J_5 applies AGC voltage to the IF amplifier when the rectified signal current exceeds the delay current. The potential at the base of J_4 is near ground potential, so the current in R_7 (100 k Ω) is approximately 200 μa . If the rectified signal current (added to the difference between the I_{co} of J_3 and J_4) is greater than this, the output of J_5 must go negative, thus reducing the IF gain. The gain control has a long time constant and so does not remove rapid variations in signal (which would interfere with the operation of the

tracker). It is necessary, therefore, to employ a large condenser in the integrating amplifier ($C_4 = 25 \mu\text{f}$). The circuit of the integrating amplifier is a conventional common emitter stage followed by a common collector stage. The resistor R_8 is added to maintain the correct polarity of dc bias on C_4 , which must be an electrolytic capacitor, and the capacitor C_5 insures ac stability. The avalanche diode D_2 and resistor R_{11} keep the collector dissipation of J_5 down to an acceptable level.

This AGC circuit holds the input to the multiplier constant within ± 5 per cent over the full 40-db variation at the input of the IF amplifier.

4. The Modulators

In both the multiplier and tracker units considerable use will be made of circuits based on the ring modulator. It will facilitate understanding of these units if the classical properties of the ring modulator are clearly in mind. A short digression will therefore recall these and describe new properties which are a logical extension from them. The classical implementations of the ring modulator are directed to transfer maximum power from signal source to output load. When power efficiency is the design criterion, then the rectifiers of the modulator must all have matched characteristics. In a new operating mode, in which this criterion is subordinate (but which permits most of the classical functions to be retained), it is shown that the modulator automatically exhibits a precise balance without rectifier selection.

The ring modulator circuit is shown in Fig. 11 where there are three sets of terminals, aa_1 , bb_1 and cc_1 . Any pair of these sets may be used as input terminals and, furthermore, there is no direct transmission of signal from any one set of terminals to any other set. In a perfectly balanced bridge it is only the resultant of the signals on two sets of terminals which is transferred.

Some of the classical operating modes are as follows.

1) Use of the bridge as a switch controlling signal polarity: When a periodic waveform is applied to the terminals aa_1 and a dc signal is applied to bb_1 , then the output signal at terminals cc_1 may be a replica of the input waveform at aa_1 but its instantaneous polarity is controlled by the sense of the dc signal. That is, the bridge provides phase reversal facilities under control of a dc polarity. If the input to bb_1 is a square wave then the output polarity will be switched by the square wave.

2) Use of the bridge as a phase-sensitive rectifier: Suppose two sinusoidal waveforms of the same frequency be applied to aa_1 and cc_1 . Then the device acts as a phase-sensitive rectifier, this producing dc at the terminals bb_1 . The amplitude of this dc depends on, among other things, the relative phases of the input waveforms. For square wave inputs the amplitude of dc becomes linearly related to the input phase difference.

3) Use of the bridge as a suppressed carrier modulator: For two input waveforms of differing frequency, applied as in 1) or 2), the output waveform does not con-

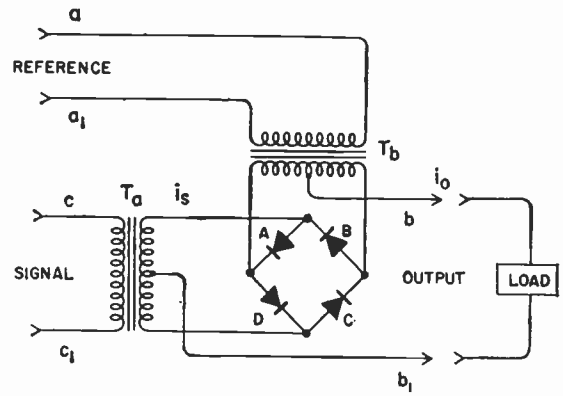


Fig. 11—Ring modulator. (The labels on the terminals, *i.e.*, signal reference and output, refer to operation in mode 2) as described in Section D-4 of Part III.)

tain input frequency, but instead gives the instantaneous product of the input waveforms. This is suppressed carrier modulation.

All of the above modes will be used in the multiplier and tracker units and, furthermore, the modulators will all operate under the condition of square wave excitation of one of the inputs.

5. Current Operation of the Modulator

It will now be shown that when one of the input waveforms is a square wave and the input and terminating impedances of the modulator are made infinite and zero, respectively, then the operation of the modulator reduces to that of a simple switch.

In these circumstances, a high degree of modulator balance is inherent. The operating principles will be described in terms of mode 2) above, but are equally applicable to the other modes.

Suppose that the input terminals (Fig. 11) are supplied from constant current sources: aa_1 with a square wave (to be called the reference) and cc_1 with a smaller current (the signal). The output is to be taken as the current in terminals bb_1 , which are short-circuited.

Then the reference wave, considered alone, will drive one rectifier pair AB (or CD) into conduction, and the alternate pair CD (or AB) will be biased to cutoff by the forward voltage generated across the conducting pair. It follows that any signal current injected into T_a can leave the transformer only by one of the secondary winding halves, and that this winding is selected by the reference polarity at that instant. Thus, when the rectifiers AB are conducting, the upper winding of T_a supplies signal current i_s , which passes to both halves of T_b secondary in opposition, to the output terminal where it is shown as i_o . It is seen that the rectifiers perform a simple switching action between the secondary windings of T_a , as determined by the reference. This switching action cannot be disturbed by the sense of the signal current, nor by its magnitude, for by definition the maximum value is insufficient to overcome the reference current. It follows that the conducting pair of diodes is still forward-biased and that the nonconducting pair therefore receive reverse bias.

If the transformers are lossless and if the reverse conduction of the cutoff diodes is zero, the whole signal current is transferred without loss to the output. Thus the diodes act as ideal switches simply routing the signal current under control of the reference. The cutoff diodes do in fact pass a small current under the influence of the small reverse voltage impressed on them by the voltage drop of the conducting rectifiers. This constitutes an error and is the main source of unbalance in the modulator. The percentage balance error may be made very small, for at maximum signal level its magnitude is simply the ratio of the leakage current to the reference current. This will depend on the type of diodes which will be chosen according to the intended purpose of the modulator. When the waveforms are of relatively high frequency, germanium diodes are preferred because of their excellent high-frequency characteristics. Then the ratio of currents may be in the order of a few thousands. For low-frequency work, the silicon junction diode can raise this ratio above one hundred thousand.¹⁴

At the frequencies used in this paper, transformer balance is controlled simply by the turns ratio. Transformer losses do not directly disturb the balance, but do make the input and output impedances finite. In a practical system the input and output impedances must necessarily depart from the ideal, but it is not difficult to make the source impedances several orders greater in magnitude than the output impedance. This does not significantly degrade performance. Care must be taken that the voltage developed across the load is not excessive, for if it is it can be shown that the bias across the nonconducting diodes will be overcome and a limiting action will ensue.¹⁴ Correct design allows signal currents to be modulated with a negligible transmission loss, and the modulators described in the multiplier have balances such that the output current at no signal is always but a few parts in a thousand of the output at full signal. The accuracy of the whole system depends on the inherent precision of these modulators.

6. The Multiplier¹⁵

The multiplication of the two Doppler spectra developed by the channels *A* and *C* for Fig. 4 is performed in a ring bridge modulator. Some transformation of signal form is required before this can be done, and hence the description will be divided into two parts. First the multiplier modulator will be described, followed by a description of the signal transforming circuits.

Suppose the phase-sensitive rectifier of Fig. 11 to be supplied with reference and signal current sources as shown. Each of these is to be in the form of a square

wave of frequency f_0 , and the amplitude of the signal source is never to exceed that of the reference. The operation depends on mode 2) of Section D-4, so that the output will be linearly related to any phase difference between the input waveforms and also to the amplitude of the signal. When square waves are used in this way, it will be recalled that the diodes act as switches, and the output is very accurately related to the phase difference and signal amplitude. The output current is given by

$$i_0 = i_s \frac{2}{\pi} (\theta - \pi/2),$$

where i_s is the amplitude of signal current at the secondary of T_a , and θ is the phase difference between signal and reference waveforms. For the present purposes it will be convenient to measure phase angle such that zero angle represents a quadrature relationship between input waveforms. Thus we introduce the new angle ϕ where

$$\phi = \theta - \frac{\pi}{2}. \quad (12)$$

It follows that

$$i_0 \propto i_s \phi \quad (13)$$

so that i_0 is zero when either i_s or ϕ is zero. The relationship of (13) may be made the basis of a multiplier in which i_s and ϕ are the analogs of the quantities whose product is desired.

The manner in which these analogs are developed will now be described by means of Fig. 12, which shows the complete multiplying system. It is seen to consist of the phase-sensitive rectifier, described above, whose inputs are developed from two further modulators, each of which receives a reference square wave of the same frequency f_0 . The output of one Doppler receiver, i_1 , is made to phase modulate the square wave f_0 and so generates the function ϕ of (13). The other Doppler output, i_2 , is caused to amplitude modulate its reference f_0 and so generate a square wave i_s of (13). Thus the output of the phase sensitive rectifier represents the product of the two Doppler inputs as required.

These preliminary modulations of the two receiver outputs are accomplished in the following manner. The modulation of i_2 simply makes use of mode 3) of Section D-4 and therefore needs no further description. The phase modulation of i_1 is a little more complex, as is seen from Fig. 13. A first step is to perform suppressed carrier modulation on i_1 , as was done for i_2 , using the same reference frequency f_0 . This results in the waveform \odot . Then the reference square wave \odot is integrated to generate waveform \ominus , whose fundamental is displaced $\pi/2$ radians to introduce the quadrature relationship required by (12). The suppressed carrier waveform \odot and the integrated reference \ominus are next added to produce the composite waveform \oplus . The required waveform, in which i_1 is represented by ϕ , results simply by the generation of

¹⁴ N. F. Moody, "A silicon diode modulator of 10^{-8} A. sensitivity for use in junction transistor dc amplifiers," *Electronic Eng.*, vol. 28, pp. 94-100; March, 1956.

¹⁵ R. J. Bibby and P. M. Thompson, "A high-speed analog multiplier with a linearity of better than $-\frac{1}{2}$ per cent," presented at the IRE-AIEE Transistor and Solid-State Circuits Conf., Philadelphia, Pa.; February, 1958. See pp. 32-33 of Digest of Tech. Papers.

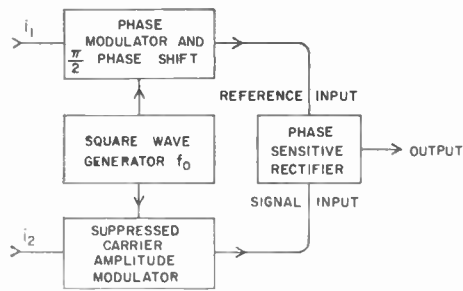


Fig. 12—Block diagram of the multiplier.

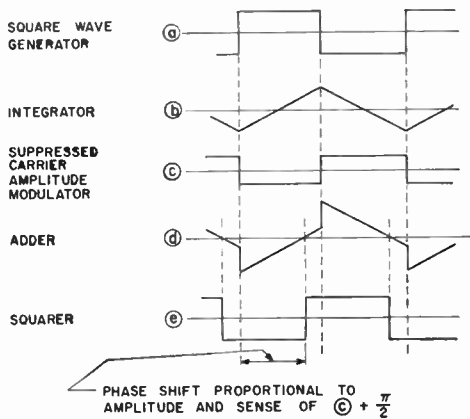
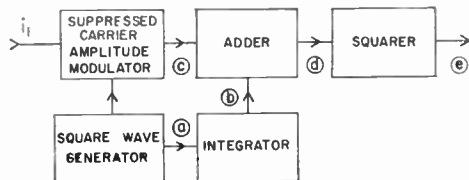


Fig. 13—Block diagram and waveforms of the phase modulator. N.B. Although the waveforms illustrate a case where the modulating signal is a dc, the technique is valid at frequencies up to half the carrier frequency. At these higher frequencies \odot will no longer be a true square wave.

a square wave \odot whose edges correspond to the times at which the composite waveform \odot crosses the zero axis. The waveform \odot is of constant amplitude and becomes the reference of the following phase-sensitive rectifier, in which the multiplication is carried out. The suppressed carrier waveform resulting from i_2 similarly becomes the signal source to the multiplying modulator.

It may be shown that the over-all system multiplies the input functions i_1 and i_2 as follows. When i_2 is zero the phase sensitive rectifier of Fig. 12 has no signal input and, hence, delivers no output. Conversely, when i_1 is zero the reference of the phase-sensitive rectifier is in quadrature with any waveform at the signal terminals, and again there can be no output. When $\pm i_1$ is present the reference is displaced $\pm \phi$ from quadrature, so that in the presence of signal there will be an output, dependent on the sign and amplitude of i_1 . Similarly, the output depends also on the amplitude of i_2 and includes the sense of i_2 since the signal waveform inverts when i_2 changes polarity.

The complete multiplier is thus seen to consist of a phase-sensitive rectifier, two suppressed carrier modulators, a square wave reference generator, integrator, adder and squarer. The phase-sensitive rectifier and suppressed carrier modulators all follow the current operated technique described in Section D-4, and so yield an accurate and predictable performance. The remaining circuits are orthodox, and it is simple to design them so that they produce little error.

The carrier frequency chosen for these operations is 50 kc, because the input Doppler spectra extend to 20 kc so the output spectrum contains frequencies up to 40 kc. It is preferable that the carrier frequency be more than twice this highest input frequency. Because of the frequency of operation, switching speed becomes a factor in the choice of diodes for these circuits. For the diodes chosen, the unbalance current was less than $1 \mu a$, while the maximum value of signal current chosen was 2 ma.

A complete circuit of the multiplier system is shown in Fig. 14. The two suppressed carrier amplitude modulators and the phase-sensitive rectifier are shown as A, B, and C, respectively. The square wave carrier for the modulators A and B and for the integrator is provided by the multivibrator J_1J_2 . This multivibrator is of a conventional cross-coupled design, whose resistance networks R_1C_1 and R_6C_3 are the prime frequency-controlling elements. The output square wave is controlled in amplitude by the emitter network $R_2C_2R_4$ by a process similar to that used in the modulator of Section A. The output current is taken by means of three current transformers connected in series $T_1T_2T_3$. T_1 and T_2 become part of the two suppressed carrier modulators, and T_3 drives the integrator.

The integrator consists of a current-limiting circuit and an integrating amplifier. The purpose of the limiting circuit $R_7R_8D_9D_{10}D_{11}D_{12}$ is to provide an accurately controlled square wave of current for the integrating amplifier. The integrating amplifier consists of a common emitter stage J_3 , followed by a common collector output stage J_4 . The constant of integration is controlled by the feedback capacitor C_5 . It will be remembered that the integrator produces a triangular waveform which must be added to the output of the suppressed carrier modulator A. These waveforms are added as currents at the emitter of J_5 , whose low impedance satisfies two requirements. It provides a satisfactory termination for the ring modulator. It also allows accurate addition of the two waveforms, since its impedance is much lower than either the reflected impedance of the ring modulator or, R_{13} the predominant part of the output impedance of the integrator.

The remaining transistors, $J_5J_6J_7J_8$, comprise the squaring amplifier. J_5 is a common base stage, followed by a common collector stage, J_6 , and an emitter coupled pair, J_7J_8 . R_{18} controls the amplitude of the output current of this last stage which drives T_6 , the reference input, to the phase sensitive rectifier C.

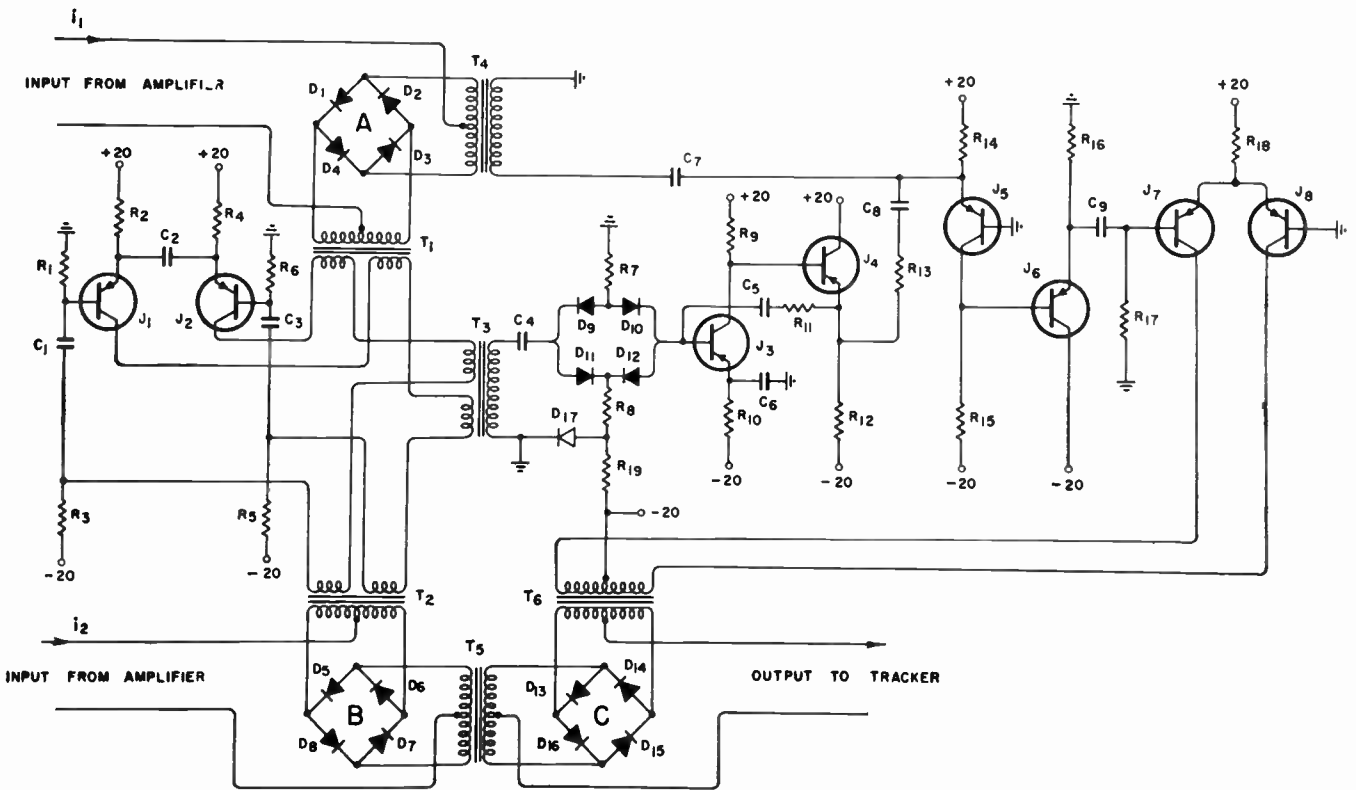


Fig. 14—The multiplier circuit.

Note that the output signal of the suppressed carrier modulator *B* is applied via *T*₅ as an input signal to the phase-sensitive rectifier *C*. Direct cascading of modulators in this way is often possible but requires some care in design. It is permissible for this reason. If the connections from the secondary of *T*₅ are lifted from *C*, observe that the impedance seen at the transformer terminals is in the order of the source impedance of *B*. Such a high impedance is appropriate for feeding modulator *C*. Conversely, if the primary leads of *T*₅ are lifted from *B*, it will be found that the transformer impedance is in the same order as the output load on *C*. Such a low impedance provides a suitable termination for *B*.

This multiplier proves to have excellent stability and linearity (± 0.3 per cent) over a temperature range extending from -40° to $+70^\circ\text{C}$. Accuracy relies only upon a few passive components and upon there being sufficient gain for the feedback and squaring amplifiers to work correctly. The over-all Janus unit proves capable of supplying the correct output spectrum [C of Fig. 15(a)] at a level constant within ± 15 per cent over the whole range of AGC control. Interference between the two channels is approximately -35 db.

E. THE FREQUENCY TRACKER

The multiplier delivers to the tracker a noisy spectrum of the type shown in Fig. 3(b). This spectrum contains the sum and difference frequencies of the Doppler signals derived from the fore and aft antenna systems. Its idealized form is shown in Fig. 15(a) where the peak

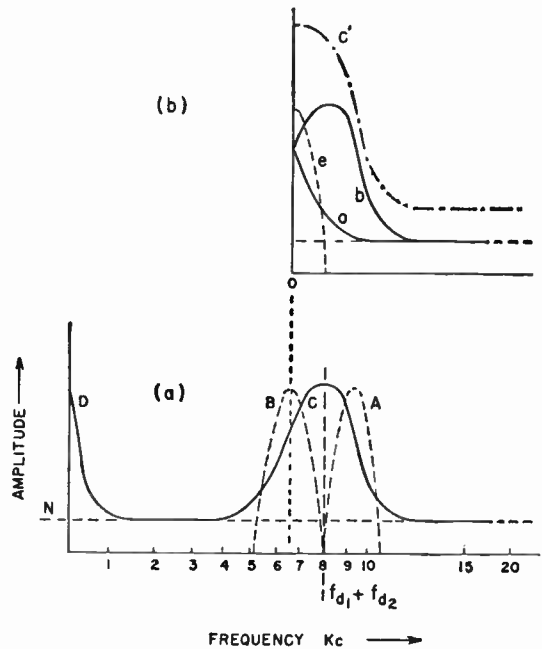


Fig. 15—(a) The form of the input spectrum for the tracker. (b) The spectrum resulting from modulating the above with frequency, f_b .

due to the sum frequency ($f_{d1} + f_{d2}$) is shown at *C*, and the difference frequency ($f_{d1} - f_{d2}$) at *D*.

The purpose of the tracker is to determine the sum frequency by examining the mean spectral energy of the peak *C* in a pair of similar frequency gates *A* and *B*. Should the gates deliver unequal energy, an error signal

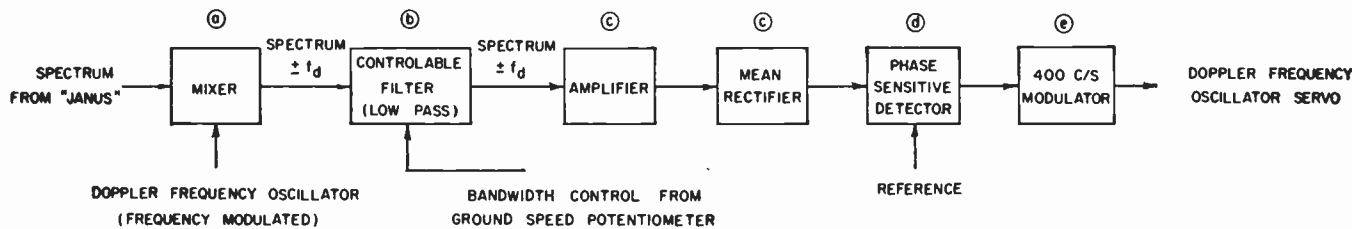


Fig. 16—Block diagram of the tracker.

arises which is made to displace the two gates bodily along the frequency axis until balanced energies result. By this means the center frequency of the gates is made to coincide with the required Doppler frequency. In fact, both gates are developed from an oscillator (called the tracker oscillator) whose mean lies at this gate center frequency. Thus the error signal is simply caused to servo the oscillator frequency until a null is found, whereupon the oscillator itself represents the required Doppler frequency.

The discussion of the tracker will commence with the spectrum to be measured, which will lead naturally to the choice of the gate bandwidths. It will conclude by outlining the practical embodiment of the device.

1. The Spectrum and Its Analysis

The spectrum to be measured consists of a background of fairly white noise N [Fig. 15(a)], on which is superimposed the sum peak C ¹⁶ and the difference peak D . The difference components give rise to frequencies lying outside the working band of the tracker, which can therefore be designed to ignore them. The sum components give rise to a noisy peak whose spectral energy distribution may be approximately defined by a Gaussian curve the width of which, to half energy points, is about 10 per cent of the center frequency.

The characteristics of the frequency gates must be so chosen that optimum use of this signal is made in the presence of the noise. To this end the center frequencies of the gates are placed at a separation corresponding to the points of maximum change of energy density on the Gaussian curve, at the half-power points. The width of the gates should be chosen to yield maximum signal-to-noise power, which leads to the choice of a bandwidth equal to one half of that defined by the half-power points of the Gaussian curve.

Because the absolute width of the Gaussian curve, expressed in cps, is proportional to the ground speed (and so to the tracker oscillator frequency), both the frequency separation, and width, of the gates should be made variable. These parameters are therefore made proportional to the tracker oscillator frequency.

2. The Method of Frequency Gate Generation

The frequency gates require a pair of variable width band-pass filters, whose widths and attenuations must

be carefully balanced at all times. It would be difficult to make such filters satisfactory under the environmental conditions. The problem of matched filters is overcome by time multiplexing a single filter between positions A and B [Fig. 15(a)]. If the energy in the filter for the time that it is in position A is compared with that received when it is in position B , the difference is the desired error signal. In the artifice (to be described) performing the above function the tunable band-pass filter also becomes replaced by a tunable low-pass filter which is much simpler to design.

The operation is most easily understood from a description of how one single-frequency gate, such as B , is formed. Let the input spectrum C from the Janus unit be modulated by a frequency f_B corresponding to the center of the desired frequency gate B [Fig. 15(a)]. Then the usual sum and difference frequency components will result. Considering only the difference frequencies, the resultant spectrum will have a component at zero frequency corresponding to any components of C lying at f_B . All components of C to the left of f_B will be reproduced to the same frequency scale as before, but with the frequency f_B subtracted from the abscissa scale, as shown in Fig. 15(b), curve a . Similarly, the components of C lying to the right of f_B will give rise to components b of Fig. 15(b), so that the output of the modulator will be the sum of a and b as shown in C' , which now becomes the converted spectrum of C . A low-pass filter of characteristic e acting upon C' now behaves in a similar manner to a band-pass filter B of twice its width acting upon the original spectrum C of Fig. 15(a). In a similar manner, the injection to the mixer of a frequency f_A , corresponding to the center of filter A , results in a frequency gate A being formed. The sum components arising from the mixer play no part in the operation for they are eliminated by the low-pass filter.

In order to form the two gates in a time multiplex arrangement, the two frequencies f_A and f_B must be generated alternately. This is accomplished by frequency modulating the tracker oscillator with a square wave at 25 cps. A block diagram of the system is shown in Fig. 16, where the output of the amplifier contains the energy represented by filter position A for half the time and the energy of filter position B for the other half. This output is rectified in a mean rectifier and then passed to a phase-sensitive rectifier whose reference is the 25-cps square wave used to modulate the tracker oscillator. The output of this final phase-sensitive rectifier will be shown

¹⁶ By peak is meant a spectral region of high energy density.

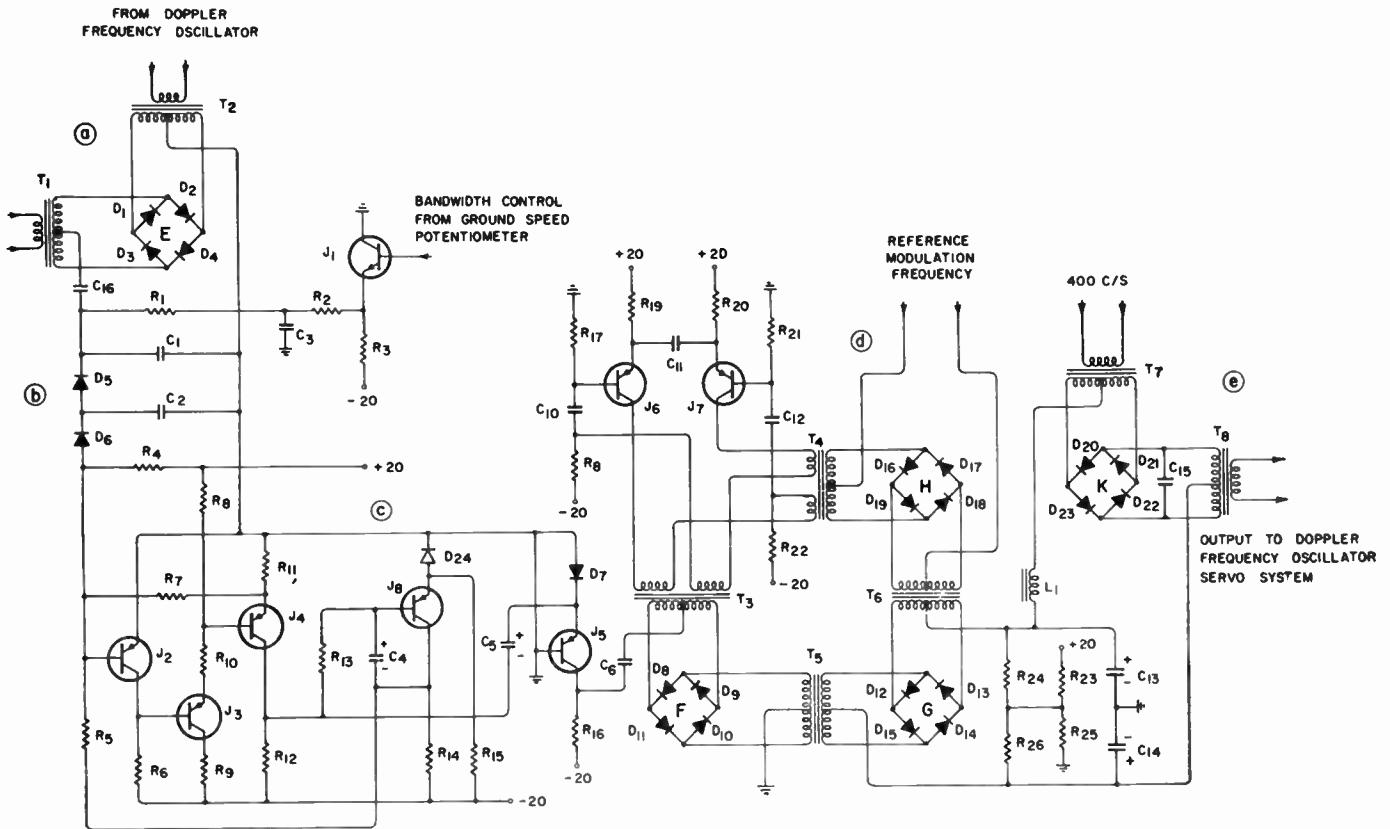


Fig. 17—Complete circuit of the tracker. (The letters relate parts of the circuit to the blocks of Fig. 16.)

to be a dc signal whose sign and amplitude are a measure of the amount by which the center frequency f_d of the tracker oscillator is misaligned with respect to the center of the Doppler spectrum $f_{d_1} + f_{d_2}$. Thus, if the center frequencies are not coincident, the energies corresponding with filters *A* and *B* will be unequal, and the mean rectifier will contain a component at the modulating frequency. The phase-sensitive rectifier converts this component to a dc error signal used to control the mean frequency of the tracker oscillator and so servo it until a null is reached.

In the particular model of the tracker to receive detailed description in this paper, the Doppler frequency is required in the form of a shaft position, and so the oscillator frequency is controlled by a motor driven resistive potentiometer. Since a 400-cps motor is used for this purpose, the dc error signal must be converted to a 400-cps error signal. A further modulator makes this conversion in the right-hand box of Fig. 16.

3. Circuit Details

Fig. 17 shows the complete circuit of the tracker. The mixer *E* is a ring modulator of the same type as those used in the Janus multiplier. This is followed by the low-pass filter $C_1 D_5 C_2 D_6$, which is tunable. The bandwidth is required to be proportional to the tracker oscillator frequency and is controlled from the same potentiometer. The filter employs silicon diodes as variable resistance elements in an RC circuit. The signal level is

reduced to 0.1 of the maximum bias current for the diodes by suitable transformer ratios in the mixer. Since the signal current must be attenuated for the filter, a low-frequency amplifier $J_2 J_3 J_4 J_8$ is employed to return the signal to a convenient level for detection (0.25 ma). The amplifier consists of three direct-coupled stages $J_2 J_3 J_4$ whose dc conditions are controlled by a Miller feedback transistor J_8 , which sets the mean output potential to the potential at its emitter. The emitter potential of J_8 is determined by avalanche diode D_{24} . If the collector potential of J_4 becomes significantly more negative than this, current in excess of $I_b - I_{c0}$ flows in R_{13} , and the potential at the collector of J_8 moves in a positive direction. This reduces the base current of the input transistor J_2 until the potential at the collector of J_8 is correct. This dc feedback loop is capable of holding the output voltage stable over wide variations of input current while employing only one electrolytic capacitor. However, it should be noted that there are limitations on the gain of an amplifier stabilized by this circuit since the feedback via this loop becomes slightly positive at high frequencies. The ac feedback network $R_{11} R_7$ holds the gain of the amplifier to a safe level, so that the dc loop does not cause oscillation. It also insures that the output will be at the correct level for detection and reduces the input impedance presented to the filter.

The output signal of the amplifier is rectified at the emitter of J_5 , and detected in the low-frequency phase-sensitive detector consisting of the three diode rings

F, *G*, and *H*. The purpose of this phase-sensitive detector is to detect components in the collector signal of J_6 at the modulation frequency (25 cps) of the tracker oscillator. Three ring modulators are used for the low-frequency detector in order to circumvent the need for constructing current transformers which operate satisfactorily at the low modulation frequency. In this arrangement neither the reference nor the signal frequencies pass through transformers. Both the reference and the signal are modulated on to a higher frequency carrier in diode rings *H* and *F*, respectively. The reference and the signal pass through T_6 and T_5 as sidebands of the carrier frequency. The diode ring *G* then performs the demodulation, and the filter $C_{13}C_{14}L_1$ removes all components except the dc error signal. The latter is passed to the 400-cps modulator *K*, which converts the error signal to the appropriate form to actuate the Doppler frequency oscillator servo.¹⁴ The 5-kc carrier frequency is generated by a multivibrator J_6J_7 , similar in design to the multivibrator in the Janus multiplier. The description of tracker oscillator is deferred to Part IV, which deals with navigation. The oscillator is a natural link between the radar and the navigational devices, which themselves may assume a variety of forms. The oscillator may even become an integral part of a navigational computer and so should be described at the same time.

F. DETERMINATION OF DRIFT ANGLE

It will be recalled from Part II that the antenna system produces two pairs of fore/aft beams [Fig. 2(b)] and that the axis of symmetry of the fore/aft beam system is aligned with the ground track in order to determine the track angle. The beam is switched by waveguide switches at a frequency of about 1 cps. As a result, the 400-cps error signal from the tracker (Section E-3) also bears information concerning any misalignment between antenna axis and track as a component modulated by the 1-cps signal. This track error signal may be regained by applying the 400-cps tracker error signal to a phase-sensitive detecting system whose reference is the square wave antenna switch control. The techniques follow those which have been described very closely, and details are therefore omitted.

G. POWER SUPPLIES AND REGULATORS¹⁷

All the transistor circuits are designed to operate from positive and negative 20-v rails. Potentials, other than where ± 20 volts or ground potential are required, are obtained at the appropriate place in the circuit by means of resistor networks or avalanche diodes, thus eliminating the need for complex inter-unit wiring. Although the circuits are designed to be noncritical of supply voltage, it is convenient to use regulators as smoothing circuits. Transistor regulators are lighter than choke-capacitor combinations; they protect the remainder of the circuit from the extreme surges present in the aircraft supply; and may also be arranged to provide a low-impedance supply which, in turn, reduces interaction between separate circuits. The regulators are designed to meet all these requirements and, in addition, to be safe under all likely conditions of surge and overload, including a short circuit.

The requirements of the klystron power supply are much more stringent, both for stability and ripple. Furthermore, the problem is made extremely difficult by the high voltages (500 to 1000) required by klystrons. The solution adopted is to convert a highly stabilized 20 volts to the required level with a dc-to-dc converter consisting of a transformer coupled square wave oscillator. Such an oscillator can be designed to have a very low internal impedance, so that it is sufficient to regulate the low voltage input only. When operating at approximately 8.5 kc, switching occurs in 2 μ sec, so that smoothing of the output to the desired degree is a relatively simple matter.

Further advantages of the square wave oscillator are that 1) it can be protected against overload by designing for a cessation of oscillation on excessive circuit demand and 2) a single oscillator can provide all the voltages required by the klystron. The supply for the 600-mw klystron provides $500 \text{ v} \pm 2 \text{ v}$ at 80 ma for the beam supply with less than 2 mv of ripple and equally stable supplies for the reflector and heater.

¹⁷ C. A. Franklin, P. M. Thompson, and W. M. Caton, "Precision high-voltage transistor-operated power regulators, with overload protection," submitted to IEE International Conv. on Transistor and Associated Semiconductor Devices, London, Eng.; May 25-29, 1959.

Part IV—Data Conversion Between the Input Organs and Positional Computer

N. F. MOODY, P. M. THOMPSON, J. MITCHELL, AND J. H. GANTON

The preceding parts of this paper are concerned mainly with the input organs of an aircraft navigational system, and in order that the significance of these various data sources could be appreciated, a broad outline of the navigational processes was given. The following

parts show the manner in which these data are used in navigation and give some treatment of the computers involved. The instruments to be described make direct use only of the following data sources: 1) a heading reference, 2) a radar which gives ground speed (or incre-

mental ground distance), and drift angle, and 3) an air-speed indicator. Auxiliary checking means, e.g., map fixes, will not be treated.

The navigational computers involved are velocity triangle, position, and steering computers. Electro-mechanical analog instruments have sufficient accuracy to satisfy the first and third functions, and their simplicity makes analog computation the preferred technique. For positional computation this is not true, but analog instruments are, nevertheless, often used. This situation arises because these instruments exist as the result of earlier systems and, indeed, are capable of giving good service where the limiting accuracy of the radar is not desired.

In order to cater to the use of existing analog positional computers, a special converter unit has been designed whose function is to make the radar output compatible with the positional computer. The converter, a description of which follows, also provides a suitable output for the velocity computer. Part V describes a digital form of positional computer which does not require the converter unit and which itself provides data in suitable form for an analog velocity computer. Whichever type of positional computer is employed, the analog velocity computer may be required, so for completeness a description of its operation is also included in this part. The functions of a positional computer are deferred to Part V, and no description of analog instruments is given since these are well known.¹⁸ The treatment of steering computers, which are auxiliary to any system, is deferred to Part VI.

A. DIGITAL-TO-ANALOG CONVERTER

The track distance output of the radar is essentially in a form suitable for a digital type of positional computer. However, in order to allow the use of the radar with existing electromechanical analog computers, a converter has been designed which converts the radar output into a form suitable for these systems.

Electromechanical analog positional computers require a track angle input and either a distance or a ground speed input. The radar supplies one of its outputs in the correct mechanical form, *i.e.* drift angle, which is simply added to the heading angle to give track angle in the form required. This shaft angle is then transmitted to the computer by a conventional synchro system. The radar Doppler output is fed to a converter, one of whose output shaft positions indicates ground speed, while a second shaft position indicates the track distance flown.

The functions of the converter are illustrated in Fig. 18. As a matter of convenience the tracker oscillator is part of this converter, for when positional computers of the type described in Part V are used, the oscillator sig-

¹⁸ S. I. Frangoulis, "Design features of the ASN-7 navigational computer," IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS, vol. ANE-3, pp. 108-112; September, 1956.

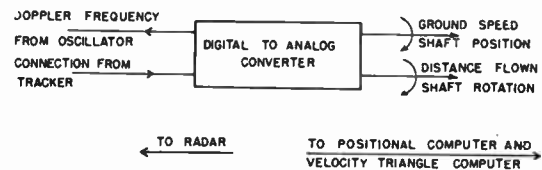


Fig. 18—Functions of the digital-to-analog converter.

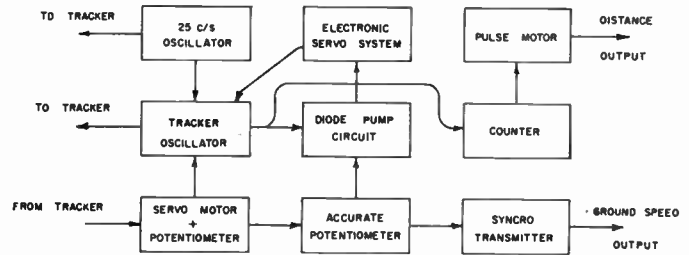


Fig. 19—Block diagram of the digital-to-analog converter.

nal is provided by one of its components. The converter delivers the tracker oscillator frequency to the tracker and receives an error signal from the tracker which corrects this frequency to the center frequency of the Doppler spectrum from the Janus multiplier in the radar.

Since either the speed or the distance output may be required for a positional computer, both must be as accurate as the over-all system. No problem exists in the conversion of distance where the accuracy is, in fact, digital since each Doppler cycle can be counted and so transmitted as an output shaft position. Speed (oscillator frequency) cannot be translated without instrumental error, but the specified accuracy, ± 0.1 per cent, is adequate for most purposes.

Fig. 19 shows a block diagram of the converter which is, in fact, digital-to-analog. This section describes first the operation of the tracker oscillator and its related circuits, followed by a brief outline of the speed measuring and distance registering system.

The tracker oscillator is connected with an electromechanical tuning system and a 25-cps frequency modulator. The tuning system consists of a transistor servo-amplifier¹⁹ and a motor which drives a potentiometer, the setting of which determines the frequency. The oscillator is tuned by a current whose value is proportional to the potentiometer setting. Two types of oscillator have been used. In one circuit the current is used to control the forward resistance of silicon diodes used as variable resistance elements in a Wien bridge; while in the other the current is used to charge a capacitance, discharged by a *p-n-p-n* trigger²⁰ when the charge has reached a predetermined value. Although the frequency of both of these oscillators is proportional to the current,

¹⁹ P. M. Thompson and J. Mitchell, "Some solutions to problems of operating germanium transistor servo amplifiers at high ambient temperatures," IRE TRANS. ON CIRCUIT THEORY, vol. CT-4, pp. 190-193; September, 1957.

²⁰ N. F. Moody and C. D. Florida, "Some new transistor bistable elements for heavy duty operation," IRE TRANS. ON CIRCUIT THEORY, vol. CT-4, pp. 241-261; September, 1957.

in neither is the relationship sufficiently linear, or stable, to allow use of the potentiometer shaft position as a measure of ground speed. The system described above simply sets the oscillator to the approximate Doppler frequency.

The oscillator is exactly related to the shaft position as follows. The oscillator frequency is compared with the setting of a second potentiometer in a "diode pump" frequency measuring circuit. The second potentiometer is also attached to the output shaft and is very linear. An error in the frequency results in an output current from the diode pump circuit; and the current, when amplified, is used to correct the frequency of the oscillator. The principle by which the frequency is corrected is illustrated in Fig. 20. The diode pump, shown in Fig. 20(a), is a semiconductor analog of the classical circuit shown in Fig. 20(b). If the frequency is adjusted until the potential at A is zero, then

$$\bar{i}_2 = \bar{i}_1 = VC'f.$$

If V and C' are known accurately, then f , the frequency of the square wave, is thereby determined to a similar accuracy. The square wave input is generated from the tracker oscillator, and precautions are taken to stabilize its amplitude. Thus the shaft position is accurately related to the Doppler oscillator frequency and becomes an accurate measure of the ground speed. An accuracy of better than ± 0.1 per cent has been achieved over the temperature range from -55°C to $+70^\circ\text{C}$. The simplified circuit of Fig. 20(a) omits many refinements which are essential to obtain this degree of precision.²¹

The operation of the tracker requires the tracker oscillator to be frequency-modulated with an accurate square wave. The frequency of modulation chosen for the converter is approximately 25 cps. Since an error in the duty cycle of the square wave would cause an error in the speed and distance measurement, special precautions are taken to control this duty cycle. The 25-cps square wave is derived by scaling the output of a 50-cps source. One complete cycle of the 50-cps determines one half a cycle of the 25-cps square wave; since successive cycles of the 50-cps input are similar, the mark-to-space ratio is determined accurately as 1:1. The output of the counter modulates the tuning current of the tracker oscillator and thus frequency modulates it at 25 cps.

The distance or ground miles output can be obtained simply by counting the number of cycles of the oscillator [see (2a) of Part II]. Thus the oscillator is coupled to a 12-stage binary counter followed by a pulse motor and a gear train (Fig. 19). The shaft rotation may be repeated at the positional computer by a synchro system or an M motor. A conventional mechanical counter connected to this output indicates ground miles flown and is used for checking and calibrating the system.

²¹ J. Mitchell, "An analog frequency measuring circuit accurate to 0.1 per cent," *AIEE Trans. (Commun. and Electronics)*, vol. 40, pp. 983-985; January, 1959.

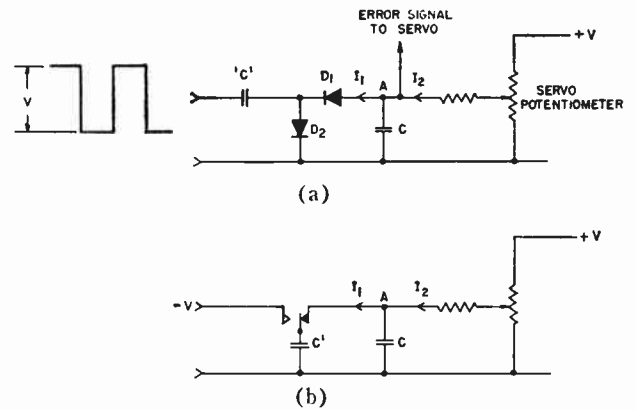


Fig. 20—Frequency measuring circuit: (a) diode pump circuit, (b) analog of above circuit.

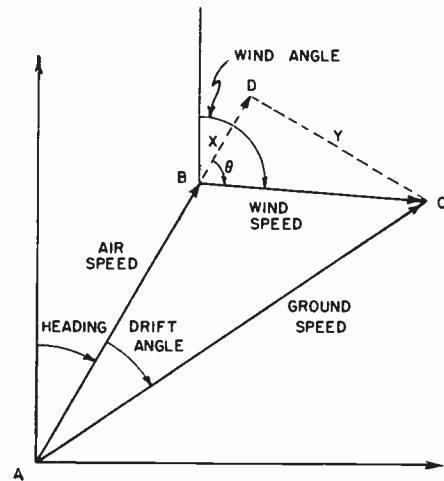


Fig. 21—The triangle solved by the velocity vector triangle computer

The digital-to-analog converter, like the remainder of the system, employs all semiconductor electronics and conventional 400-cps instrument servosystems. All parts have been fully tested over a temperature range from -55°C to $+65^\circ\text{C}$ and function well within the desired accuracy.

B. VELOCITY VECTOR TRIANGLE COMPUTERS

During any short period of time when the radar gives no output, the ground speed and track angle are computed from remembered wind information, airspeed and heading. This function is performed by a velocity vector triangle computer, which at all other times computes the wind vector from airspeed, heading, and the radar information. The velocity computer is mechanically coupled to the shafts bearing the outputs of the radar to the positional computer and is able to drive these when the radar is inoperative. Thereby both the radar frequency gates and antenna axis are correctly set to receive signals when the radar once more becomes operative. It will be seen that when the radar is inoperative, the velocity computer may be regarded as replacing the tracker. A sensing circuit is employed to detect whether there is sufficient Doppler signal to operate the tracker accurately, and when this signal falls below the level at

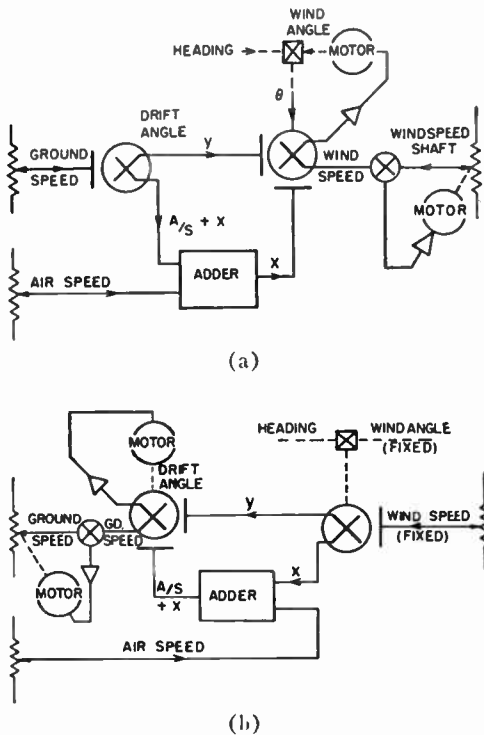


Fig. 22—Block diagram of the velocity vector triangle computer (a) while solving for wind, (b) while solving for ground speed and drift angle.

which the radar is more accurate than the velocity computer. At that time the computer takes over.

There are several suitable computers in existence, one being used in the AN/APN-81 navigation system.⁷ This computer uses two synchro resolvers of the type discussed in Section A of Part VI to solve the vector triangle shown in Fig. 21. Fig. 22(a) shows a block diagram of the computer arrangement used to calculate the wind. Potentiometers are set to provide voltages proportional to the airspeed and ground speed. The ground speed is resolved by the drift angle into components perpendicular to the airspeed y and parallel to the airspeed $x + A/S$. When x and y are known, a single resolver solves the

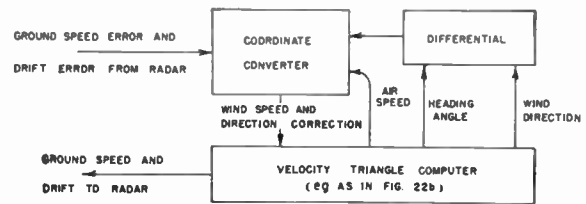


Fig. 23—Arrangement of velocity vector triangle computer for connection within the tracker feedback loop. (It should be noted that the conversion of error signal which uses airspeed, as shown, is imperfect, but adequate for most purposes. A better system, requiring an extra servomotor when used with the type of positional computer described in Part V, uses ground speed instead of airspeed.)

right-angled triangle BDC for the wind speed and for θ . The wind direction is found by adding the heading to θ with a differential gear, and a potentiometer is servoed to represent wind speed as a shaft position.

When there is no radar signal, relays are actuated to lock the shafts representing wind speed and wind direction and to change the configuration to that shown in Fig. 22(b). The solution for ground speed and drift angle uses the same two resolvers and the same method of calculation. The wind speed is resolved by θ into components x and y , and a resolver is used to solve the right-angled triangle ADC for the ground speed and the drift angle.

A second type of velocity triangle computer is shown in Fig. 23. This computer is connected within the tracker feedback loop, and the drift and ground speed correction signals are resolved in such a way that they will correct the wind direction and speed. The wind direction and speed are used to recompute ground speed and drift which are fed back again to the radar. The radar then holds the output of the velocity triangle computer to the radar accuracy, independent of the accuracy of the analog components in the triangle. When the radar signal fails, the wind direction and speed remain at their last setting, and the computer continues to provide ground speed and drift angle.

Part V—A Positional Computer of Digital Accuracy

N. F. MOODY, P. M. THOMPSON, AND J. H. GANTON

The accuracy with which an aircraft can be flown to a destination is determined by the accuracy with which its position is known, since it is assumed that the exact position of the destination is readily available. It is important, then, that a positional computer should be capable of using the full accuracy of the input organs. A system of input organs, already demonstrated capable of measuring distance to ± 0.2 per cent and track angle to $\pm 1^\circ$ over land²² has been described in Parts II-IV.

²² The error in track angle includes both drift angle errors due to the radar and heading errors.

It is probable that this accuracy will be improved even before this paper is published, and a distance accuracy of 0.1 per cent may be possible. An exact knowledge of distance from base would be of great use to a navigator even if he did not know the exact direction in which he had flown, since a radio bearing would allow him to correct his position; or, failing that, the knowledge that he has flown the exact distance of his destination would enable him to fly on a line which will pass directly over it.

The preceding part, in describing the functions of the digital-to-analog converter, introduced a concept of considerable importance. This concept is that the Doppler

radar output can be considered as yielding incremental digital information concerning the distance flown, and it was shown that the distance along the aircraft's track could be provided as a shaft position without loss of digital accuracy. Usually a positional computer will resolve this incremental distance into coordinates suitable for the frame of reference which it employs. This resolution may be performed by the electromagnetic type of resolver, described in Section A of Part VI, which would introduce the uncertain errors common to analog devices. These errors would be avoided if the resolution were performed digitally on the Doppler output signal directly.¹ In this case the distance error would be only that of the digital resolution, and since the output of the radar may be regarded as digital, little further error need be introduced by the computer.

This part of the paper describes positional computers which employ a digital technique having much in common with analog techniques. The first section consists of some brief remarks on the mathematical operations to be performed within the instrument. This is followed by a discussion of the digital analog computing techniques employed, and a description of their incorporation in some practical positional computers.

A. THE MATHEMATICS OF POSITIONAL COMPUTATION (DIGITAL DISTANCE)

The purpose of the positional computer is to provide continuously information about the aircraft's position with respect to a known reference grid. The two types of reference grid considered are: 1) a square grid used over small areas of the earth which may be considered plane, and 2) latitude and longitude. This part of the paper describes computer systems for use with each type of reference.

The inputs to both positional computer systems consist of a train of pulses representing the distance the aircraft travels and a synchro repeated shaft position representing the track direction θ . For sufficiently small increments of distance the track angle may be assumed constant, so that the computer will operate on incremental information in polar coordinate form. This motion can be resolved into rectangular coordinates as follows:

$$\Delta_{(\text{distance north})} = \Delta r \cdot \cos \theta,$$

$$\Delta_{(\text{distance east})} = \Delta r \cdot \sin \theta.$$

Position on a square grid is given as distance north and distance east of a datum, and therefore

$$\text{distance north} = \sum \Delta r_j \cdot \cos \theta_j + \text{initial distance north},$$

$$\text{distance east} = \sum \Delta r_j \cdot \sin \theta_j + \text{initial distance east}.$$

These equations completely describe the mathematical operations necessary for a square grid positional computer. Fig. 24 shows the logical block diagram of such a computer.

Position on a latitude and longitude reference system may be computed by converting the increments of

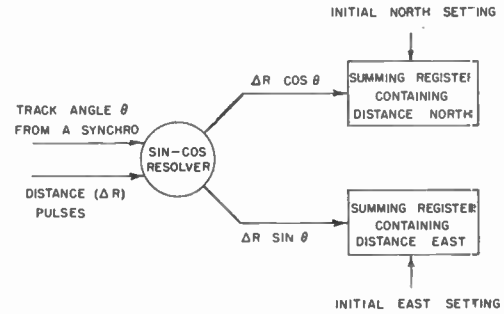


Fig. 24—Logical block diagram of the square grid positional computer.

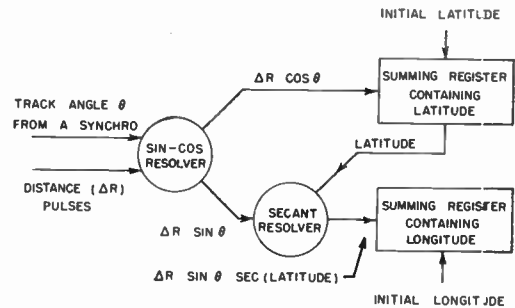


Fig. 25—Logical block diagram of the latitude and longitude positional computer.

travel north and east into increments of latitude and longitude as follows:

$$\text{latitude} = \sum \Delta r_j \cdot \cos \theta_j + \text{initial latitude},$$

and

longitude

$$= \sum \Delta r_j \cdot \sin \theta_j \cdot \secant(\text{latitude})_j + \text{initial longitude}.$$

These equations completely describe the mathematical operations necessary for a latitude and longitude analog positional computer. Fig. 25 shows the logical block diagram of such a computer.

In both computers the mathematical operations are performed by means of the digital analog techniques to be described in the next section. Such techniques offer an ideal solution since the radar system delivers the natural form of input required and since the resulting computers prove to be simple, economical and highly accurate.

B. DIGITAL ANALOG COMPUTATION

Digital analog computation is a form of analog computation where the values are transmitted as trains of pulses. It follows that the logical block diagram of a digital analog computer is identical to that of any other form of analog computer, e.g., voltage or mechanical. However the data are always handled in the form of pulses, such that one pulse represents a single increment of magnitude. This single weight for a pulse distinguishes the digital analog computer from the conventional digital computer, with which, however, it shares

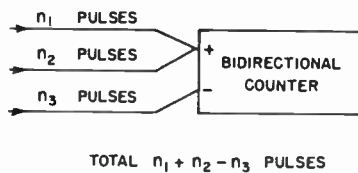


Fig. 26—Digital analog addition and subtraction

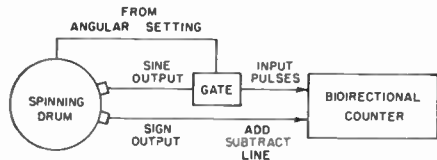


Fig. 27—Logical block diagram of a spinning drum device used as a sine function generator.

the property of delivering results where the accuracy is limited only by the resolution. As a member of the analog computer family it is most suited to a fixed program, and is often far simpler than a conventional digital computer.

This part of the paper illustrates how mathematical operations may be performed within a digital analog system, in preparation for a description of their practical embodiment in the positional computers. This paper is restricted to the three operations necessary for the positional computers: addition, subtraction, and resolution by trigonometrical functions (to be described in this order).

1. Addition and Subtraction

Addition and subtraction are usually performed within a binary counter, as illustrated in Fig. 26. The counter must be bidirectional if both addition and subtraction are required, and also there must be some provision to prevent coincidence between any two input pulses. The usual practical arrangement is that there is a single input for digits and that the function of the counter *i.e.*, addition or subtraction, is determined at any time by another input, usually referred to as the add-subtract line. A counter of this type is shown in the right-hand block of Fig. 27.

2. The Generation of Trigonometrical Functions

The digital analog resolving system to be described receives two inputs: one in the form of a pulse train is the quantity to be resolved, and the other in the form of a shaft position sets in the angle θ . The resolver also delivers its output as a pulse train, but this train must be summed in a counter to obtain the required function. The action of any trigonometrical resolver can be considered as based on two properties. First the resolver must be able to generate the function in the form of a ratio between two numbers of pulses, *e.g.*, $\sin \theta$, and then must be able to control the magnitude of the vectors according to the input magnitude $.1$, so as to generate the required output, $.1 \sin \theta$. The following description will commence by showing how the ratio $\sin \theta$ is generated.

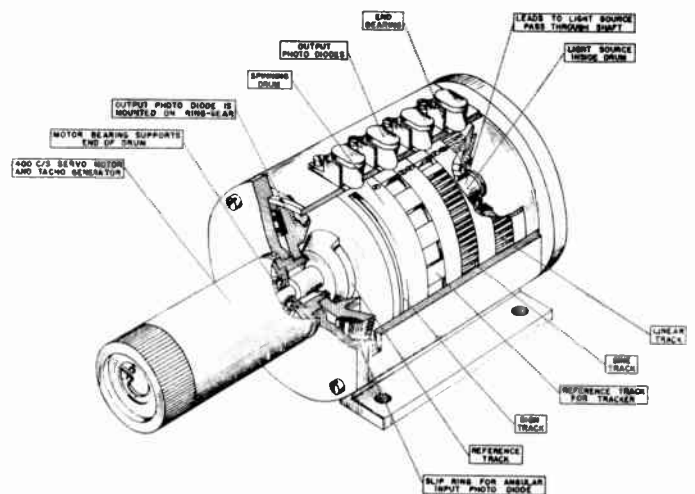
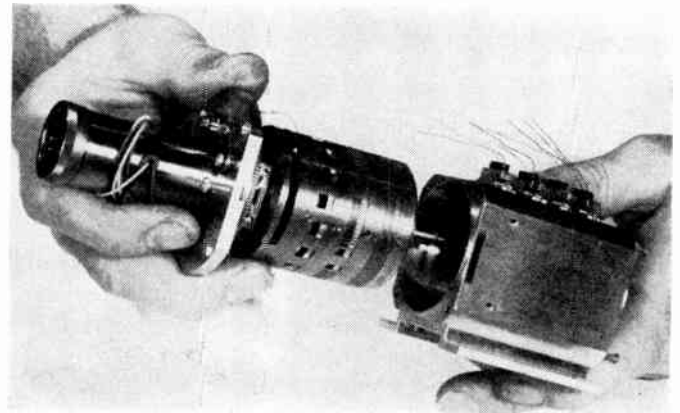
Fig. 28—An early example of a spinning drum digital resolver (resolution 1°).

Fig. 29—A photograph of an early spinning drum resolver. The device is taken apart in order to show the drum.

The resolver consists of a motor driven drum (Figs. 28 and 29), on which are several tracks of perforations (or marks). In the center of the drum is a light source and on the outside, opposite the tracks, are photodiodes mounted in the outer casing. As the drum rotates the marks interrupt the light, thus causing the photodiodes to generate the pulses used in the system.

Three of the five tracks are required to generate the ratio $\sin \theta$: the function track, the reference track and the sign track (to be introduced in this order). The function track delivers output pulses whose sum will be the required sine. Its perforations are so graduated that their density per unit angle ϕ (measured around the drum) varies according to the law $|k \cos \phi|$, where k is an arbitrary constant. The zero position of ϕ is defined by the reference track. It is necessary that the function $k \cos \phi$ shall bear the sign for the value of ϕ in each of the four quadrants. Since the function track cannot deliver sign it is supplemented by the sign track, which contains a single mark extending from 90° – 270° . Throughout these quadrants the photodiode output of this track is treated as a negative quantity in accordance with the properties of a cosine.

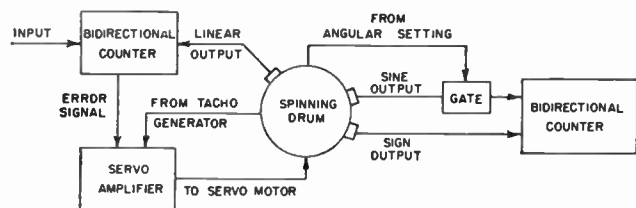


Fig. 30—Logical block diagram of a spinning drum device used as a digital sine resolver.

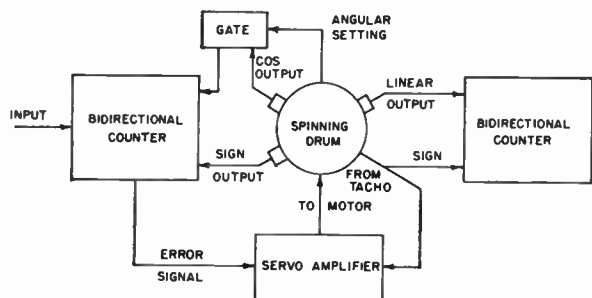


Fig. 31—Logical block diagram of a spinning drum device used as a digital secant resolver.

It is now possible to show that the sum of the output pulses (with sign determined by the sign track) is proportional to the sine of the angle ϕ through which the drum is rotated. This operation is performed by connecting the output of the function track to the input of a counter whose add-subtract line is actuated by the output of the sign track. This is illustrated in Fig. 27, where for the moment the gate should be ignored. As the drum is rotated from a position $\phi = 0$, the scaler total at any angle ϕ' is

$$\sum_0^{\phi'} k \cos \phi_j \cdot \Delta\phi_j \sim \int_0^{\phi'} k \cos \phi \cdot d\phi = k \sin \phi'$$

The sum may be made to approach the true integral as the function track is more finely graduated, and thus $k \sin \phi'$ may be generated to the required digital accuracy.

In the practical system the drum rotates continuously, but a similar effect can be had by gating the output pulses to the counter so that they are accepted only over the interval 0 to ϕ' . Then each rotation of the drum will generate a score in the counter of $k \sin \phi'$. Such gating can be performed by introducing a second photodiode on the reference track, whose angular position can be offset by θ° from the zero index position. If θ is set equal to ϕ' and the gate is opened and closed by impulses from these photodiodes, the scaler will score $k \sin \theta$ pulses per revolution. The block diagram of Fig. 27 represents such a system. The angle θ is set in by means of a second shaft (not shown), which causes the appropriate photodiode to be displaced by turning the ring gear on which it is mounted.

The final step, controlling the magnitude A of the output vector, $k \sin \theta$, is arranged as follows. The drum is caused to rotate once for every set n of input pulses. Then, since the counter receives a total of $k \sin \theta$

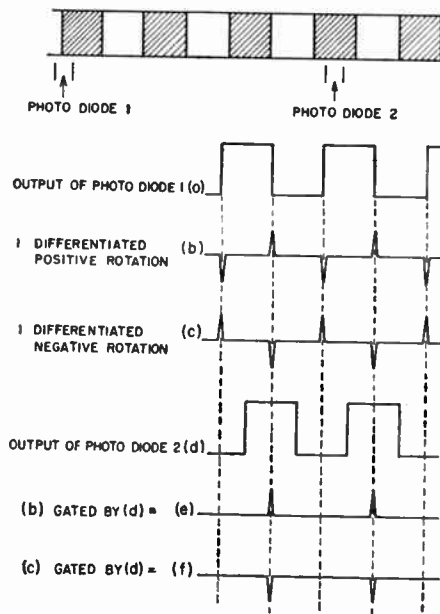


Fig. 32—Method by which the direction of rotation of the drum in the secant resolver is detected.

pulses per revolution, the output/input relationship will be

$$\frac{\text{counter store}}{\text{no. of input pulses}} \sim \frac{k}{n} \sin \theta = A \sin \theta,$$

which approaches equality when the number of input pulses is very large compared to a single set n . The required control of drum rotation is provided by a fourth track linearly divided by n marks. If, as shown in Fig. 30, the input pulses are added in a bidirectional counter from whose score the output of the linear track is subtracted, the instantaneous total in the scaler will represent an error signal which may be used to correct the drum position. This error signal may therefore be made to actuate the drum motor to form a servosystem, which may be damped by employing the usual tachometer generator.

The general navigation equations (Section A) were seen to require both cosine and secant functions. It can be shown that cosine functions are obtainable from another pair of photodiodes mounted on the function and sign tracks, respectively, at right angles to the existing pair. These new diodes feed a separate counter which will generate $A \cos \theta$ when gated identically with the $\sin \theta$ counter. Secant can be resolved as follows. It will be noticed from Fig. 30 that the resolver generates pulses at both linear and sine (or cosine) outputs. Since secant is simply the reciprocal of cosine it is necessary only to reverse the roles of the cosine and linear outputs, as shown in Fig. 31, so that the cosine pulses are now compared with the incoming train and the linear pulses are summed to produce the solution. The direction of rotation of the drum determines the sign of the secant and, in this block diagram, the output from the tachometer generator indicates the direction. However, this

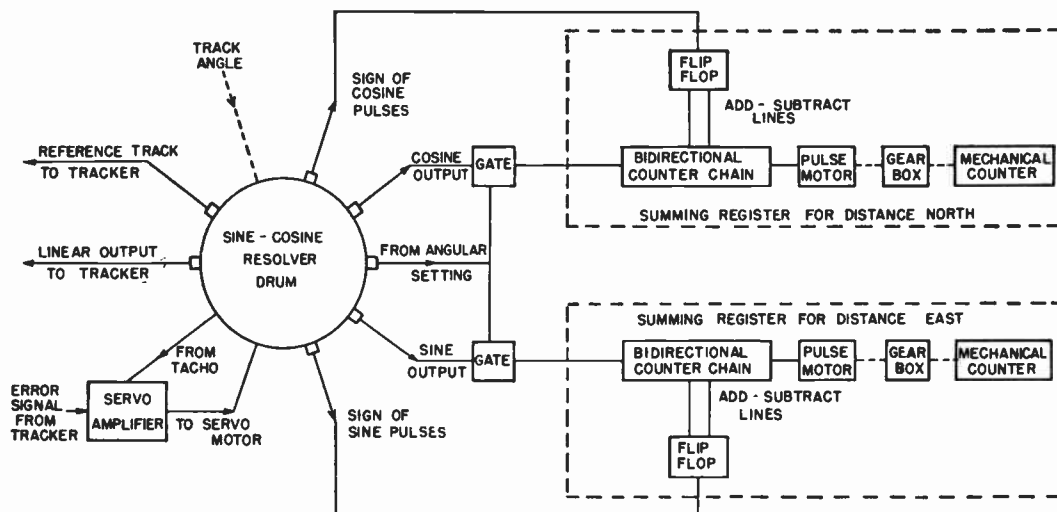


Fig. 33—Practical square grid positional computer.

method has limitations, for at very low speeds the desired output from the tachometer becomes masked by spurious signals. In the computer a second photodiode is positioned on the linear track in such a way that its output is 90° out of phase with the other photodiode on the linear track, as in Fig. 32. The differentiated output of the first photodiode is gated by the output of the second photodiode, so that positive pulses (e) represent positive rotation, while negative pulses (f) represent negative rotation.

The way in which the resolvers are used in the practical embodiment of the positional computer is discussed further in the next part of this paper.

C. PRACTICAL POSITIONAL COMPUTERS

This part shows the practical embodiment of the digital analog components in the square grid and the spherical grid positional computers shown logically in Figs. 24 and 25.

Fig. 33 shows the practical block diagram of the square grid computer. A single spinning drum is used to resolve the distance input pulses by $\sin \theta$ and $\cos \theta$, where θ is the track angle set from a synchro. The spinning drum also generates two of the waveforms required by the tracker. The output of the linear track is used directly as the tracker oscillator which, it will be remembered from Section A of Part IV, required a superimposed square wave FM. In fact the linear track bears step changes in the density of the marks so that the two generated frequencies lie above and below the mean track frequency. The pattern of density changes is in accordance with a second track (reference track for tracker, Fig. 28) which itself supplies the reference frequency required by the tracker for the demodulation. The output of the tracker becomes the input of the servoamplifier and regulates the speed of revolution of the drum.

At this point it should be noted that a tachometer generator mounted on the servomotor provides a 400-

cps voltage analog of the speed of the drum and, hence, the ground speed of the aircraft. This voltage may be used directly by the velocity triangle computers of Section B of Part IV in place of the ground speed input shaft, since the purpose of the shaft is to set a potentiometer to such a voltage.

The summing registers are composed of a bidirectional counter chain, a pulse motor, and a mechanical counter. The add-subtract lines of the bidirectional counter chain are controlled by the sign output of the spinning drum. A train of pulses from the spinning drum is summed in the scaler chain and the most significant figures of the number are converted into shaft rotation by the pulse motor. The pulse motor, in turn, drives the mechanical counter where distances are stored and displayed. A gear box between the pulse motor and the mechanical counter introduces a scale factor, so that the counter reads directly in nautical miles. The mechanical counters are provided with facilities by which the original position may be entered or corrections may be made during flight. The sine-cosine resolver and the two summing registers comprise a complete square grid computer.

A comparison of the logical block diagrams (Figs. 24 and 25) shows that a latitude and longitude computer is similar to the square grid computer with the addition of a secant resolver. Similarly, the digital analog latitude and longitude position computer of Fig. 34 is based on Fig. 33, with the addition of a second spinning drum used as a secant resolver. The sine output pulses of the sine-cosine resolver and the cosine output pulses of the secant resolver are retimed to avoid coincidence and then compared in an error register (a bidirectional counter chain). The error signal which is proportional to the number in the error register, *i.e.* the difference between the two trains of pulses, controls the speed of revolution of the secant drum. The add-subtract lines of the error counter chain are controlled by gates connected to the two inputs, by the sign of sine output from the sine-

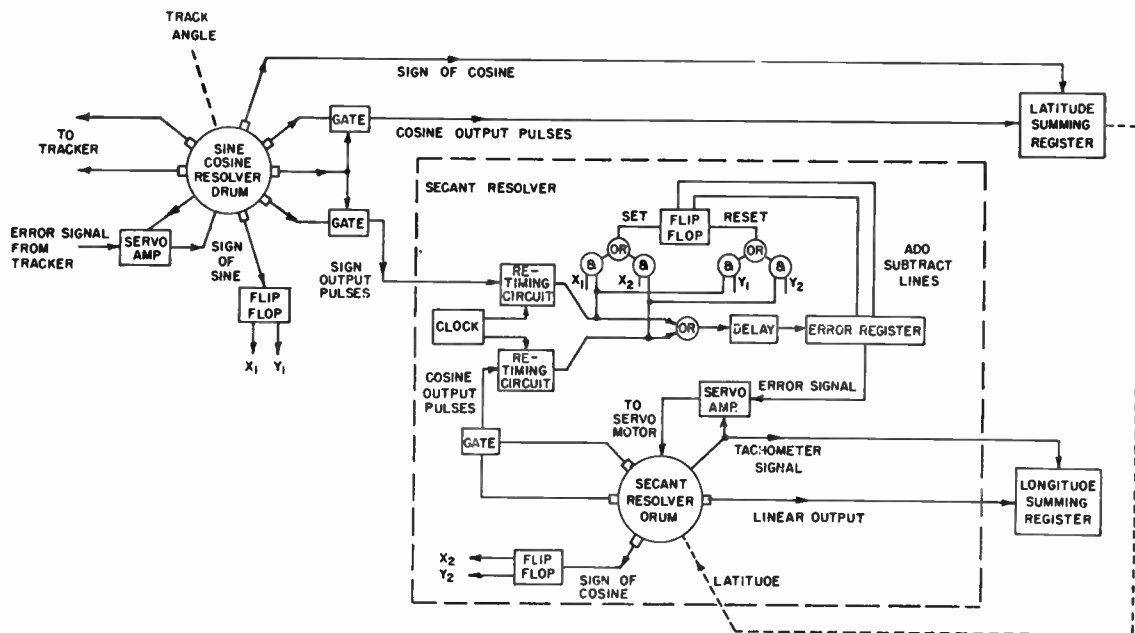


Fig. 34—Practical latitude and longitude positional computer.

cosine resolver and the sign of cosine output from the secant resolver. The input pulses to the error register are delayed to allow the add-subtract lines to switch. The linear output of the secant drum and the cosine output of the sine-cosine drum are summed in registers similar to those in the square grid computer with suitable mechanical counters substituted. The latitude summing register sets the angular input of the

secant resolver by means of a suitable gear train. The addition of these parts to the square grid computer converts it to a latitude and longitude computer.

These combinations of semiconductor electronics and mechanical components allow the construction of both latitude and longitude, and square grid positional computers with digital accuracy but without the complication of a general purpose digital computer.

Part VI—Steering Computers and the Choice of a Complete Navigational System

J. H. GANTON, P. M. THOMPSON, AND N. F. MOODY

The previous parts have been concerned with the equipment necessary to define the position of an aircraft. This part discusses the remaining navigational problem, that of computing a route to a destination. The route is usually presented to the pilot as a heading and a distance. The accuracy required need be no greater than the precision with which an aircraft can be flown and it is postulated that the pilot cannot make use of solutions more accurate than one degree for heading and one per cent of the distance to be flown. The resultant errors are not cumulative because the positional computer provides a constant check on true aircraft position. The accuracy of a conventional analog computer suffices, and the steering computers here described are based on this method since it leads to the simplest instrumentation.

The equations to be solved depend on the type of reference grid used. Steering computers are described for use with both the square grid type of positional computer and the latitude and longitude positional com-

puter. The output of the steering computer is track angle and distance to the destination; however, the pilot needs heading rather than track angle so that drift angle is added before display. The square grid steering computer solves a right-angle triangle, while for the latitude and longitude steering computer, more complex solutions are required. Since latitude and longitude are spherical coordinates, an accurate distance and bearing require the solution of a spherical triangle. Where the position is calculated on a square grid, one axis of which is referenced along the desired track, the computer outputs are distances along and across track. Then no separate computer is required; the steering information is provided by the positional computer directly.

The purpose here is to compare the relative merits of several methods of steering computation. Block diagrams of a square grid steering computer, a latitude and longitude spherical steering computer, and two plane earth approximations are developed from their equations. Since these computers are based on synchro re-

solvers,²³ the characteristics of these components dictate the form of the equations.

A. RESOLVER CHARACTERISTICS

A synchro resolver is essentially a transformer whose coefficient of coupling can be controlled by rotating a shaft. There are normally two stator windings at right angles to each other and two rotor windings also at right angles. If one stator winding is energized by a voltage E , then the voltages on the rotor windings are:

$$E_{r1} = E \sin \phi \tag{14a}$$

$$E_{r2} = E \cos \phi, \tag{14b}$$

where ϕ is the shaft angle of the rotor. When both stator windings are energized, one with voltage E_1 and the other with voltage E_2 , then the voltages on the rotor windings are given by

$$E_{r1} = E_2 \cos \phi - E_1 \sin \phi \tag{15a}$$

$$E_{r2} = E_1 \cos \phi + E_2 \sin \phi. \tag{15b}$$

In the following discussion of the steering computers, the equations are manipulated in order to make use of (14) and (15).

B. SQUARE GRID STEERING COMPUTER

When the position of the aircraft and the destination are known on a square grid, then the distance to destination and the required track angle are computed by solving a right-angled triangle (see Fig. 35). If the following notation is used:

x = distance east

y = distance north

Δx = (destination distance east) - (aircraft distance east)

Δy = (destination distance north) - (aircraft distance north)

R = distance to destination

θ = desired track angle,

then

$$\tan \theta = \frac{\Delta x}{\Delta y}$$

The relationships

$$\Delta x \cos \theta - \Delta y \sin \theta = 0 \tag{16a}$$

and

$$R = \Delta y \cos \theta + \Delta x \sin \theta \tag{16b}$$

follow immediately, and notice that they are of the same form as the set (15).

Fig. 36 shows the block diagram of a computer for solving these equations. Voltages proportional to Δx and Δy are set on potentiometers coupled to the inputs of a

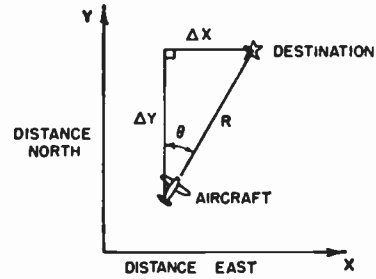


Fig. 35—Triangle solution for square grid steering computation.

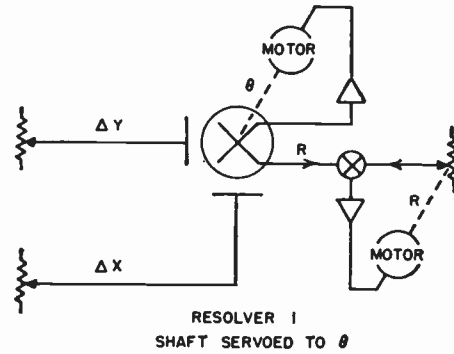


Fig. 36—Logical block diagram of a square grid steering computer.

resolver. Since the voltage on one rotor winding is proportional to $\Delta x \cos \phi - \Delta y \sin \phi$, the triangle is solved when this output equals zero. Then the rotor shaft angle ϕ equals the track angle θ (or $\theta + \pi$), and the magnitude of the voltage on the other rotor winding is proportional to the distance R of (16b).

The practical computer employs a servomotor to position the rotor, and with such a system there is no ambiguity in ϕ since only one null is stable. The distance can also be displayed as a mechanical position by means of the servomotor and potentiometer shown in Fig. 36. The accuracy of this computer depends upon the size of grid covered and upon the accuracy of the resolver and potentiometers.

C. LATITUDE AND LONGITUDE STEERING COMPUTER

An exact great circle course is obtained by solving the spherical triangle joining the north pole, the aircraft and the destination.²³ For convenience, the equations are manipulated using the standard spherical trigonometric notation shown in Fig. 37. Symbols representing the known and required quantities are then substituted. These symbols are shown and the two types of notation are related in Table I.

Consider the spherical triangle ABC where A is the north pole, B is the aircraft, and C is the destination. The known quantities are A , b , and c ; the desired quantities are B (the track angle) and a (the distance to destination). The spherical triangle equations used are the law of cosines,

$$\cos a = \cos b \cos c + \sin b \sin c \cos A, \tag{17}$$

and the identity

$$\cos c \cos A = \sin c \cot b - \sin A \cot B. \tag{18}$$

²³ E. B. Brown, "Resolver analog computers," IRE Second Natl. Conv. on Military Electronics, Washington, D. C.; June 17-18, 1958. See pp. 163-167 of the Conference Proceedings.

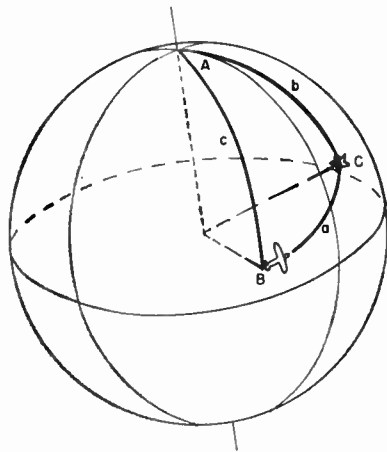


Fig. 37—Spherical trigonometry notation. In spherical triangle ABC , a, b, c represent the angular measure of sides BC, CA, AB , respectively. A, B, C represent the angles of the triangle.

TABLE I

La_A	latitude of aircraft	La_A	$90^\circ - c$
Lo_A	longitude of aircraft	La_D	$90^\circ - b$
La_D	latitude of destination	ΔLo	A
Lo_D	longitude of destination	R	a
ΔLa	$La_D - La_A$	θ	B
ΔLo	$Lo_A - Lo_D$		
R	distance to destination		
θ	track angle		

If the identity (18) is multiplied by $\sin b \sin B$ and rearranged it becomes

$$\cos B [\sin b \sin A] + \sin B [\cos c (\sin b \cos A) - \sin c (\cos b)] = 0, \quad (19a)$$

which may be rewritten as

$$\cos \theta [\cos La_D \sin \Delta Lo] + \sin \theta [\sin La_A (\cos La_D \cos \Delta Lo) - \cos La_A (\sin La_D)] = 0. \quad (19b)$$

This equation is in a form which may be readily solved for θ by means of four resolvers (Fig. 38). The shafts of resolvers 1, 2, and 3 are respectively set to the latitude of the destination, the difference in longitude between the aircraft and the destination, and the latitude of the aircraft. The fourth resolver is used to solve for the track angle θ . Then the unused rotor winding on resolver 3 gives

$$\begin{aligned} &(\cos b) \cos c + (\sin b \cos A) \sin c \\ &= \cos a \text{ [by the law of cosines (17)]} \\ &= \cos R, \end{aligned}$$

and the unused rotor winding on resolver 4 gives

$$\sin B [\sin b \sin A] - \cos B [\sin b \cos c \cos A - \cos b \sin c] = \sin a = \sin R.$$

The distance R is calculated by use of a fifth resolver (Fig. 38).

Since the spherical computer is scaled for long distances, it becomes inaccurate at short ranges, and a supplementary computer is required. A short-range computer uses a plane earth approximation to the

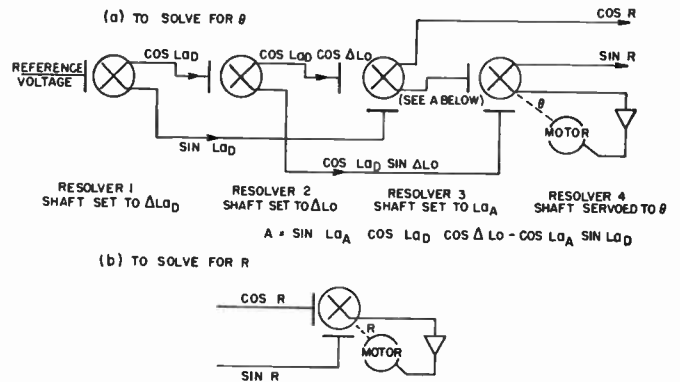


Fig. 38—Logical block diagram of latitude and longitude steering computer: (a) solution for θ , and (b) solution for R .

spherical latitude and longitude reference grid. A number of approximations can be used, two of which will be discussed. The range over which accurate results are required determines the approximation used and hence the complexity of the computer.

A computer based on a simple rectangular approximation has sufficient range to supplement the spherical computer. A suitable approximation is based upon the assumption that the meridians of longitude are parallel within the area under consideration. In this approximation

$$\begin{aligned} \Delta_{(\text{distance north})} &= \Delta La \\ \Delta_{(\text{distance east})} &= \Delta Lo \cos La_D. \end{aligned}$$

Fig. 39 shows a computer suitable for such a grid. It is a modified form of the square grid computer shown in Fig. 36. The only addition is a resolver which calculates distance east from ΔLo .

Increased range results from representing latitude and longitude by the trapezoid $ABCD$ ¹⁸ shown in Fig. 40. The trapezoid is a better approximation to latitude and longitude than the rectangle since it takes into account the convergence of the meridians between the latitude of the aircraft and the latitude of the destination. Triangle BDF is solved to find the range R and the track angle θ .

Since angle ABF is small, it may be assumed that

$$FB = AB = \Delta Lo \cos La_D \quad (20)$$

and

$$DF = \Delta La + d, \quad (21)$$

where

$$d = \Delta Lo \cdot \cos La_D \cdot \cos \phi,$$

and

$$\begin{aligned} \cos \phi &= \frac{\Delta Lo}{2\Delta La} [\cos La_D (\cos \Delta La - 1) \\ &\quad + \sin La_D (\sin \Delta La)]. \end{aligned}$$

The derivation of the expression for $\cos \phi$ is straightforward but tedious and is therefore omitted. If ΔLa is small, the expression for $\cos \phi$ may be simplified, for then it may be assumed that

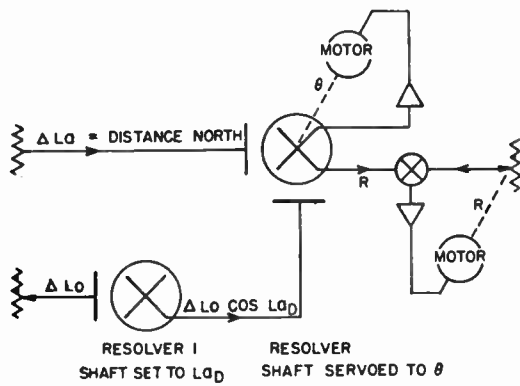


Fig. 39—Logical block diagram of steering computer based on a rectangular approximation.

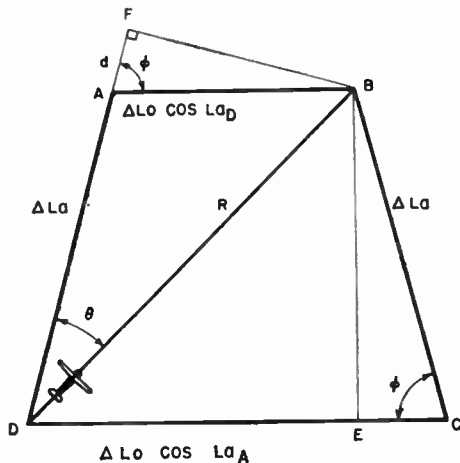


Fig. 40—A trapezoid approximation for a spherical triangle.

$$\sin \Delta La = \Delta La \quad \text{and} \quad \cos \Delta La = 1.$$

When these approximations are made, the expression for $\cos \phi$ reduces to

$$\cos \phi = \frac{\Delta Lo}{2} \sin L\alpha_D.$$

Eq. (21) now becomes

$$DF = \Delta La + (\Delta Lo \cdot \cos L\alpha_D) \left(\frac{\Delta Lo}{2} \sin L\alpha_D \right). \quad (21a)$$

With the sides FB and DF given by (20) and (21a), the right-angled triangle BDF may be solved for R and θ in the manner discussed previously.

Fig. 41 shows the block diagram of a suitable computer similar in form to the computer based on a rectangular approximation (Fig. 39) with the addition of two blocks: a multiplier to calculate d , and an analog adder to add this to ΔLa . The multiplier may be either a conventional electromechanical servomultiplier or an all electronic multiplier. The utility of these two approximations is discussed in the following section.

D. OPERATIONAL REQUIREMENTS—ACCURACY VS SIMPLICITY

The method of steering computation chosen for a navigational computer depends upon its operational re-

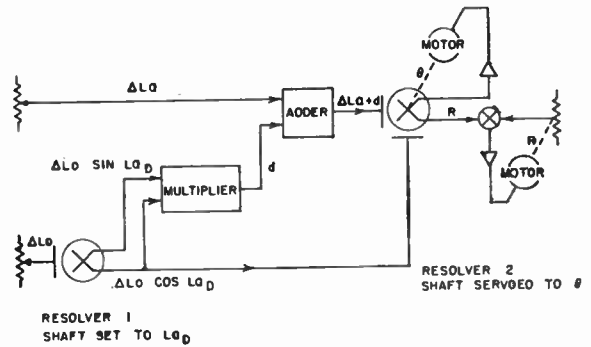


Fig. 41—A logical block diagram for a steering computer using the trapezoid approximation for a spherical triangle.

quirements since no one method investigated performs every possible requirement well. The requirements fall under two general headings, those of accuracy and those of simplicity. It is possible to construct a hybrid steering computer using several methods of computing course, so that the method used at any time is dependent upon the respective positions of the aircraft and its destination. Such a complex system would be heavier and less reliable than a simpler system. It follows that the design of a steering computer must be a compromise between accuracy and complexity. In this part, possible requirements for a steering computer will be related to the accuracy and complexity resulting from the systems described.

1. A Comparison of the Accuracies of the Various Steering Computers

It is instructive to show the relation between the spherical solution and the two approximations. The equations for $\tan \theta$ are

spherical:

$$\tan \theta = \frac{\sin \Delta Lo \cos L\alpha_D}{\cos L\alpha_A \sin L\alpha_D - \sin L\alpha_A \cos L\alpha_D \cos \Delta Lo}, \quad (22)$$

trapezoid approximation:

$$\tan \theta = \frac{\Delta Lo \cos L\alpha_D}{\Delta La + \frac{(\Delta Lo)^2}{2} \cos L\alpha_D \sin L\alpha_D}, \quad (23)$$

and rectangular approximation:

$$\tan \theta = \frac{\Delta Lo \cos L\alpha_D}{\Delta La}. \quad (24)$$

These approximations may be deduced directly from the spherical formulas as follows. When ΔLo and ΔLa are small:

$$\sin \Delta Lo = \Delta Lo, \quad \cos \Delta Lo = 1 - \frac{(\Delta Lo)^2}{2},$$

and

$$\sin \Delta La = \Delta La. \quad (25)$$

TABLE II
EFFECT OF SYSTEM ACCURACIES ON TYPICAL COURSES

Triangle			Great Circle Course			Rectangular Approximation		Trapezoid Approximation		REMARKS
La_A	La_D	ΔLo	Track	Angle	Distance	Track Angle Error	Distance Error	Track Angle Error	Distance Error	
75°N	80°N	10°	18°	35'	326	34'	8 nm or 2½ per cent	03'	<1 nm	Accuracy of both approximations decreases with increasing range due to assumption (25)
50°N	70°N	20°	18°	05'	1327	48'	59 nm or 4½ per cent	07'	5 nm or 0.4 per cent	
40°N	70°N	30°	17°	54'	2028	59'	189 nm or 9 per cent	20'	19 nm or 1 per cent	
05°S	05°N	20°	63°	42'	1340	21'	40 nm or 3 per cent	1° 03'	6 nm or 5 per cent	Trapezoid approximation becomes inaccurate when course crosses equator due to assumption (27)
44° 30'N	45°N	5°	80°	04'	215	1° 53'	1 nm	10'	<1 nm	Rectangular approximation is inaccurate for mid-latitude courses where $\Delta Lo > \Delta La$ due to assumption (28)
60°N	70°N	30°	38°	43'	952	6° 03'	102 nm or 11 per cent	37'	20 nm or 2 per cent	

The spherical equation (22) then becomes

$$\tan \theta = \frac{\Delta Lo \cos La_D}{\Delta La + \frac{(\Delta Lo)^2}{2} \sin La_A \cos La_D} \quad (26)$$

If it is assumed that

$$\sin La_A = \sin La_D, \quad (27)$$

then (26) becomes the equation of the trapezoid approximation. This last assumption, (27), does not hold when the aircraft and the destination are on opposite sides of the equator. However, near the equator $\sin La_D$ is small and the resulting errors are small. If it is assumed that

$$\frac{(\Delta Lo)^2}{2} \sin La_A \cos La_D = 0, \quad (28)$$

then (26) becomes the equation of the rectangular approximation. Since $\sin La_A \cos La_D$ has a maximum value of one half, $(\Delta Lo)^2/4$ must be small with respect to ΔLa for the assumption (28) to be valid. This approximation is least accurate for midlatitude courses which are approximately east-west, and thus the computer based on this rectangular approximation has a severe range limitation.

The effects of these assumptions on some typical courses are shown in Table II. This table compares some typical courses computed by the two plane earth approximations against a true great circle bearing and distance.

It will be noted that the simple rectangular approximation is fairly accurate near the equator, but is not suitable for flights over 200 miles at medium latitudes. However, at these latitudes the trapezoidal approximation is adequate for flights of up to 1000 miles between turning points. Both these approximations have severe limitations near the poles, where it is common practice to use another system of coordinates.

While the great circle computer has no system error since it solves true spherical triangles, the requirements for its components are the most stringent of any of the systems discussed. This will become apparent when it is realized that one revolution of a resolver corresponds to a complete circumference of the earth. Because these errors are due solely to the components, some typical results of an experimental model are tabulated in Table III and give guidance on realizable accuracy.

The computer (Fig. 38) was constructed from five size-15 amplifier compensated resolvers, for which suitable transistor amplifiers have been designed. The table demonstrates that such a computer gives a course within $\pm \frac{1}{2}^\circ$, or misses the destination by approximately 12 miles, whichever is the greater. The computer is therefore suitable for providing steering information at distances greater than 200 miles; but at distances closer than this a plane earth auxiliary computer, such as the one using a rectangular approximation (Fig. 39), is necessary. The errors due to components in such short-range computers are much less than for the spherical triangle computer and may be reduced to less than a mile by suitable care in the electrical design.

E. THE CHOICE OF A COMPLETE NAVIGATIONAL COMPUTING SYSTEM¹

The final form of any complete navigational computing system will be determined by the requirements of a particular user, who will restrict the facilities to his particular needs in the interests of equipment simplicity.

It is quite possible, however, to make a flexible assembly of interchangeable computing units so that any, or all, of the facilities of a general purpose system can be selected at will. Such a system has great advantages and may be developed with little further effort.

Several guiding principles, used to choose the facilities appropriate to a particular operation need, will now be discussed. There seems little doubt that navigation

TABLE III
 ERRORS IN EXPERIMENTAL SPHERICAL STEERING COMPUTER

Triangle			Correct* Track Angle	Measured Error	Correct* Distance	Measured Error
La_A	La_D	ΔLo				
72°S	69°N	-66°	322° 32'	31'	7362 nm	8 nm
84°N	69°N	113°	54° 20'	02'	1444 nm	9 nm
30°N	45°N	15°	33° 57'	19'	1148 nm	3 nm
30°N	36°N	6°	38° 23'	15'	469 nm	2 nm
30°N	32°N	2°	40° 05'	1' 21'	158 nm	2 nm
30°N	31°N	1°	40° 29'	2° 48'	79 nm	1 nm

* Assuming the earth is a perfect sphere.

based upon latitude and longitude is the best all-round system. A position defined thus is unique and universal. Such a system could comprise a latitude-longitude positional computer and a great circle steering computer supplemented by a square grid steering computer. The great circle computer would be used to guide the aircraft on the shortest path to its destination and the square grid computer to give a high accuracy of closing on the last 200 miles.

A slightly simpler system results if the great circle and square grid steering computers are replaced by the trapezoidal type. Often the operational requirements, particularly for commercial flights, restrict air traffic to predetermined lanes and turning points. Where such routes do not extend too close to the polar regions and when the turning points are separated by less than 1000 miles, the trapezoidal steering computer is a suitable compromise.

For flights within an area of a few hundred miles square, a simple square grid positional computer, together with a square grid steering computer, will suffice; and if it is planned to fly a straight course from one place to another, the steering computer may be eliminated by referencing the square grid computer to the desired track.

The instrumentation of a basic system favored by the authors makes use of three input organs, the radar air-speed indicator and north reference, coupled to a velocity triangle computer and a square grid positional computer. Then for many flights, the square grid computer may be referenced along and across track. If more complex patterns must be flown, or if there is a need for greater flexibility, the basic square grid computer may be used as the first section of one of the more complex computer systems described above.

A CONCLUDING REVIEW

The foregoing paper has described in varying detail the parts which make up a complete navigational system. The resulting instrument designs are flexible of application, so that several complete systems may be based on selections of them, or on arrangements consisting of some of these instruments used in conjunction with units developed elsewhere.

Any of these navigational systems is a hybrid of three techniques: Doppler radar, velocity computation based on air data, and inertial navigation; and as is usual with a hybrid system, the relative importance of the various

parts within the system is determined largely by the state of development of each separate technique.

In the systems here described the dominant technique is the Doppler radar. The velocity computation is used only for the short periods of time when there is no radar signal, and the inertial device is merely a north reference. If this paper had been written six years ago, the velocity computer would have dominated; while in six years' time the inertial devices may become far more important and the radar may play the supporting role. Other input organs such as automatic radio star trackers may become feasible due to increased mastery of UHF techniques. As input devices of greater accuracy become available, the accuracy of the positional computer may also need improvement. This is expected to favor the development of the digital over the electrical or mechanical analog devices.

There has been, for some time, a need for a lightweight, self-contained navigational system, which is independent of time, place, and weather conditions. This paper has presented some of the more recent work in this field; in particular, a system of navigation consistent with present techniques and sufficiently flexible to permit the incorporation of future developments.

ACKNOWLEDGMENT

The navigation project has required the cooperation of many workers and several organizations. The authors have done their best to make the many acknowledgments required, and ask indulgence for any inadvertent omissions.

The design authority also responsible for the early system analysis, some of the electronic design, some of the early mechanical design, the supervision of flight trials, and interpretation of results was the D.R.T.E. Electronics Lab., Defence Res. Board, Can. The personnel included B. A. Walker, R. K. Brown in charge, and J. N. Barry, J. Bloom, D. R. Boulding, W. W. H. Clarke, G. T. Lake, J. A. Moffatt, and W. W. Russell.

The industrial contractor responsible for the design of antennas, microwave system, and RF amplifiers, and also the construction of the user trials models was the Canadian Marconi Company of Montreal. The personnel here included K. C. M. Glegg, H. Hansard, A. Poznanski, and R. Reeves.

The transistor radar and digital position computer work was performed within the D.R.T.E. Electronics Lab. The circuit and system design was by N. F. Moody, R. K. Brown, and P. M. Thompson, with contributions by J. Mitchell, C. A. Franklin, R. J. Bibby, R. Gagne, G. T. Lake, G. H. Booth, J. H. Ganton, R. S. Gruno, A. E. Gagnon, Y. Fujimoto, and G. St. Amand.

The mechanical design and installation of radome and antenna were performed by G. W. Morton, N. A. Harrison, and E. P. Doherty of the D.R.T.E. Mechanical Section. All the flight testing used the aircraft and facilities of the Central Experimental and Proving Establishment, R.C.A.F., Ottawa.

The CAA Doppler Omnirange*

S. R. ANDERSON†, MEMBER, IRE, AND R. B. FLINT‡, ASSOCIATE MEMBER, IRE

Summary—This paper describes a VHF omnirange, the operation of which is based upon an application of the Doppler effect. Progress in the development of a Doppler VOR is outlined. The Doppler VOR has a measured 7-to-1 improvement with respect to siting effects when compared to the conventional four-loop VHF omnirange. The wave transmitted by the Doppler VOR is horizontally polarized. The unique interchange in the functions of the 30-cps amplitude modulation and the 30-cps frequency modulation of a 9.96-kc subcarrier is described. The Doppler VOR is shown to be compatible with present VHF omnirange receiving equipment.

The results reported in this paper were obtained by experimentation and mathematical analysis.

INTRODUCTION

SINCE 1952, the VHF omnirange (VOR) has been widely used and has become the International Civil Aviation Organization (ICAO) short-range air navigational system.¹ The increased use of the VOR in air traffic control operations places more stringent requirements on the location of omnirange stations. A VOR location best suited geographically for Federal Airways use is occasionally in an environment where reflecting objects cause excessive course bends and scalloping. Further, the environment of an established facility may change.

The effects of various siting conditions on omnirange courses have been studied for the purpose of defining an offending object and establishing criteria for VOR site location.^{2,3} The period and amplitude of course bends and scalloping are related to the distance between the facility and the reflecting object, and to the projected area of the object normal to the arriving signal. This relationship allows the source of a reflection to be identified from a flight recording of the course information. Because the effective azimuthal distribution of total VOR RF energy is of necessity nondirectional, the solution to the reflection problem has been to remove the offending object or relocate the station. These procedures are expensive and in some cases are impossible.

The recent development at the CAA Technical Development Center of the compatible Doppler VOR⁴ ap-

pears to be the most promising solution to VOR siting problems at difficult locations. It was decided in February, 1958, to determine the characteristics of a compatible Doppler VOR. Development progressed rapidly, leading to a flight test at Indianapolis on April 4, 1958, of a Doppler VOR utilizing 20 loops with an aperture of 17 feet.

Although a substantial improvement in siting effect was obtained, it was realized that a larger aperture would afford a further improvement. In June, 1958, an experimental 50-loop Doppler having an aperture of 44 feet was tested. Preliminary tests showed a minimum improvement of 5 to 1 over the conventional omnirange.

In August, 1958, a prototype of the experimental Doppler VOR was installed and tested in a wooded area at Charleston, S. C., where the conventional four-loop VOR was already installed. Trees had been cleared for approximately 1000 feet from the station and the four-loop antenna system was mounted on a 75-foot high, 150-foot diameter counterpoise. The maximum scalloping error of the four-loop VOR was $\pm 2.8^\circ$. Under the same conditions, when the Doppler VOR was installed, maximum scalloping error was $\pm 0.4^\circ$.

This paper presents a study of the Doppler VOR, and the results of the development at Indianapolis, Ind., and operational tests conducted at Charleston, S. C.

DESCRIPTION

The Doppler VOR, as the name implies, creates an effect in the received signal that is similar to the commonly observed Doppler effect normally associated with sound. The cause of the Doppler effect in the received signal is the motion or position change of one of the two sources of radiation. The motion of this source is accomplished by feeding successively each of 50 VHF horizontal loop antennas located on a circle 44 feet in diameter, thereby causing the point of the radiation source to rotate around this circle. The virtual source is rotated at 30 revolutions per second. A Doppler shift of ± 480 cps created during each revolution imposes a sinusoidal frequency modulation upon the signal at the receiver.

A second source of radiation is obtained from a VHF loop antenna located in the center of the circle of 50 antennas. The RF energy fed to the center antenna is amplitude-modulated at 30 cps to produce the reference phase 30-cps signal in the receiver.

Since the center antenna operates on the VOR channel frequency while the 50 antennas operate 9960 cps above the VOR channel frequency, the ± 480 -cps Doppler shift is also imposed upon the 9960-cps difference

* Original manuscript received by the IRE, December 2, 1958; revised manuscript received, February 20, 1959.

† U. S. Dept. of Commerce, Civil Aeronautics Administration, Tech. Dev. Center, Indianapolis, Ind.

¹ H. C. Hurley, S. R. Anderson, and H. F. Keary, "The CAA VHF omnirange," *PROC. IRE*, vol. 39, pp. 1506-1520; December, 1951.

² S. R. Anderson and H. F. Keary, "VHF Omnirange Wave Reflections from Wires," CAA Tech. Dev. Center, Indianapolis, Ind., Rep. No. 126; May, 1952.

³ R. B. Flint and A. E. Frederick, "VHF Omnirange Reflections from a Single Tree," CAA Tech. Dev. Center, Indianapolis, Ind., Rep. No. 314; June, 1957.

⁴ A non-compatible Doppler omnirange is reported by P. G. Hansel, "Doppler-effect omnirange," *PROC. IRE*, vol. 41, pp. 1750-1756; December, 1953.

frequency.⁵ The phase of the 30-cps modulation of this frequency-modulated difference frequency varies directly with the bearing of the receiver from the station. The resulting 30-cps signal in the reference channel of the receiver therefore varies in phase with azimuth.

Two transmitters are used to supply power to the antenna system. The transmitter that supplies RF energy to the center antenna is high level modulated with voice and identification signals. The RF output from this transmitter is also mechanically modulated at 30 cps prior to being applied to the center antenna. The output power of this transmitter is approximately ten times that of the transmitter output power fed to the circle of 50 antennas.

The transmitter that supplies power to the circle of antennas has no modulation applied to it. The distribution of the RF energy for the 50 antennas on the circle is accomplished through the use of a differential-capacitor distributor. The action of the distributor, by feeding successively each antenna on the circle, simulates one antenna on an arm 22 feet long rotated at 1800 rpm.

The radiation pattern of the center antenna is circular. The radiation pattern of each antenna on the circle approaches an ellipse with some variation depending upon the size of the counterpoise and the parasitic currents in adjacent antennas.

Fig. 1 is a view of the experimental Doppler omnirange at Indianapolis, Ind. The counterpoise is 150 feet in diameter and 10 feet above ground level.

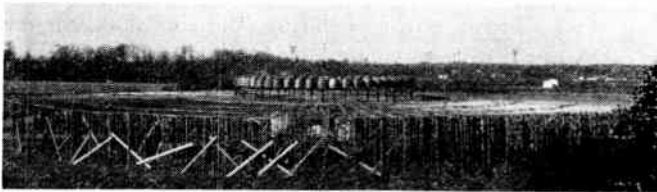


Fig. 1—The experimental Doppler omnirange at Indianapolis, Ind.

THEORETICAL CONSIDERATIONS

In an attempt to explain briefly the operation of the CAA Doppler omnirange, an ideal antenna will be assumed, *i.e.*, an antenna that is located on the circumference of a circle and rotated around the center of the circle at a uniform velocity of 30 revolutions per second.⁶ The aircraft is assumed to be motionless. Fig. 2 displays the situation in the horizontal plane where:

- r = radius of the circle
- ϕ = azimuth angle of the aircraft
- R = distance from the center of antenna rotation to the aircraft

⁵ The concept that the Doppler omnirange be made compatible with the conventional omnirange by providing a 9.96-kc difference frequency was first suggested to the authors by H. C. Hurley, Chief, Receiver Branch, CAA Tech. Dev. Center, Indianapolis, Ind.

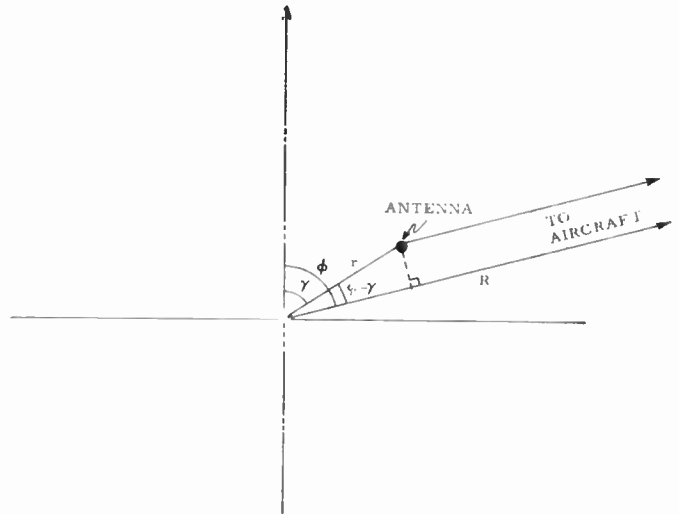


Fig. 2—The coordinate system for one transmitting antenna.

- γ = azimuth angle of antenna
- $\beta_2 = 2\pi/\lambda_2$
- λ_2 = wavelength.

The wave traveling from the antenna to the aircraft undergoes an RF phase delay.

$$\psi = \beta_2 [R - r \cos(\phi - \gamma)]. \tag{1}$$

The antenna is rotated in the counterclockwise direction at a uniform velocity ρt , where t is time and $\rho/2\pi$ is a frequency, hence:

$$\gamma = \rho t. \tag{2}$$

The signal at the aircraft may then be expressed as:

$$E_2 = \cos [\omega_2 t + \beta_2 r \cos(\phi - \rho t) - \beta_2 R] \tag{3}$$

where $\omega_2/2\pi = f_2 =$ frequency of the current in the rotating antenna.

$$m_f = \frac{\text{variation of radio frequency away from the mean frequency}}{\text{modulation frequency}} = \text{modulation index}$$

as given by Terman,⁶ and therefore:

$$m_f = \beta_2 r. \tag{4}$$

The instantaneous frequency, f_{inst} , of the wave represented by (3) is obtained by differentiation of the instantaneous angular velocity:

$$f_{inst} = \frac{1}{2\pi} \frac{d}{dt} [\omega_2 t + \beta_2 r \cos(\phi - \rho t) - \beta_2 R]$$

$$f_{inst} = f_2 + \frac{r\rho}{\lambda} \sin(\phi - \rho t). \tag{5}$$

Up to this point it has been shown that by rotating an antenna at a uniform angular velocity, driven by RF

⁶ F. E. Terman, "Radio Engineers' Handbook," McGraw-Hill Book Co., Inc., New York, N. Y.; 1943.

current of frequency f_2 , about a vertical axis, a phase-modulated wave is transmitted to the aircraft. The significant fact about the phase-modulated wave is that the phase of the modulation frequency is the azimuth angle of the aircraft, ϕ . This is clearly shown in (3). It is interesting to observe that no error exists between the electrical phase angle and the azimuth of the aircraft. In order to extract this bearing intelligence, it is necessary to put the signal into a frequency discriminator where the amplitude of the output voltage is a linear function of the instantaneous frequency of the input signal. Eq. (5), containing the bearing information ϕ , expresses the instantaneous frequency of the signal impressed upon the discriminator input. The discriminator output is therefore a wave of frequency $\rho/2\pi$ which varies as the aircraft azimuth angle, ϕ .

Applying the constants of the CAA Doppler omnirange to the preceding equations results in the following:

$$r = 21.796'$$

$$\frac{\rho}{2\pi} = 30 \text{ cps}$$

$$118 \text{ mc} > f_2 > 108 \text{ mc}$$

$$m_f = \beta_2 r = \frac{2\pi}{8.56} \times 21.796 = 16 \text{ for } f_2 = 115 \text{ mc}$$

$$f_{\text{inst}}(\text{cps}) = 115 \times 10^6 + 480 \sin(\phi - \rho t)$$

It should be noted that an *fm* wave of $m_f = 16$ and modulation frequency of 30 cps is exactly compatible with the conventional VOR.

An additional antenna is located at the center of rotation of the circle for the purpose of: 1) conveying a reference 30-cps signal, and 2) providing a carrier 9.96 kc removed from the frequency-modulated signal for converting the frequency of this signal to 9.96 kc at the output of the second detector of a conventional VOR receiver. The signal at the aircraft from the center antenna is:

$$E_1 = \cos(\omega_1 t - \beta_1 R)(1 + m \cos \rho t) \quad (6)$$

where

$\omega_1/2\pi = f_1$ = frequency of the current in the center antenna:

$$\beta_1 = 2\pi/\lambda_1$$

λ_1 = wavelength

m = amplitude modulation index.

Combining (3) and (6) represents an expression of the total signal at the aircraft:

$$E = E_1 + nE_2 \quad (7)$$

where n is a constant regulating the relative amounts of the two signals.

$$E = \left\{ \cos(\omega_1 t - \beta_1 R)(1 + m \cos \rho t) + n \cos[\omega_2 t + \beta_2 r \cos(\rho t - \phi) - \beta_2 R] \right\} \quad (8)$$

The first term of (8) represents an amplitude-modulated wave which conveys the reference or fixed-phase 30-cps signal, while the 30-cps variable phase signal is contained in the second term in the form of a frequency modulation.

The CAA Doppler omnirange is made compatible with the conventional omnirange and therefore the variables and constants of (8) are fixed at:

$$\begin{aligned} \frac{\omega_2}{2\pi} - \frac{\omega_1}{2\pi} &= \frac{\Delta}{2\pi} = 9.96 \text{ kc} \\ m &= n = 0.3. \end{aligned} \quad (9)$$

Fig. 3 depicts the distribution of the signals radiated by the Doppler omnirange where the reference 30-cps signal is in the form of AM sidebands while the variable 30-cps signal is contained by the FM sidebands.

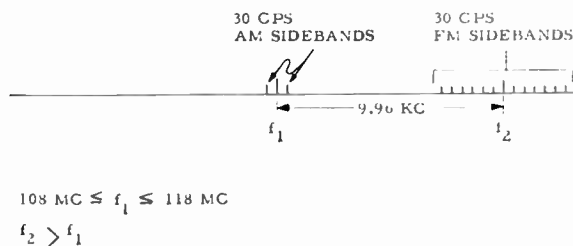


Fig. 3—Frequency spectrum showing sidebands and carrier.

Will the signal radiated by the CAA Doppler omnirange theoretically be exactly compatible with the conventional omnirange, and with a high degree of bearing accuracy? To answer this question, the signal expressed by (8) is put into a linear detector and the output is analyzed. This is ideally accomplished mathematically by analyzing the envelope of (8).⁷ Combining (8) and (9) and assuming $\beta_2 R = \beta_1 R$

$$E = \sqrt{(1 + m \cos a + n \cos b)^2 + n^2 \sin^2 b} \cdot \cos(\omega_1 t - \beta_1 R - \psi), \quad (10)$$

where

$$a = \rho t$$

$$b = \Delta t + \beta_2 r \cos(\rho t - \phi)$$

$$\psi = \tan^{-1} \left\{ \frac{-n \sin \Delta t + \beta_2 r \cos(\rho t - \phi)}{1 + m \cos \rho t + n \cos[\Delta t + \beta_2 r \cos(\rho t - \phi)]} \right\}$$

The output of the linear detector is proportional to

$$|E| = \sqrt{A^2 + B^2} \quad (11)$$

where

$$A = 1 + m \cos a + n \cos b$$

$$B = n \sin b.$$

By means of the binomial expansion where $(B/A)^2 \leq 1$

⁷ See C. B. Aiken, "Theory of the detection of two modulated waves by a linear rectifier," *Proc. IRE*, vol. 21, pp. 601-629; April, 1933.

$$|E| = A + \frac{B}{2} - \frac{1}{8} \frac{B^2}{A} + \frac{3}{48} \frac{B^3}{A^2} - \dots \quad (12)$$

$$|E| = 1 + m \cos a + n \cos b + \frac{1}{2} n \sin b - \frac{1}{8} \frac{n^2 \sin^2 b}{1 + m \cos a + n \cos b} + \frac{3}{48} \frac{n^3 \sin^3 b}{(1 + m \cos a + n \cos b)^2} - \dots \quad (13)$$

The fifth term in (13) is expanded by means of the binomial theorem with the condition $(m \cos a + n \cos b)^2 < 1$

$$\frac{1}{1 + m \cos a + n \cos b} = 1 - (m \cos a + n \cos b) + (m \cos a + n \cos b)^2 - (m \cos a + n \cos b)^3 + \dots \quad (14)$$

By placing (14) into (13) and retaining only the first five terms of (13):

$$|E| = 1 + m \cos a + n \frac{\sqrt{5}}{2} \cos \left(b - \tan^{-1} \frac{1}{2} \right) - \frac{n^2}{8} [\sin^2 b - \sin^2 b (m \cos a + n \cos b) + \sin^2 b (m \cos a + n \cos b)^2 - \dots] \quad (15)$$

Finally, by combining terms of the same frequency in (15), letting $m=n=0.3$, and neglecting the dc components, the output of the linear detector becomes:

$$|E| = 0.3017 \cos \rho t + 0.337 \cos [\Delta t + \beta_2 r \cos (\rho t - \phi) - 26.5^\circ] - 0.00169 \cos \rho t \cos [2\Delta t + 2\beta_2 r \cos (\rho t - \phi)] - 0.00084 \cos [3\Delta t + 3\beta_2 r \cos (\rho t - \phi)] + \dots \quad (16)$$

Eq. (16) indicates clearly that the ideal linear detector will transform the signal from the CAA Doppler omnirange into a distortionless reference 30-cps signal, and a distortionless 9.96-kc subcarrier signal which is frequency modulated with a 30-cps variable phase signal. The phase of the variable phase signal is the azimuth angle. Therefore the bearing accuracy is expected to be excellent. The third and fourth terms of (16) are at a frequency of 2×9.96 kc and 3×9.96 kc, respectively, and would be expected to have negligible effect on the omnirange receiver performance. The terms neglected in the series of (16) are expected to be smaller than any term included, and of higher frequency.

It is demonstrated that a Doppler omnirange compatible with the conventional VHF omnirange is feasible. It is generally found in the field of navigational aids for aircraft that siting error in bearing, caused by reflections of RF energy from objects near the transmitter, is inversely proportional to some function of the size of the antenna system. It is desired to show theoretically the

amount and nature of the improvement in siting errors that might be expected from the CAA Doppler omnirange compared with the conventional omnirange.

Fig. 4 depicts the situation to be studied where one reflecting object is located d_1 distance east of the transmitting antenna of the Doppler omnirange or the conventional omnirange. An aircraft is located at P , azimuth angle ϕ . The reflecting object radiates a circular pattern. The signals arriving at the aircraft follow two paths, i.e., along r_0 and r_1 . The signal traveling along r_0 contains the correct bearing information, ϕ , while the reflected signal contains bearing information of $\phi = 90^\circ$.

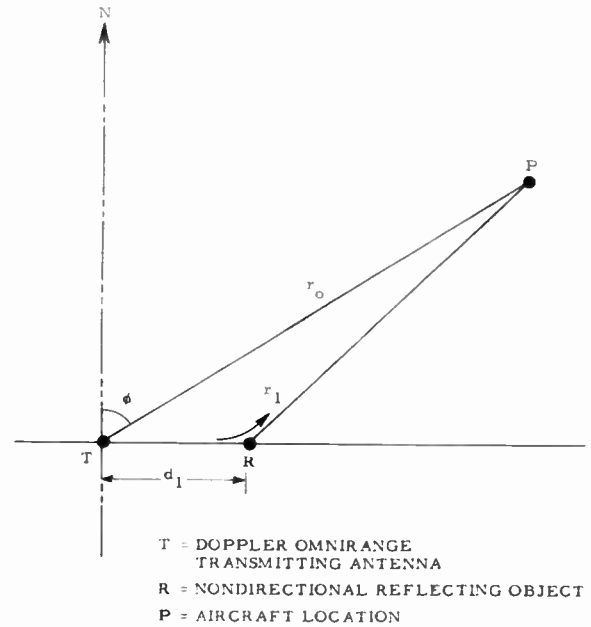


Fig. 4—A plan view of the Doppler omnirange and one reflecting object.

The symbols in the following equations have the same meaning as the corresponding ones in the first part of this paper. The signal⁸ at point P is derived as follows: Direct wave:

$$E(\phi) = \frac{\cos \omega_1 t - \beta_1 r_0}{r_0} (1 + m \cos \rho t) + \frac{n}{r_0} \cos [\omega_2 t + \beta_2 r \cos (\rho t - \phi) - \beta_2 r_0] \quad (17)$$

Reflected wave:

$$E(90^\circ) = \frac{A \cos (\omega_1 t - \beta_1 r_1 - \delta)}{r_1} (1 + m \cos \rho t) + \frac{A_n}{r_1} \cos (\omega_2 t + \beta_2 r \sin \rho t - \beta_2 r_1 - \delta) \quad (18)$$

⁸ The mathematical analysis is similar to one appearing in a paper by M. G. Crosby, "Frequency modulation propagation characteristics," Proc. IRE, vol. 24, pp. 898-913; June, 1936. It was suggested to the authors in an unpublished memorandum by F. Moskowitz of Rome Air Dev. Center, Griffiss AFB, Rome, N. Y.

where

- δ = RF phase change caused by the reflector.
- r = radius of circle of Doppler omnirange antenna system.
- r_0 = direct path.
- r_1 = indirect path via reflecting object.
- A = ratio of reflected signal divided by direct signal appearing at aircraft.

$$\phi_P = \beta_2 r \cos(\rho t - \phi) - \beta_2 r_0 + .1 \sin [\beta_2 r \sin \rho t - \beta_2 r_1 - \delta - \beta_2 r \cos(\rho t - \phi) + \beta_2 r_0]. \quad (20)$$

$$\phi_P = \beta_2 r \cos(\rho t - \phi) - \beta_2 r_0 + .1 \sin \left[2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \cdot \sin \left(\rho t - \frac{90^\circ + \phi}{2} \right) + \beta_2(r_0 - r_1) - \delta \right]. \quad (21)$$

$$\begin{aligned} \phi_P &= \beta_2 r \cos(\rho t - \phi) - \beta_2 r_0 \\ &+ 2.1 \cos [\beta_2(r_0 - r_1) - \delta] \sum_{n=0}^{\infty} J_{2n+1} \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \sin \left[(2n+1) \left(\rho t - \frac{90^\circ + \phi}{2} \right) \right] \\ &+ .1 \sin [\beta_2(r_0 - r_1) - \delta] J_0 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \\ &+ 2.1 \sin [\beta_2(r_0 - r_1) - \delta] \sum_{n=1}^{\infty} J_{2n} \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \cos \left[2n \left(\rho t - \frac{90^\circ + \phi}{2} \right) \right], \end{aligned} \quad (22)$$

Since

$$\frac{1}{r_0} \approx \frac{1}{r_1},$$

the total signal at the aircraft becomes

$$\begin{aligned} E(\phi) + E(90^\circ) &= \frac{1}{r_0} \{ [\cos(\omega_1 t - \beta_1 r_0) + A \cos(\omega_1 t - \beta_1 r_1 - \delta)] \\ &\cdot (1 + m \cos \rho t) + n \cos [\omega_2 t + \beta_2 r \cos(\rho t - \phi) - \beta_2 r_0] \\ &+ .1 n \cos [\omega_2 t + \beta_2 r \sin \rho t - \beta_2 r_1 - \delta] \}. \end{aligned} \quad (19)$$

The last two terms in (19) represent the desired phase modulation signal which conveys the azimuth intelligence and the reflected phase modulation signal, respec-

where:

$$\left. \begin{aligned} J_0(\) \\ J_{2n+1}(\) \\ J_{2n}(\) \end{aligned} \right\} \text{ are Bessel functions of the first kind.}$$

The output of an ideal discriminator is proportional to the instantaneous frequency of the input signal. The instantaneous frequency is equal to

$$\frac{1}{2\pi} \frac{d\phi_P}{dt}.$$

Therefore,

$$K \frac{d\phi_P}{dt} = i_P = \text{discriminator output current}$$

where $K = a$ constant of proportionality.

$$\begin{aligned} i_P &= -\beta_2 r \rho \sin(\rho t - \phi) + 2.1 \rho \cos [\beta_2(r_0 - r_1) - \delta] J_1 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \cos \left(\rho t - \frac{90^\circ + \phi}{2} \right) \\ &- 4.1 \rho \sin [\beta_2(r_0 - r_1) - \delta] J_2 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \sin (2\rho t - 90^\circ - \phi) \\ &+ 6.1 \rho \cos [\beta_2(r_0 - r_1) - \delta] J_3 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \cos \left(3\rho t - \frac{270^\circ - 3\phi}{2} \right) + \dots \end{aligned} \quad (23)$$

tively. In order to determine the bearing error incurred because of a reflection from an object at 90° azimuth, the last two terms of (19) are combined. It is only the instantaneous phase angle of the resultant signal that is required for a determination of bearing error.

$\phi_P \approx$ instantaneous phase angle of the resultant signal, $.1 \leq 0.1$

In the conventional omnirange receiver discriminator output are filters that favor the fundamental frequency, $\rho/2\pi$, and tend to reject the harmonics of $\rho/2\pi$. Assuming that only the fundamental frequency terms of (23) are used in the receiving equipment for determining the bearing, the infinite number of terms after the first two are neglected. Combining the first two terms of (23):

$$i_P = \pm \sqrt{(\beta_2 r \rho)^2 + (2.1 \rho)^2 \cos^2 [\beta_2 (r_0 - r_1) - \delta] J_1^2 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) + 4\beta_2 r \rho^2 A \cos [\beta_2 (r_0 - r_1) - \delta] J_1 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \sin \left(\frac{\phi - 90^\circ}{2} \right)} \cdot \sin \left[\rho l - \phi + \tan^{-1} \left[\frac{2.1 A \cos [\beta_2 (r_0 - r_1) - \delta] J_1 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \cos \left(\frac{\phi - 90^\circ}{2} \right)}{\beta_2 r + 2.1 A \cos [\beta_2 (r_0 - r_1) - \delta] J_1 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \sin \left(\frac{\phi - 90^\circ}{2} \right)} \right] \right]. \quad (24)$$

Finally, the resultant $\rho/2\pi$ phase angle is obtained as shown in (24). The error caused by the reflected wave is the $\tan^{-1}(\)$ term. In (24) it can be seen that if no reflection occurs, $A = 0$, then

$$i_P = -\beta_2 r \rho \sin(\rho l - \phi). \quad (25)$$

The error, designated by x , caused by the reflecting object located at 90° azimuth becomes:

$$x \approx \frac{2.1 A J_1 \left(2\beta_2 r \sin \left(\frac{90^\circ - \phi}{2} \right) \right) \cos \left(\frac{\phi - 90^\circ}{2} \right) \cos [\beta_2 (r_0 - r_1) - \delta]}{\beta_2 r}, \quad (26)$$

since $A \leq 0.1$.

Eq. (26) makes clear that x , the bearing error, varies as the aircraft moves, because $(r_0 - r_1)$ is a function of the aircraft position.⁹ Hence, this error can be seen to oscillate as $\cos [\beta_2 (r_0 - r_1) - \delta]$ varies from plus one to zero, then to minus one, etc. This is called scalloping of the bearing indication. The maximum values of (26) are plotted on Fig. 5 with $A = 0.1$. For comparison, the scalloping associated with a conventional omnirange¹⁰ is plotted on Fig. 6 where the scalloping of the Doppler omnirange is plotted to the same scale. A comparison of the scalloping of the two omnirange systems is shown in Fig. 7. It is interesting to note that the scalloping of the Doppler omnirange, as indicated by Fig. 7, has a reduction factor of 0.1 or less for most azimuths. This means that the scalloping associated with a conventional omnirange will be reduced to at least 0.1 at most azimuths by replacing the conventional omnirange with a Doppler omnirange. Fig. 7 also indicates that there is a sector of approximately 10° centered on $\phi = 90^\circ$ where little reduction in scalloping will be obtained by replacing a conventional omnirange by a Doppler omnirange. This phenomenon is borne out qualitatively in practice where

⁹ This is treated in some detail by S. R. Anderson and H. F. Keary, "VHF Omnirange Wave Reflections From Wires," CAA Tech. Dev. Center, Indianapolis, Ind., Rep. No. 126; May, 1952.

¹⁰ See S. R. Anderson, R. B. Flint, W. L. Wright, and J. Turk, "Preliminary Tests of a Precision VHF Omnirange," CAA Tech. Dev. Center, Indianapolis, Ind., Rep. No. 355; August, 1958.

the reflecting source consists of two vertical towers approximately 60 feet high located 880 feet at 266° azimuth and 285° azimuth, respectively, from the omnirange. The offending towers are the two located immediately behind the Doppler omnirange shown in Fig. 1. The sinusoidal scalloping on Fig. 18 in the vicinity of 250° azimuth and 270° azimuth is caused by the two towers.

EQUIPMENT

A block diagram of the Doppler omnirange equipment is shown in Fig. 8. The antenna system uses 51 VHF loop antennas designed to operate in the 108 to 118-mc range. They are supported by pedestals at approximately one-half wavelength above a wire screen counterpoise. Type RG-8/U transmission line is used throughout. Fig. 9 is a view of the Doppler VOR antenna system.

The transmitter and modulation equipment used to supply RF power to the center antenna are of conventional design. This transmitter is capable of supplying 200 watts RF power continuously. Plate modulation is employed to provide the station with voice intelligence and a keyed 1020-cps identification signal.

The 30-cps modulation of the transmitter output is accomplished by means of a capacity goniometer and a transmission line VHF bridge network. It will be noted in Fig. 8 that the transmitter output is fed directly to a power divider which is followed by a modulation eliminator. The purpose of the modulation eliminator used here is the same as that in the conventional VOR system.¹¹ The smaller amount of the RF power divider apportionment, having the modulation removed, is fed to the input of a capacity goniometer. The capacity goniometer is the conventional type used in the standard VOR. One output of the goniometer is fed to a Terma-

¹¹ H. C. Hurley, S. R. Anderson, and H. F. Keary, "The CAA VHF Omnirange," CAA Tech. Dev. Center, Indianapolis, Ind., Rep. No. 113; June, 1950.

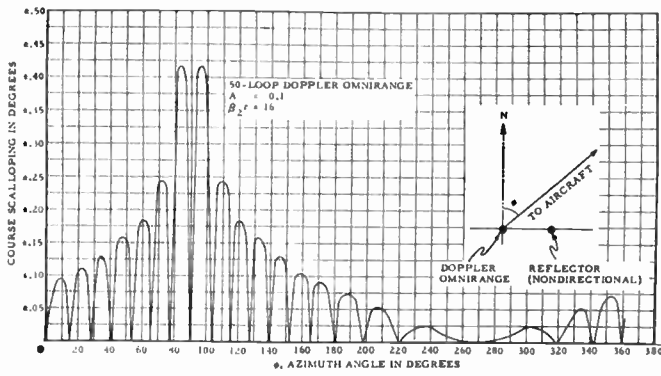


Fig. 5—The theoretical Doppler omnirange course scalloping caused by one reflecting object vs azimuth angle.

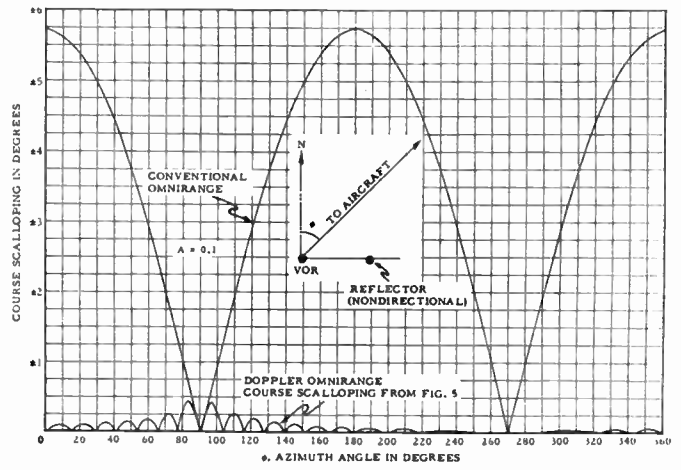


Fig. 6—Theoretical course scalloping caused by one reflecting object vs azimuth angle.

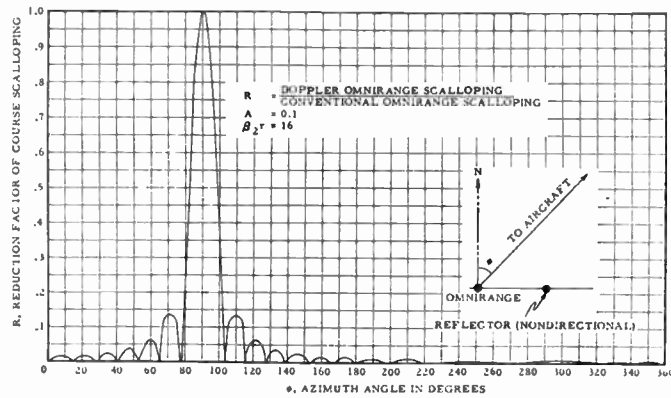


Fig. 7—Theoretical reduction in VOR course scalloping effected by the Doppler omnirange.

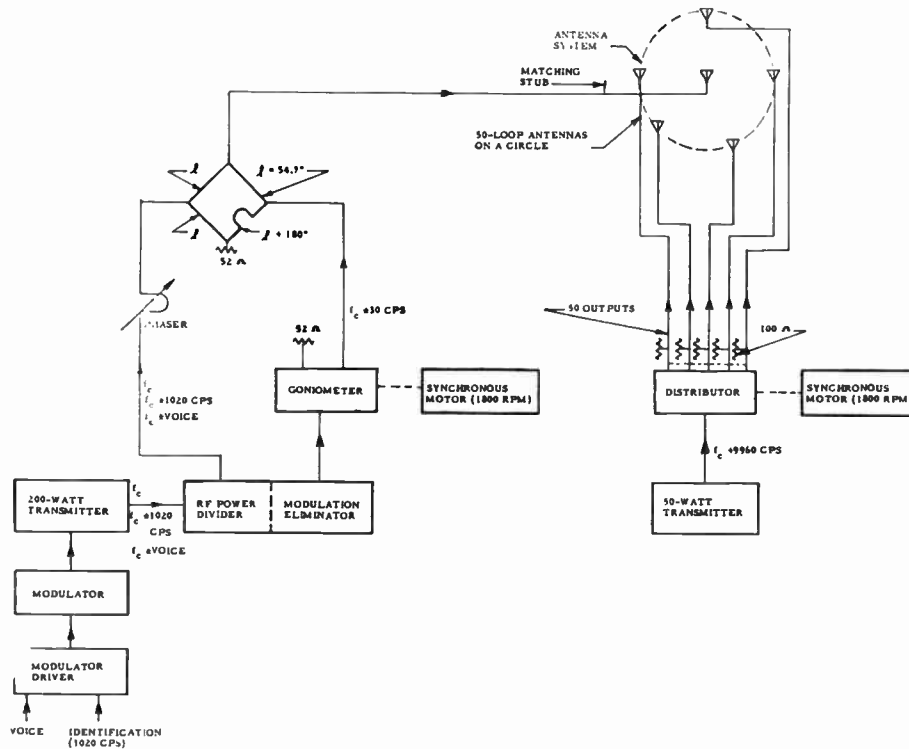


Fig. 8—Block diagram of Doppler omnirange equipment.

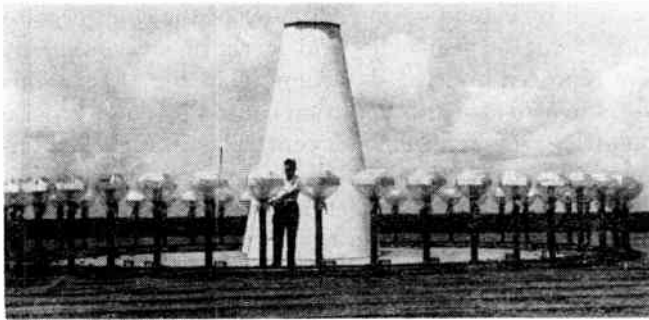
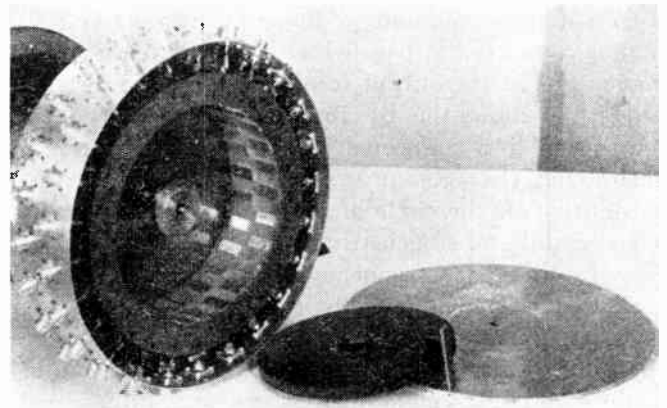


Fig. 9—The Doppler VOR antenna system.

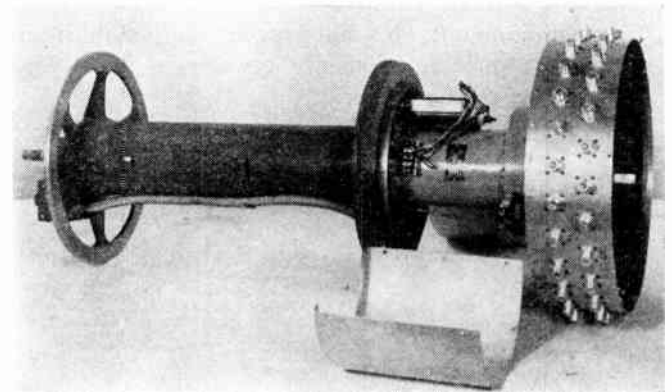
line dummy load. The other output, consisting only of 30-cps sideband energy, is fed to an input corner of the VHF bridge network. This sideband energy is combined with the carrier energy at the output corner of the bridge, thereby amplitude-modulating the carrier to approximately 30 per cent. The RF power at the output corner of the bridge is fed through coaxial transmission line to the center antenna.

Present plans include the possible use of a 30-cps alternator to modulate the carrier. This would preclude the necessity of the bridge network, goniometer, and the modulation eliminator used to provide 30-cps modulation of the carrier energy.

A separate transmitter is used to supply RF power to the circle of 50 antennas. This transmitter is crystal-controlled at a frequency 9960 cps higher than the frequency of the carrier energy of the center antenna. The differential-capacitor distributor feeds each antenna in succession. The RF power applied to the input of the distributor unit is fed through a rotating element of a capacitor in the form of concentric cylinders, thence through coaxial line in a bore along the motor shaft to a single plate located on the periphery of a disk of insulating material. The disk, approximately 9 inches in diameter, is attached to the shaft of the synchronous motor and rotates at 1800 rpm. Fifty stationary plates encompass the arc of the rotor plate.



(a)



(b)

Fig. 10—The Doppler VOR differential capacitor distributor. (a) Distributor head, (b) distributor unit.

the rotor revolves. The distributor output terminations are connected directly to the stator plates. The voltage waveshape at the output terminations is sinusoidal. Ideally, any one output is zero during 345.6° of each rotation of the rotor. During the remaining 14.4°, the output voltage rises to a maximum and falls back to zero. The sinusoids are equally spaced on a circle with adjacent spacing (*i.e.*, maximum to maximum) of 7.2°. Expressed mathematically:

$$\begin{array}{ll}
 E_1 = \sin 12.5\theta & 0^\circ \leq \theta \leq 14.4^\circ \\
 E_2 = \sin (12.5\theta - 90^\circ) & 7.2^\circ \leq \theta \leq 21.6^\circ \\
 E_3 = \sin (12.5\theta - 180^\circ) & 14.4^\circ \leq \theta \leq 28.8^\circ \\
 E_n = \sin [12.5\theta - 90^\circ(n - 1)] & 7.2(n - 1) \leq \theta \leq 14.4^\circ + 7.2^\circ(n - 1) \\
 \dots & \dots \\
 \dots & \dots \\
 E_{50} = \sin (12.5\theta - 4410^\circ) & 352.8^\circ \leq \theta \leq 367.2^\circ,
 \end{array}$$

The disassembled distributor is shown in Fig. 10. Note that alternate stator plates are displaced vertically. This is done to minimize coupling between adjacent plates. The RF energy in the rotor plate is capacitively coupled to each stator plate in succession as

where

- θ = the angle of rotation of the rotor
- E_1 = voltage of No. 1 output
- E_n = voltage of No. n output
- n = the output number.

To minimize transients, a 100-ohm nonreactive resistor was connected in parallel with each antenna transmission line at the output terminations of the distributor. Fig. 11 shows the distributor unit in normal operation with all lines connected.

Although the present system, utilizing a separate transmitter for the circle of 50 antennas, has proved to be successful and sufficiently stable with respect to the 9960-cps difference frequency, a modification is contemplated that will possibly replace this transmitter with a frequency converter unit. The frequency converter unit will be excited and controlled by the transmitter used in conjunction with the center antenna. Also, the distributor unit will be mechanically connected to the synchronous motor and reference 30-cps alternator or goniometer. This will prevent the possibility of an erroneous phase relationship between the two 30-cps signals in the VOR receiver.

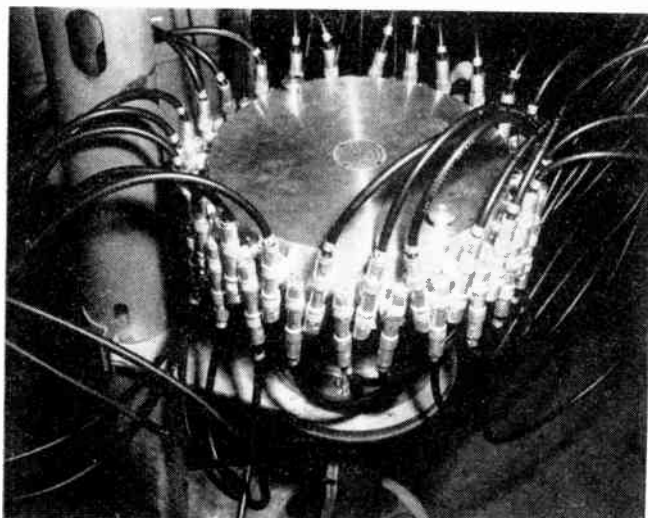


Fig. 11—Distributor unit with all lines connected as in normal operation.

TESTS

During the development of the CAA Doppler omnirange, a counterpoise 75 feet in diameter and 10 feet high, was used to support 50 loop antennas mounted 48 inches above the counterpoise and equally spaced around a 22-foot radius circle. The 22-foot radius circle is concentric with the counterpoise.

A special test was conducted where the 50 loop antennas were removed from atop the counterpoise and a single loop placed on the 22-foot radius circle. Field strength measurements were made at a distance of one-half mile with the single loop fed directly from the transmitter. The single loop was moved to every even-numbered mounting position for each measurement of field strength. The data are plotted in Fig. 12 and show a variation in field strength of ± 22.6 per cent about a mean value due to the counterpoise eccentricity effect.

A reduction of the counterpoise eccentricity effect was

obtained by extending the diameter of the counterpoise from 75 to 150 feet. The field strength measurements were repeated and the data plotted in Fig. 13. In this case, the field strength variation caused by counterpoise eccentricity effect is 9.82 per cent. This is a marked improvement over the 75-foot diameter counterpoise. There is, however, considerable modification of these patterns when the distributor and 50 loop antennas are used. The RF coupling between any driven stator plate and an adjacent stator plate in the engineering model distributor and mutual couplings between antennas cause a different pattern to be radiated in operation. The envelope of the 9960-cps signal is proportional to the amplitude of the field strength radiated by the 50 loops. An oscillogram of the 9960-cps signal in a receiver located 200 feet distant is shown in Fig. 14. The large variations in amplitude of the 9960-cps signal, which occur at a 30-cps rate, cause a cross modulation of the 30-cps signal in the AM channel of the receiver. The cross modulation creates a bearing error which varies with azimuth, as evidenced by the shape of the Doppler bearing error curve of Fig. 15. The resulting maximum bearing error is approximately $\pm 1^\circ$.

The small sinusoidal changes in amplitude in Fig. 14 occur at a rate of 1500 cps, and are due to the variation in field strength caused by a transition from one pattern when a single antenna is being fed to another pattern when two antennas are being fed. This amplitude modulation of the subcarrier causes an audible 1500-cps signal in the audio output of the receiver only in strong signal areas. Such areas are encountered when the aircraft is at an altitude of 2000 feet and less than 5 miles from the ground station. The 1500-cps signal has no effect on the course deviation indicator (CDI).

Fig. 16 shows the antenna current distribution where the rotor was stopped on one stator. The field pattern associated with these currents and the counterpoise eccentricity effect determine the field strength pattern radiated by the 50 loops.

The bearing error data obtained from orbital flights at approximately 6-mile radius are shown in Fig. 15. The bearing error curve of a conventional VOR is also presented on Fig. 15 in order that a comparison may be made between the two VOR systems.

Radial flights were made over the station to record various signals in the receiver. The recordings are displayed in Fig. 17. The reference channel 30-cps signal, which is the variable phase signal in the Doppler omnirange system, decreases in amplitude with an increase in elevation angle. This would be expected theoretically, because the motion toward or away from the observer of the virtual source decreases with an increase of the elevation angle. This decrease has a "softening" effect on the course deviation, to-from, and flag indicators. While the plane passes through the cone over the station, there is a short period of time when the flag alarm shows and the 1500-cps background level increases sharply in the audio channel.

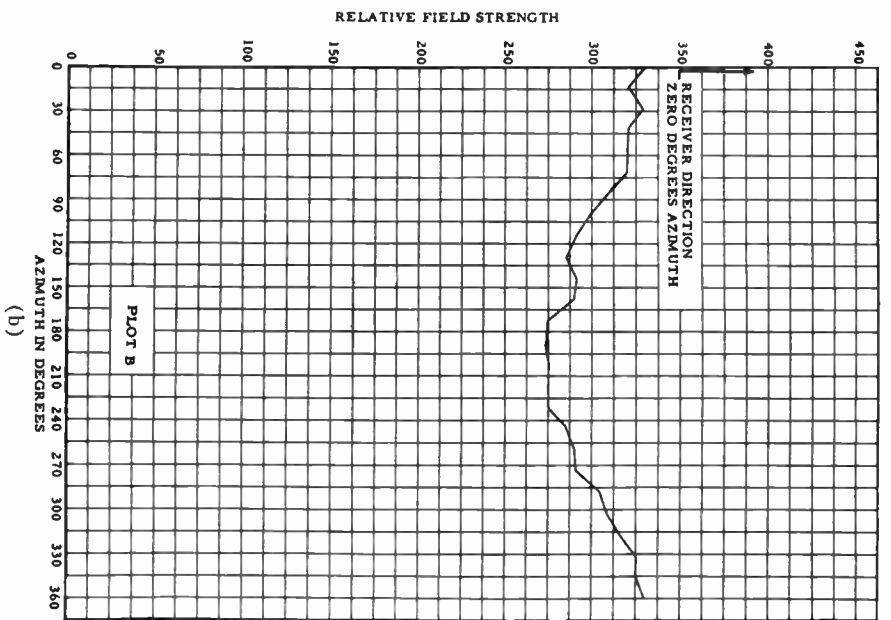
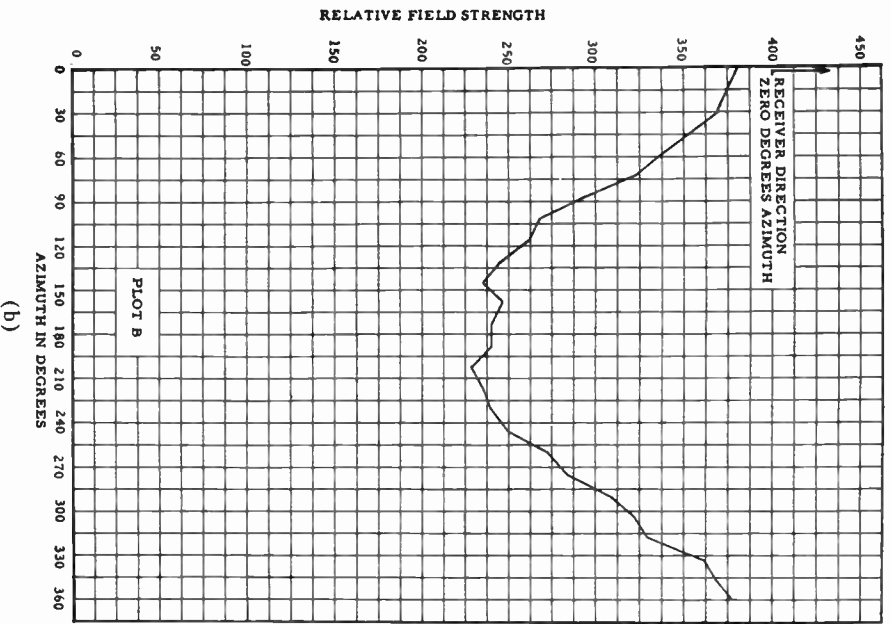
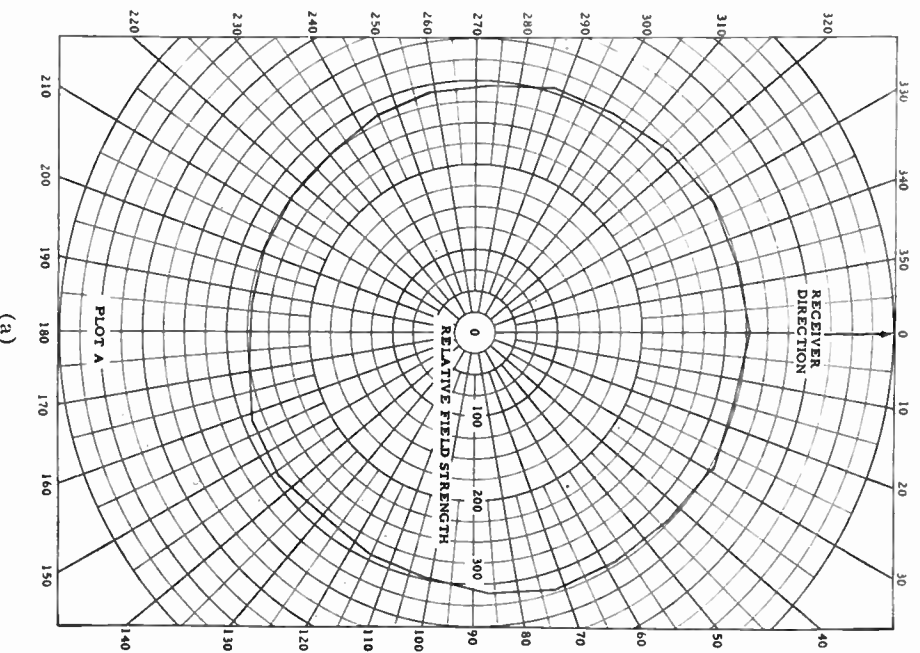
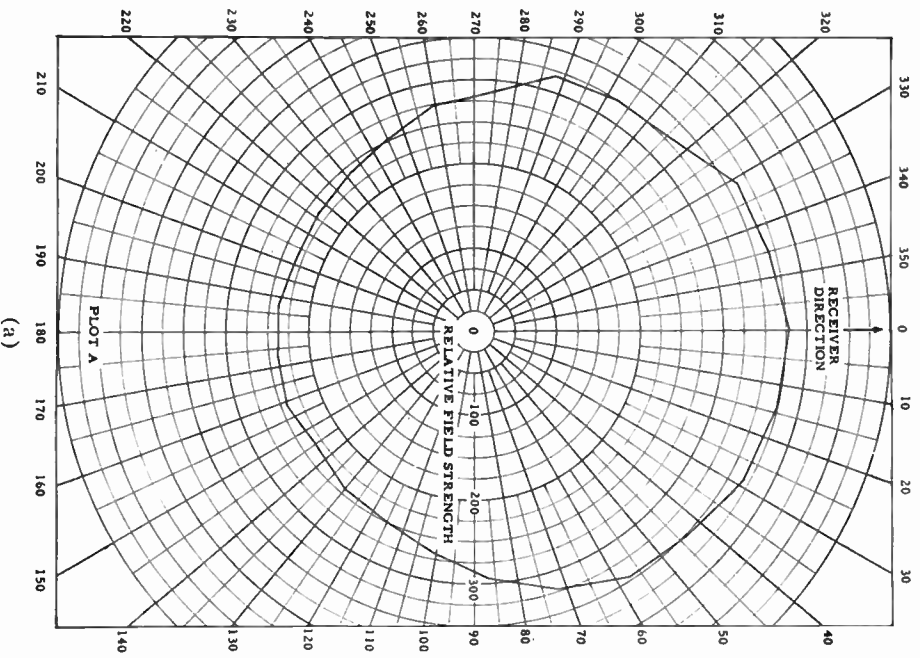


Fig. 12—Variation in field strength of one antenna caused by eccentricity effect of 75-foot-diameter counterpoise.

Fig. 13—Variation in field strength of one antenna caused by eccentricity effect of 150-foot-diameter counterpoise.

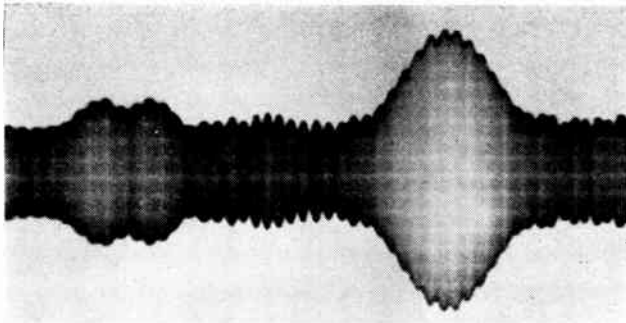


Fig. 14—Photo oscillograph of 9960-cps signal in receiver 200 feet from station.

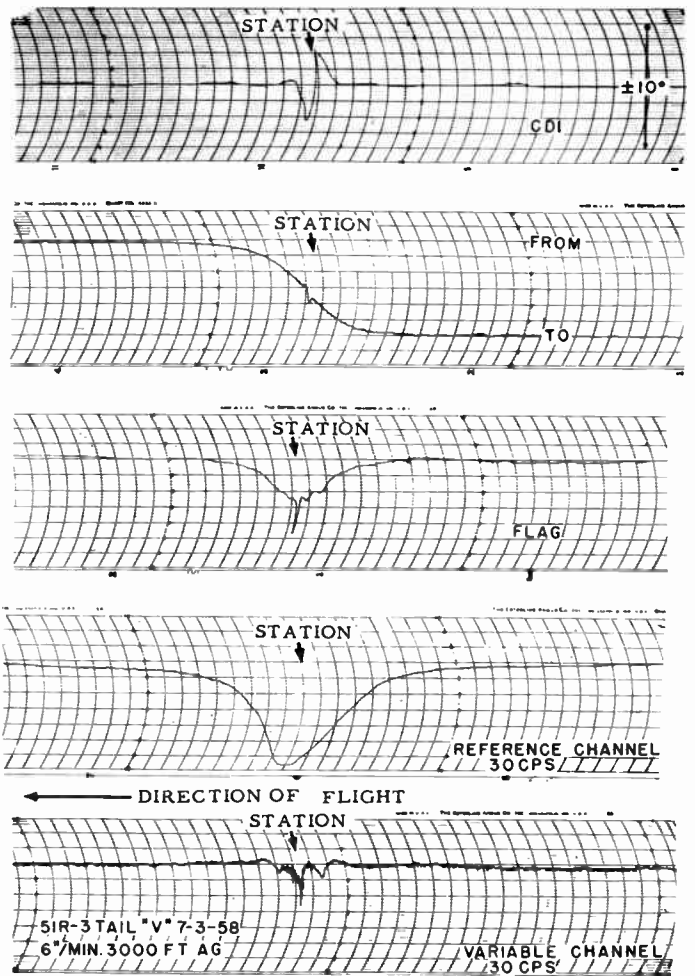


Fig. 17—Doppler VOR flight recordings.

tests on several Doppler omnirange stations and found to be less than $\pm \frac{1}{2}^\circ$.

TESTS AT CHARLESTON, S. C.

The tests on the good site at the Technical Development Center at Indianapolis, Ind., indicated so much reduction of siting error that further tests were conducted at Charleston, S. C., where a 150-foot diameter counterpoise, 75 feet above the ground, had been installed. The site consists of trees beyond 1000 feet from the tower in all directions for miles. Previous tests with a conventional four-loop VOR antenna on the tower showed a maximum scalloping of $\pm 2.8^\circ$ during a 20-mile orbital flight. The flight data on the Doppler VOR showed a maximum scalloping of $\pm 0.4^\circ$. Fig. 19 is a reproduction of the flight recordings made at the Charleston, S. C., site of the conventional omnirange and the Doppler omnirange.

CONCLUSIONS

This paper has shown by means of reasoning and experimentation that the principle of moving a source of electromagnetic radiation can be used to provide azi-

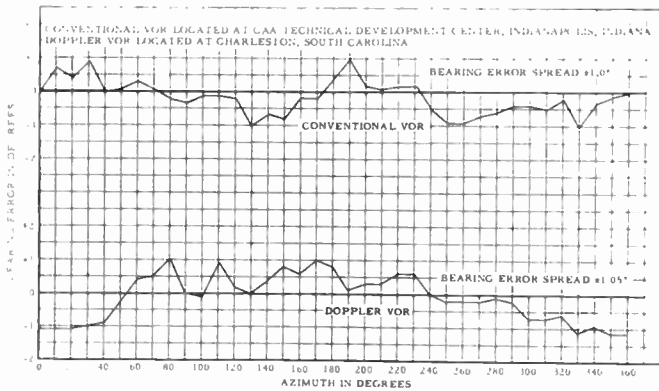


Fig. 15—Doppler VOR and conventional VOR bearing error curves.

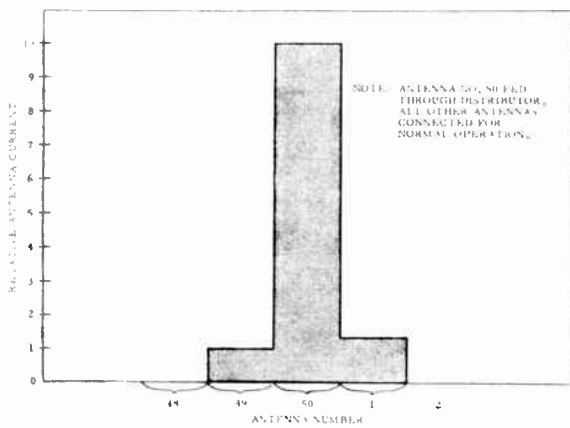


Fig. 16—Antenna current distribution with distributor rotor positioned to feed a single antenna.

Fig. 18 is a reproduction of flight recordings of the conventional VOR and the Doppler VOR located at the TDC test site. The course roughness and scalloping caused by trees and buildings noted on the conventional VOR recording are essentially eliminated in the recording of the Doppler VOR. The remaining scalloping in the 250° and 270° sectors is caused by two steel towers 60 feet high located west of the VOR site at a distance of 880 feet, pictured on Fig. 1, foreground.

Polarization error has been measured during flight

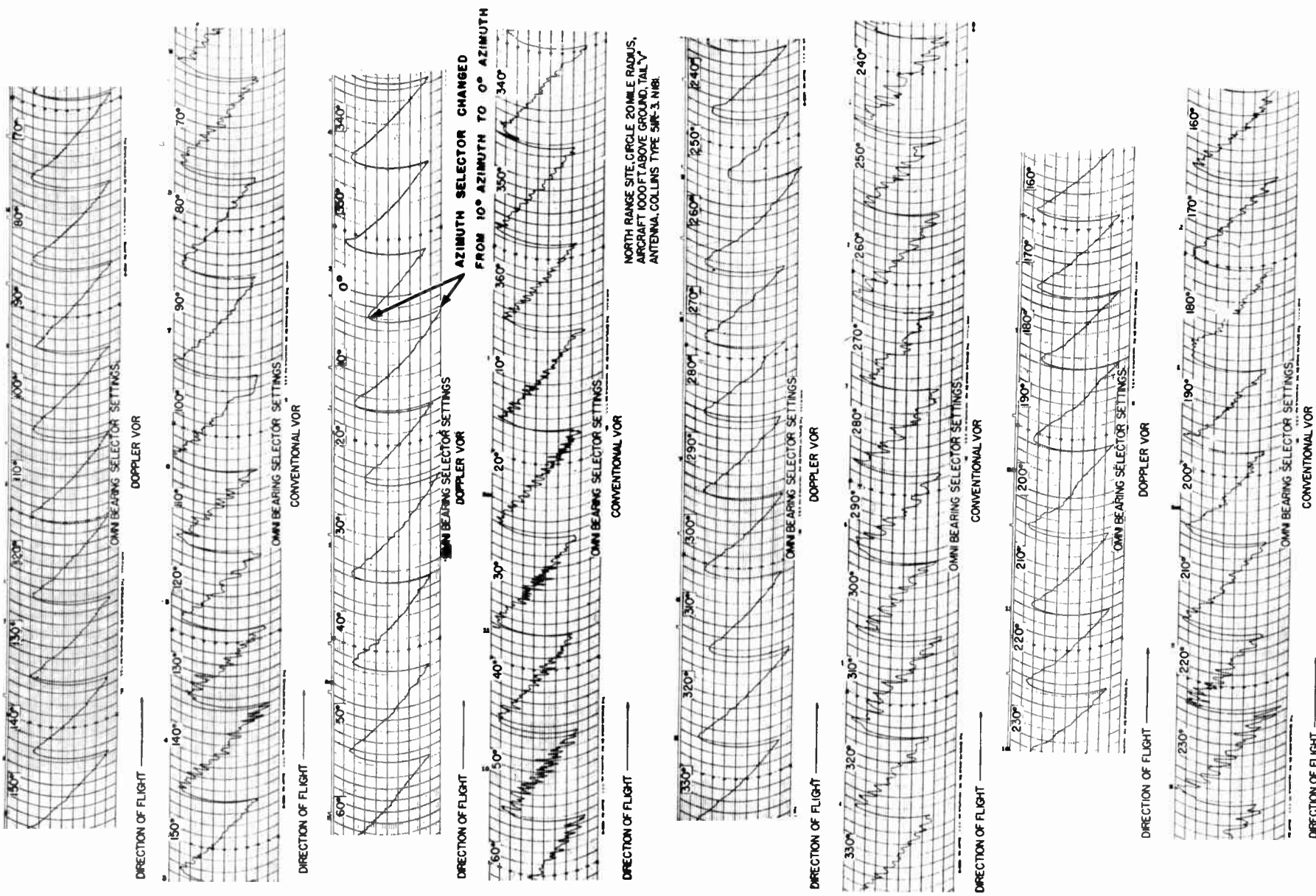


Fig. 18—Orbital flight recordings of the Doppler VOR and the conventional VOR at the Technical Development Center, Indianapolis, Ind.

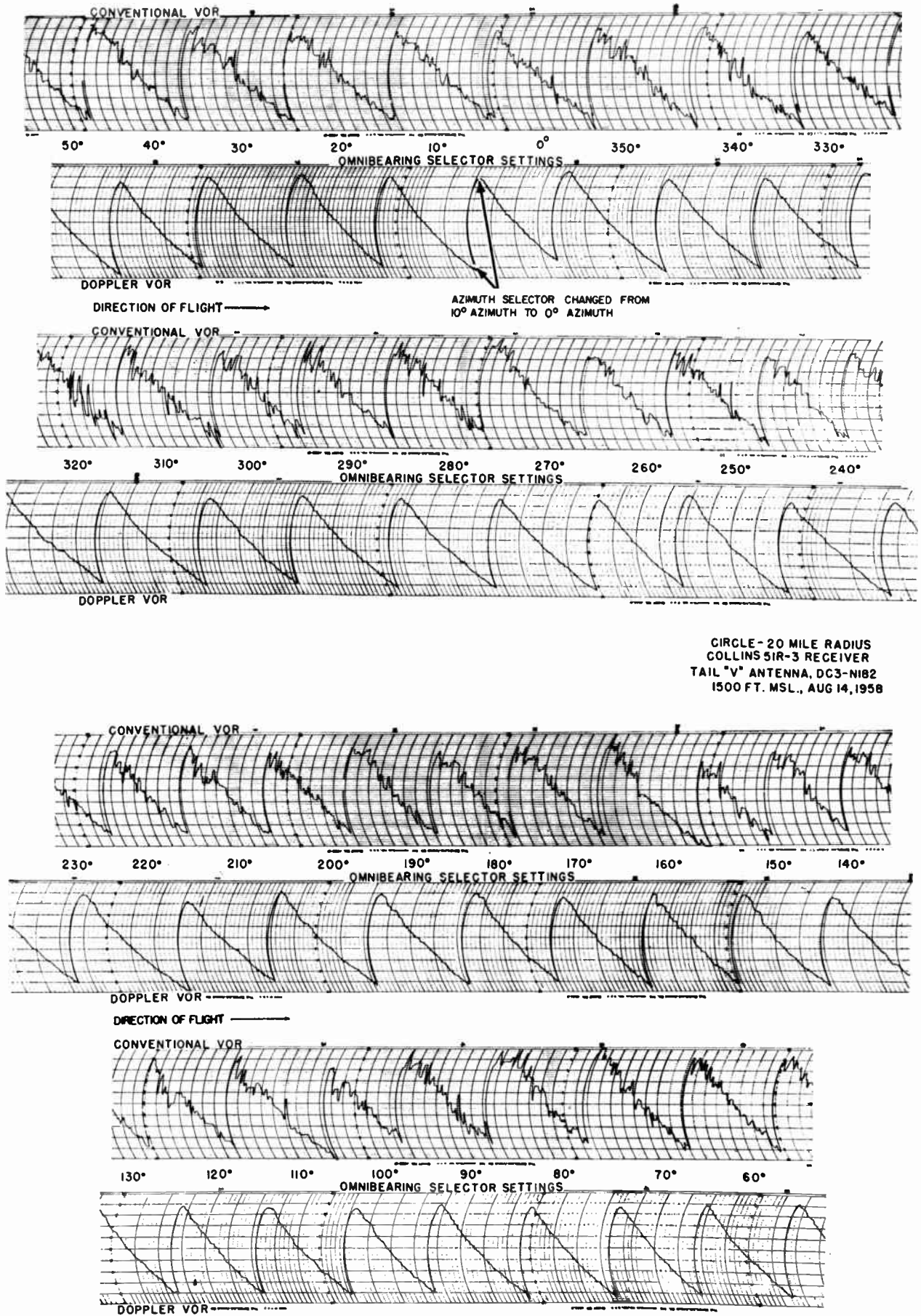


Fig. 19—Orbital flight recordings of the Doppler VOR and the conventional VOR at Charleston, S. C.

muth information for aircraft. Furthermore, it has been demonstrated that there is a distinct advantage in the Doppler principle over the pattern directivity principle of the conventional omnirange. Specifically, the Doppler principle has the characteristic that the antenna aperture does not introduce ambiguities into the bearing data. On the other hand the conventional omnirange can be made to have large aperture, but only by introducing ambiguities into the bearing information.

The following is a list of results pertaining to the development of the CAA Doppler omnirange:

- 1) The Doppler omnirange principle, when applied to an antenna array with an aperture of approximately 5 wavelengths and consisting of 50 loop antennas fed by a capacity type distributor, is feasible.
- 2) The Doppler omnirange markedly reduces the siting error compared with the conventional omnirange.

- 3) The Doppler omnirange is compatible with the conventional omnirange.
- 4) The maximum bearing errors of the two omnirange systems are comparable and approximately equal to $\pm 1.0^\circ$.
- 5) The cone characteristics of the Doppler omnirange are satisfactory. Reduced course sensitivity was measured over the station.
- 6) The Doppler omnirange polarization error is small compared to the conventional omnirange polarization error.

ACKNOWLEDGMENT

The authors wish to acknowledge the valuable aid of the following personnel in the Transmitter and Antenna Branch: William E. Adkins, John E. Bunk, Oliver J. DeZoute, Dennis E. Mason, Samuel E. Taggart, John Turk, Michael Wakovsky, Virgil E. Willey, William L. Wright.

Noise-Modulated Distance Measuring Systems*

B. M. HORTON†

Summary—Distance-measuring systems using random noise as the modulating function are described. The distance measurement is accomplished by correlating the modulation on the transmitted and received signals. The spectrum of the modulating function determines the way in which this correlation, and hence system output, depends on distance to a reflecting target. Physical realizability of filters limits the output-to-distance behavior of linear, noise-modulated systems. Theoretically, either amplitude or frequency modulation can be used, but the latter has distinct advantages in overcoming incidental spurious signals generated within the system. Actual multiplication of signals is avoided through use of a conventional mixer. The resulting system is similar to existing altimeters but is free of the ambiguities inherent in periodically modulated systems, avoids the "fixed error," and is capable of measuring distances down to a few feet. This makes it particularly suited for use as an altimeter in blind landing systems.

INTRODUCTION

THE purpose of this paper is to discuss the general characteristics of distance-measuring, or ranging systems in which the transmitted energy is modulated with random noise. One reason for interest in this type of system is that it appears to be promising for use as an altimeter for blind landing. Discussion will primarily concern CW systems.

Distance-measuring systems which depend upon reflection of energy from a target usually are based on a predictable detailed relationship between the transmitted and returning signals. Pulse radar systems measure time delay,¹ and FM systems measure phase shift or frequency difference²⁻⁹ to determine distance to the target. Those systems which are modulated with a periodic function are subject to ambiguities for targets whose time delay is greater than the repetition period. "Jitter" of the repetition period is often used to reduce

¹ L. N. Ridenour, ed., "Radar system engineering," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 1; 1947.

² S. Matsuo, "A direct reading radio-wave-reflecting-type absolute altimeter for aeronautics," Proc. IRE, vol. 26, pp. 848-858; July, 1938.

³ "FM radar altimeter," *Electronics*, pp. 130-134; April, 1946.

⁴ D. G. C. Luck, "Frequency Modulated Radar," McGraw-Hill Book Co., Inc., New York, N. Y.; 1949.

⁵ F. Penin and G. Phelizon, "Le mesureur de distance DME," *Onde Elect.*, vol. 33, pp. 309-318; January, 1953.

⁶ A. Black, K. E. Buecks, and A. H. Heaton, "Improved radio altimeter," *Wireless World*, vol. 60, pp. 138-140; March, 1954.

⁷ G. Collette and R. Labrousse, "Un altimetre radioelectrique a modulation de frequence," *Ann. Radioelec.*, vol. 10, pp. 387-398; January, 1955.

⁸ H. Gutton, H. Familier, and B. Ginger, "Etude de la modulation de frequence appliquee a la mesure des distances," *Ann. Radioelec.*, vol. 11, pp. 107-117; April, 1956.

⁹ M. A. W. Ismail, "A Study of the Double Modulated FM Radar," Verlag Leeman, Zurich, Switz., pp. 22-23; 1955.

* Original manuscript received by the IRE, January 15, 1959.
† Diamond Ordnance Fuze Labs., Washington, D. C.

the effect of this ambiguous response. When used in this way, the random variations are a perturbation to the orderly (and usually periodic) modulation of the transmitter.

In addition to the ambiguities arising from the repetition period, there is, in a periodic FM altimeter, a quantizing effect, sometimes called the "fixed error," which causes discrete changes in system output with changes in target distance. Although this quantizing effect can be avoided,^{10,11} a system which is inherently free of such effects would be desirable.

Elimination of the inherent ambiguities and the quantizing effect in a distance-measuring system implies that the modulation is not periodic and hence the spectrum of the modulating signal is not a harmonic series. One way of insuring that the modulation is not periodic is to use random noise as the modulating function. It is the purpose of this paper to discuss the general characteristics of noise-modulated systems, to point out some of the novel features which can be achieved, to discuss some of their limitations, to describe several types of noise-modulated systems, and to report the successful operation of one type of system. The method of analysis gives opportunity, within the limits of realizability and practicality, to design a system with a particular desired output-vs-distance relation.

Recent investigations¹²⁻¹⁴ of the power spectrum of a carrier which is modulated by random noise give pertinent information on the spectral characteristics of the signals in the systems to be described here. Although those results were not available at the time that the work being reported here was done, they are very useful in determining the relation between system parameters and the signal spectrum. Use of noise itself as the carrier has also been investigated,¹⁵ but the results do not apply to the present system.

APPLICATION OF THE CORRELATION FUNCTION TO DISTANCE MEASUREMENT¹⁶

Suppose a system radiates energy which is modulated in amplitude, phase, or frequency by a "random" func-

tion, *e.g.*, Gaussian noise. If a target very close to the transmitter reflects back some of the energy with a very short time delay, the transmitter will not have had sufficient time to make a very great change in amplitude (or phase, or frequency) during this short time. For a target at a greater distance, the time delay of the return signal will be greater, and there will be an increased probability of a large change in the signal which is being transmitted. If we now make a comparison between the outgoing and incoming signals, we have a result whose statistical properties depend upon distance. At zero distance (no time delay) the transmitted and return signals are identical, and we find perfect *correlation* between them. At very large distances, the return signal will seem to be almost completely unrelated to what is being transmitted, *i.e.*, the transmitted and return signals are *uncorrelated*.

The way in which correlation between the transmitted and return signals changes from perfect correlation at zero distance to lack of any correlation at very great distances depends upon the frequency components in the random function, *i.e.*, the spectrum of the noise. If only very-low-frequency components are present, then the transmitted signal cannot change rapidly in time, and the correlation between the outgoing and return signals changes very slowly with increasing distance. If, however, very-high-frequency components are present, the transmitter can make rapid changes in amplitude, phase, or frequency, and the correlation between the two signals goes rapidly toward zero as the time delay between transmitted and return signals increases.

It should be noted that if the modulating function is truly a random function, there is some probability that the outgoing and returning signals can be very different even for a very close target, but such large differences are not as likely as smaller ones. In the system being proposed here, only the average relationship between the outgoing and returning signals is used to determine distance. The accuracy with which the distance determination is made depends upon the length of time during which the relationship is measured as well as upon the accuracy of system parameters.

Correlation of a function with itself is called its "correlation function" or "autocorrelation" in communication theory, and "serial correlation" in the field of statistics. In an actual system, one of the signals can be modified by noise, Doppler effect, and other forms of interference, hence the correlation between the outgoing and return signals is not precisely its "autocorrelation" or "correlation function." However, since the system is based on the principle implied by these terms, they are appropriate, and departures from the principle will be treated as error signals.

The correlation-function principle has the possibility of being used in systems which employ radio waves, light, sound, or other effects as carrier. It is only necessary that the random function be impressed upon the

¹⁰ H. P. Kalmus, J. C. Cacheris, and H. A. Dropkin, "Nonquantized frequency-modulated altimeter," IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS, vol. ANE-1, pp. 15-21; June, 1954.

¹¹ M. A. W. Ismail, "A precise new system of FM radar," PROC. IRE, pp. 1140-1145; September, 1956.

¹² J. L. Stewart, "The power spectrum of a carrier frequency modulated by Gaussian noise," PROC. IRE, vol. 42, pp. 1539-1542; October, 1954.

¹³ R. G. Medhurst, "The power spectrum of a carrier frequency modulated by Gaussian noise," PROC. IRE, vol. 43, pp. 752-753; June, 1955.

¹⁴ P. R. Karr has shown in unpublished reports the relation between the spectrum of a carrier which is frequency modulated with random noise and the spectrum of the modulating function. For a high modulation index, the resulting spectrum asymptotically approaches the probability density of instantaneous resulting frequency.

¹⁵ A. A. Kharkevich, "The transmission of signals by modulated noise," in "Telecommunications," Pergamon Press, Inc., London, Eng., pp. 43-47; 1957.

¹⁶ Material in this section was presented in Natl. Bur. Stand. Classified Rep. O.E.D. 13.4-205R, issued May 23, 1952, and reissued in a slightly shorter version as Natl. Bur. Stand. Classified Rep. 17-104, November 7, 1952.

outgoing energy and detected in that which is reflected back from the target.

In some applications it will be advantageous actually to measure, not the correlation function itself, but a closely related function which might be called the "anti-correlation" function. This will be discussed after some features of the correlation function principle have been presented.

The Correlation Function

Let $F(t)$ be a dimensionless continuous random-noise function with a Gaussian distribution of amplitudes and no dc component. The correlation function $\psi(\tau)$ of $F(t)$ is defined by ^{17,18}

$$\psi(\tau) = \left(\frac{1}{K}\right) \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T F(t)F(t - \tau)dt \quad (1)$$

where the normalizing factor K is given by

$$K = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [F(t)]^2 dt. \quad (2)$$

The amplitude spectrum $S(f)$ of a function $G(t)$ which is equal to $F(t)$ in the interval $0 \leq t \leq T$ and zero outside the interval is:

$$S(f) = \int_0^T F(t)e^{-i2\pi f t} dt, \quad (3)$$

and the corresponding normalized "power" spectrum $w(f)$ is

$$w(f) = \left(\frac{1}{K}\right) \lim_{T \rightarrow \infty} \frac{2|S(f)|^2}{T}, \quad (4)$$

considering only positive frequencies. The correlation function is related to the power spectrum by the Wiener-Khinchine relation

$$\psi(\tau) = \int_0^\infty w(f) \cos 2\pi f \tau df. \quad (5)$$

and

$$w(f) = 4 \int_0^\infty \psi(\tau) \cos 2\pi f \tau d\tau. \quad (6)$$

We see that a system which performs the operation of correlating $F(t)$ with $F(t-\tau)$, as indicated by (1), will behave in accordance with (5), *i.e.*, its distance dependence will be governed by the power spectrum $w(f)$ of the random noise. It can be seen from (6) that some desired system behavior $\psi(\tau)$ can be postulated, and the corresponding noise power spectrum determined. Many types of correlation function, however, lead to unrealizable, or impractical power spectra.

¹⁷ M. C. Wang and G. E. Uhlenbeck, "On the theory of the Brownian motion II," *Rev. Mod. Phys.*, vol. 17, pp. 323-342; April-July, 1945.

¹⁸ S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 23, pp. 282-332; July, 1944. For convenience in later use of this function, we have used $f(t-\tau)$ in place of Rice's $f(t+\tau)$, but the difference vanishes as $T \rightarrow \infty$ when τ is finite.

We now consider several special cases of noise spectra and give the corresponding correlation functions. In each case, we use a normalized power spectrum *i.e.*, a spectrum for which

$$\int_0^\infty w(f)df = 1, \quad (7)$$

and a normalized $\psi(\tau)$ in order that $\psi(0) = 1$.

Case I—Ideal Band-Pass Filter: Let

$$\begin{aligned} w_1(f) &= \frac{1}{f_2 - f_1} \text{ for } f_1 \leq f \leq f_2 \\ &= 0 \text{ for } f < f_1 \\ &= 0 \text{ for } f > f_2. \end{aligned} \quad (8)$$

Using (5) and integrating, we obtain

$$\psi_1(\tau) = \frac{1}{(f_2 - f_1)2\pi\tau} [\sin 2\pi f_2 \tau - \sin 2\pi f_1 \tau]. \quad (9)$$

When $f_1 = 0$, the low-pass filter case,

$$\psi_1(\tau) = \frac{1}{2\pi f_2 \tau} \sin 2\pi f_2 \tau. \quad (10)$$

Both of these functions oscillate, and thus have the disadvantage that a certain value of $\psi(\tau)$ may correspond to more than one value of τ . When f_1 and f_2 are almost equal, $\psi_1(\tau)$ looks like a very slightly damped sine wave. As the bandwidth is increased, the "damping" increases. When $f_1 = 0$, $\psi_1(\tau)$ appears highly damped.

Case II—Exponential Filter: Now suppose

$$w_2(f) = ae^{-af} \quad (11)$$

where a is a constant.

Again using (5) and integrating, we obtain

$$\psi_2(\tau) = \frac{1}{1 + \left(\frac{2\pi\tau}{a}\right)^2}. \quad (12)$$

Here we have a correlation function which decreases monotonically with increasing τ for $\tau > 0$, hence ambiguity in the relation between τ and $\psi_2(\tau)$ is avoided. It would be impossible, however, to obtain a noise spectrum conforming exactly to (11) and would be difficult to approximate.

Case III—Low-Pass RC Filter: In this case we assume that the noise spectrum is that which would be produced if "white" noise from a generator of zero output impedance is passed through the circuit of Fig. 1. The output power spectrum is

$$w_3(f) = \frac{4/\lambda}{(2\pi f)^2 + (1/\lambda)^2}. \quad (13)$$

where $\lambda = R_1 R_2 C / (R_1 + R_2)$ is the time constant of the noise-shaping network. If we substitute (13) into (5) and integrate, we obtain

$$\psi_3(\tau) = e^{-\tau/\lambda}. \quad (14)$$

This correlation function, the inverse of Case II, is a monotonically decreasing function of delay time τ , hence avoiding ambiguities in the relationship between $\psi(\tau)$ and τ . It also has the advantage that the noise spectrum can be shaped with a simple circuit.

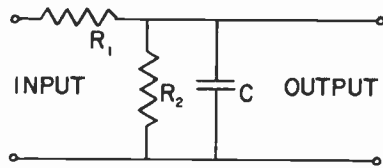


Fig. 1—Noise filter for Case III.

Case IV—Gaussian Filter With Maximum at Zero Frequency:

$$w_4(f) = \frac{1}{f_c} \sqrt{\frac{2}{\pi}} e^{-f^2/2f_c^2} \quad (15)$$

i.e., the power spectrum follows a one-sided normal distribution with maximum density at zero. This leads to

$$\psi_4(\tau) = e^{-2\pi^2 f_c^2 \tau^2} \quad (16)$$

which also follows the normal distribution with maximum at zero. Such a correlation function may be especially adapted to some applications where, for example, greater range sensitivity is desired at a particular range. The required power spectrum $w_4(f)$ would, as in Case II, however, be impossible to obtain exactly.

Case V—Gaussian Filter With Maximum at \bar{f} : In this case we assume that the noise power follows a normal distribution with mean value at \bar{f} , and has a standard deviation (in frequency) of σ , *i.e.*,

$$w_5(f) = \frac{1}{\sigma\sqrt{\pi}} e^{-(f-\bar{f})^2/2\sigma^2} \quad (17)$$

If we put (17) in (5) and integrate, we obtain an approximate expression for the corresponding correlation function $\psi_5(\tau)$, which is valid when $\bar{f} \gg \sigma$,

$$\psi_5(\tau) = [\cos 2\pi\bar{f}\tau] e^{-2\pi^2\tau^2\sigma^2} \quad (18)$$

an oscillating function with a Gaussian envelope.

The required power spectrum of the noise may be obtained approximately with tuned circuits, but $\psi_5(\tau)$ has the disadvantage that it oscillates and hence has ambiguities in the relationship between $\psi(\tau)$ and τ .

Of the several cases of noise spectrum which have been presented, three have led to correlation functions which give no ambiguities, but two of these may require a complex noise band-shaping network while the other requires only a resistance-capacitance network. Its correlation function $\psi_3(\tau)$, (14) is

$$\psi_3(\tau) = e^{-\tau/\lambda} \quad (14a)$$

If one wants $\psi_3(\tau)$ to have the value $1/e$ (*i.e.*, $\tau = \lambda$) for a distance of 100 feet, then

$$\lambda = \frac{2 \times 100}{C} \approx 2 \times 10^{-7}$$

seconds, approximately. The corresponding half-power point frequency for the RC network is about 0.8 mc. This case has the property that a certain percentage change in range gives a constant percentage change in $\psi_3(\tau)$ for all ranges. This is shown by the fact that

$$\frac{d}{d(\tau/\lambda)} \psi_3(\tau) = \frac{-e^{-\tau/\lambda}}{e^{-\tau/\lambda}} = -1. \quad (19)$$

A disadvantage of the correlation function is that it has its maximum value when $\tau = 0$. This means that a small amount of leakage from the transmitting antenna to the receiving antenna might obscure a return signal of greater strength but with a greater time delay. Similarly, a small return from a near target may obscure a large target farther out. This is discussed more fully in later sections. This property of the correlation function led to the search for a function which is complementary to the correlation function, *i.e.*, one which starts at zero for zero distance and increases monotonically to a maximum value at infinite distance.

“Anticorrelation” Distance Measurement

In the correlation function system, a function was multiplied by its delayed version, and the average value of this product was obtained. [See (1).] The average value of this product is a maximum when the delay is zero. In the “anticorrelation” system to be discussed now, we find the mean square value of the difference between the value of a function and its delayed version. When the time delay τ is zero, the instantaneous difference is zero, hence the mean-square difference is zero, and the effect of leakage from transmitter to receiver is lessened.

The “anticorrelation” function $H(\tau)$ will be defined by

$$H(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_0^T [F(t) - F(t - \tau)]^2 dt \quad (20)$$

$$= \lim_{T \rightarrow \infty} \left[\frac{1}{2T} \int_0^T [F(t)]^2 dt + \frac{1}{2T} \int_0^T [F(t - \tau)]^2 dt - \frac{1}{T} \int_0^T F(t)F(t - \tau) dt \right].$$

If $F(t)$ is normalized, *i.e.*, in (2), $K = 1$, then

$$H(\tau) = 1 - \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T F(t)F(t - \tau) dt \quad (21)$$

or

$$H(\tau) = 1 - \psi(\tau). \quad (22)$$

Thus we see that $H(\tau)$, defined as half the mean-square difference between normalized $F(t)$ and $F(t - \tau)$, is the complement of their mean product $\psi(\tau)$.

Using (22) and the correlation functions obtained in Cases I-V in the previous section, we have immediately, from (22), $H(\tau)$ for the noise spectra considered there.

For a noise band which is shaped by the RC network, (Case III)

$$H_3(\tau) = 1 - e^{-\tau/\lambda}. \quad (23)$$

If a response of

$$H_3(\tau) = 1 - \frac{1}{e}$$

is desired for a target at 100 feet, λ must be approximately 2×10^{-7} seconds, which again corresponds to a half-power frequency of about 0.8 mc.

At very small values of τ/λ , $H_3(\tau)$ is approximately proportional to τ/λ ; thus the output of this system is, for near targets, nearly a linear function of distance.

The fractional range sensitivity is:

$$\frac{\frac{d}{d(\tau/\lambda)} H_3(\tau)}{H_3(\tau)} = \frac{e^{-\tau/\lambda}}{1 - e^{-\tau/\lambda}} = \frac{1}{e^{\tau/\lambda} - 1}. \quad (24)$$

When $\tau=0$ this is unlimited, and as $\tau \rightarrow \infty$, it approaches the value zero. Thus the system has its highest sensitivity at the shortest distances.

When $\tau/\lambda=1$, the fractional range sensitivity has the value 0.58.

PRACTICALITY OF VARIOUS SYSTEMS

The foregoing discussion shows that a system which impresses noise modulation on a carrier, and then correlates it with the modulation reflected from a target, is able to determine distance to the target. It has been shown that measurement of the mean-square difference between the modulation on the transmitted energy and the modulation on the received energy gives a system output which is the complement of the correlation function. Any type of system capable of performing the above operations can be used to measure distance.

Modulating a particular characteristic of the transmitted energy always involves some incidental modulation of other characteristics, and this leads to practical difficulties in achieving good system sensitivity. Some of the transmitted energy invariably arrives at the receiver input terminals with a very short time delay and tends to obscure small signals. Target signal levels vary over a very wide range, may obscure one another, may cause abnormal operation of parts of the system, or may require balancing or normalizing operations which are not practical to achieve.

Leakage from the transmitter to the receiver is conveniently avoided by operating the receiver at either different times and/or different frequencies from those being used by the transmitter. Pulse systems accomplish this in time by turning off the transmitter during signal receiving times. Conventional FM altimeters accomplish this by making the receiver very insensitive to frequencies in the modulation waveform, and thus they are insensitive to signals having a small time delay. Practi-

cal noise-modulated systems must deal with these same problems.

AM Systems

An AM correlation-function system would give an output which would depend on signal strength as well as target distance, unless complete normalizing of the return signal were accomplished. This implies an essentially perfect automatic gain control in the receiver, if system output is to be dependent upon distance only. Thus such a system is not very practical.

In an AM correlation-function system, the response $\psi(\tau_a, \tau_b)$ to two signals of amplitude a and b , having delay times τ_a and τ_b respectively, would be:

$$\psi(\tau_a, \tau_b) = \frac{\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T F(t) [aF(t - \tau_a) + bF(t - \tau_b)] dt}{\sqrt{K} [aF(t - \tau_a) + bF(t - \tau_b)]_{\text{rms}}} \quad (25)$$

where the terms in the denominator of (25) serve to normalize $\psi(\tau_a, \tau_b)$. We can obtain from (25) and from the definition of (1), that

$$\psi(\tau_a, \tau_b) = \sqrt{K} \frac{a\psi(\tau_a) + b\psi(\tau_b)}{[aF(t - \tau_a) + bF(t - \tau_b)]_{\text{rms}}}, \quad (26)$$

which is a weighted mean of $\psi(\tau_a)$ and $\psi(\tau_b)$ where a and b are the weighting factors. The contributions of signals a and b to $\psi(\tau_a, \tau_b)$ have the ratio,

$$R_{ab} = \frac{a\psi(\tau_a)}{b\psi(\tau_b)}. \quad (27)$$

This result shows how a weak signal of strength a can obscure a larger signal of strength b , if $\psi(\tau_a)$ is much larger than $\psi(\tau_b)$. In a radiating system where the strength of signal declines rapidly with distance, this may be a very objectionable property.

An AM anticorrelation system must obtain the difference between two normalized functions, otherwise system output will again depend upon signal strength as well as target distance. Normalizing the received signal is so difficult that such a system is not very practical.

Interference between two signals in an AM anticorrelation system is somewhat involved, and an analysis will not be given here because it is of no immediate practical interest.

FM Systems

An FM correlation type system has the advantage that a comparison of the frequencies of the transmitted and returning signals can be made essentially independent of the signal level. The "capturing effect" of an FM receiver which greatly suppresses receiver output for all but the largest signal,¹⁹ can either be beneficial or detrimental, depending on relative signal levels.

¹⁹ J. Granlund, "Interference in Frequency Modulation Reception," Ph.D. dissertation, Mass. Inst. Tech., Cambridge, Mass.; 1950.

An FM correlation-function system must convert the frequency modulation on the returning signal to some convenient form, say voltage, in order to multiply it by the modulation being impressed on the transmitted signal. Such a conversion would involve either 1) a discriminator operating at the carrier frequency or 2) a mixer and local oscillator to heterodyne the signal to a lower frequency and a discriminator at the lower frequency. Such a system could be made, but like the AM systems, would be very sensitive to leakage from transmitter to receiver and would be substantially more complex than a conventional FM altimeter.

Simple Noise-Modulated System

A simpler system using noise modulation is an FM anticorrelation system. A sample of the transmitted signal is used as the local oscillator in a conventional mixer, as is done in conventional FM altimeters. The mixer output has an instantaneous frequency equal to the magnitude of the instantaneous frequency difference of the transmitted and received signals. Thus, direct conversion to a conveniently low frequency is accomplished without a separate local oscillator. The type of receiver best suited to handling the difference-frequency signal depends upon the application. If a wide-band limiting receiver is used, it will preserve the "zero crossings" of the mixer output. When the modulation index of the mixer output is high, the zero crossings give a good measure of the average magnitude of the instantaneous difference frequency.²⁰ This high modulation index corresponds to large phase excursions of the difference-frequency phasor between changes in sign of the instantaneous difference frequency. Large phase excursions occur when the transmitter is swept monotonically over a wide frequency range and/or when the time delay of the received signal is large. If the rate of zero crossing is converted to a fluctuating voltage, squared, and averaged, the system output will behave, to the extent that zero crossing rate is a good measure of instantaneous frequency, as an anticorrelation system.

We now take note of the fact that the random noise function $F(t)$ assumed has a Gaussian probability distribution of amplitudes. If the frequency-deviation characteristic of the transmitter is linear, the transmitted signal will have a Gaussian probability density with mean value at the carrier frequency. The returning signal will, for a noise-free, stationary target likewise have a Gaussian probability with the same mean value. Since the difference frequency,

$$\Delta\omega = \omega(t) - \omega(t - \tau) \quad (28)$$

is the difference of frequencies having a Gaussian distribution, it too will have a Gaussian probability distribution, but with zero mean. The mixer does not, however, preserve the algebraic sign of the difference frequency,

so the mixer output is a signal whose probability distribution of instantaneous frequency is a one-sided Gaussian distribution, with its maximum at an instantaneous frequency of zero.

Since the instantaneous difference frequency follows a Gaussian distribution, the average squared difference frequency $\overline{\Delta\omega^2}$ is related to the average magnitude of the difference frequency $|\overline{\Delta\omega}|$ by

$$|\overline{\Delta\omega}| = \sqrt{\frac{2}{\pi} \overline{(\Delta\omega^2)}} \quad (29)$$

Because of this relationship, which holds as long as $\Delta\omega$ has a Gaussian probability distribution, the FM anticorrelation system can, by measuring the average magnitude of the difference frequency, determine through (29) the mean-square frequency difference which is in turn related to the correlation function and the noise power spectrum. This noise-modulated system is no more complex than a conventional altimeter.

The precise relationship between mean magnitude of the difference frequency and the noise power spectrum will now be obtained. Suppose the transmitter is frequency modulated so that the instantaneous frequency deviation of the transmitter is $\omega(t)$. Let

$$F(t) = \frac{\omega(t)}{\omega(t)_{\text{rms}}} \quad (30)$$

where

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \omega(t) dt = 0, \quad \text{and} \quad \omega(t)_{\text{rms}} = \{\overline{\omega(t)^2}\}^{1/2}; \quad (31)$$

then

$$H(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_0^T \left[\frac{\omega(t)}{\omega(t)_{\text{rms}}} - \frac{\omega(t - \tau)}{\omega(t)_{\text{rms}}} \right]^2 dt \quad (32)$$

$$= \frac{1}{[\omega(t)_{\text{rms}}]^2} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_0^T [\omega(t) - \omega(t - \tau)]^2 dt \quad (33)$$

$$= \frac{1}{[\omega(t)_{\text{rms}}]^2} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_0^T [\Delta\omega]^2 dt. \quad (34)$$

Using (29)

$$H(\tau) = \frac{\pi \{ |\overline{\Delta\omega}| \}^2}{4\omega(t)_{\text{rms}}^2}. \quad (35)$$

But

$$\begin{aligned} H(\tau) &= 1 - \psi(\tau) \\ &= 1 - \int_0^\infty w(f) \cos 2\pi f \tau df. \end{aligned}$$

Hence

$$\{ |\overline{\Delta\omega}| \}^2 = \frac{4}{\pi} [\omega(t)_{\text{rms}}]^2 \left\{ 1 - \int_0^\infty w(f) \cos 2\pi f \tau df \right\}, \quad (36)$$

²⁰ P. R. Karr, "A Note on Zero Crossings and Average Frequency," Diamond Ordnance Fuze Labs., Washington, D. C., Tech. Rep. No. 133; October 25, 1954.

or

$$|\Delta\omega| = \frac{2}{\sqrt{\pi}} \omega(t)_{\text{rms}} \sqrt{1 - \int_0^{\infty} w(f) \cos 2\pi f \tau df}, \quad (37)$$

which shows the way in which this system depends upon the power spectrum $w(f)$ of the modulating noise.

If the RC filter of Case III is used to shape the noise spectrum, the mean magnitude of the difference frequency is

$$|\Delta\omega| = \frac{2}{\sqrt{\pi}} \omega(t)_{\text{rms}} \sqrt{1 - e^{-\tau/\lambda}}. \quad (38)$$

The normalized mean magnitude m_1 of the difference frequency for this case is

$$m_1 = \frac{|\Delta\omega|}{\omega(t)_{\text{rms}}} = \frac{2}{\sqrt{\pi}} \sqrt{1 - e^{-\tau/\lambda}}. \quad (39)$$

This theoretical behavior is plotted in Fig. 2. The system output changes most rapidly at short time delays.

A similar plot for the Gaussian filter of Case IV is also given in Fig. 2. In this case the normalized mean magnitude m_2 of the difference frequency is

$$m_2 = \frac{2}{\sqrt{\pi}} \sqrt{1 - e^{-2\pi^2 f_c^2 \tau^2}}. \quad (40)$$

Test of Simple System

A system, shown in the block diagram of Fig. 3, was built in accordance with the above description of the "Simple Noise-Modulated System." The results of tests with this system are shown in Fig. 4. The noise spectrum used was that provided by a commercial noise generator having a spectrum intermediate between the Cases III and IV. The receiver was of the broad-band limiting type with suppression of the low frequency response to avoid effects of the leakage signal. Only the zero crossings of the difference frequency signal (excluding the very low frequency portion of the spectrum) were preserved. A linear discriminator provided an average output voltage proportional to the average frequency of these zero crossings. With very short time delay or very low frequency deviations, the phase excursions of the difference frequency phases were insufficient to "capture" the receiver from its own noise, and this placed a lower limit on the system output. At the largest frequency deviations and longest time delays, the discriminator was driven out of its linear region part of the time, and this is indicated by the saturation effect in system output. It can be seen from the figure that the general behavior of system output does conform to expectations.

LIMITATIONS, ERRORS, MODIFICATIONS

Limitations

Since the Fourier transform exists for a wide class of functions, one might be tempted to expect that a system might be designed to give almost any desired out-

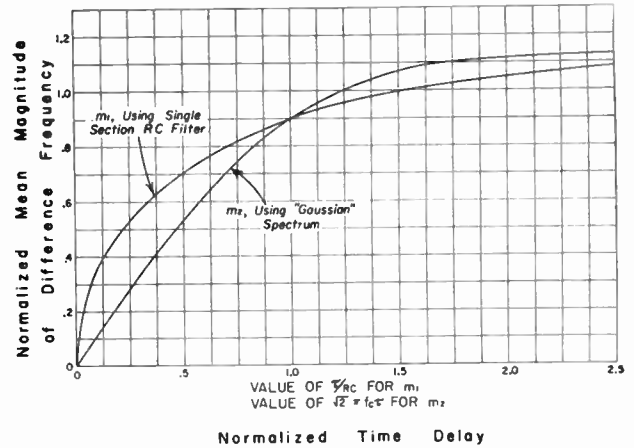


Fig. 2—Theoretical output for simple system.

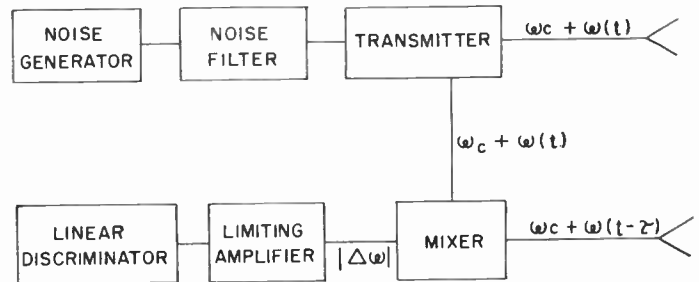


Fig. 3—Block diagram of simple noise-modulated system.

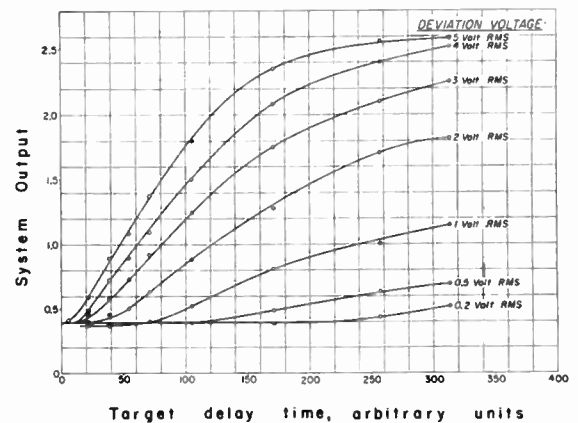


Fig. 4—Performance of simple noise-modulated system.

put with distance. Suppose, for example, that a correlation function which is a step function is desired. The corresponding power spectrum is, according to (6), the Fourier transform of this step function and hence would involve *negative* power in some spectral regions. This, of course, is not even theoretically achievable.

Errors

Certain errors are inherent in a noise modulated system because of the short term variability of the random noise and the fact that unlimited averaging time is not available. A complete analysis of this limitation has not been made, but in the case of an FM system which counts zero crossings of the difference frequency signal,

an upper limit can be set. The individual zero crossings can be no more random than if they were completely independent, and in that case, the standard deviation of samples of zero crossings for which the average number is \bar{n} would have the approximate value $\sqrt{\bar{n}}$ and a *relative* standard deviation of $\sqrt{\bar{n}}/\bar{n} = 1/\sqrt{\bar{n}}$. Thus, if \bar{n} were 1,000, the rms fluctuation would be approximately thirty-two and the root-mean-square relative fluctuation would be about 3 per cent.

If, in an FM system, the return signal is shifted by a fixed amount, as would occur by the Doppler effect, the resulting probability distribution of the magnitude of the instantaneous difference frequency will be that of a Gaussian distribution which is folded (by the mixer action) at a point other than its mean value. If the Doppler shift is small compared to the mean magnitude of the difference frequency brought about by the modulation, $|\Delta\omega(t)|$, the effect on system output would be very small, since the fixed frequency shift would alternately add and subtract from $|\Delta\omega(t)|$.

Modifications

Numerous variations to the basic system are possible, and no attempt will be made to list them all, but several have interesting possibilities and should probably be mentioned.

A system which provides for an internally delayed version of the transmitted signal (or of the modulation on it) could use this delayed signal for comparing with the received signal. In a correlation system, this would give maximum correlation for a signal delay equal to the internal delay. An anticorrelation system would yield a null at the same point.

The saturation effect at high difference frequencies and the small phase excursions at small time delays in

the simple system suggest the desirability of maintaining an essentially constant difference-frequency spectrum by means of a servo system which would control the root-mean-square frequency deviations of the transmitter. The servo output could then be used to indicate range. This modification would make better use of receiver bandwidth and would tend to maintain constant relative statistical error.

SUMMARY

A new type of modulation function, random noise, has been introduced in distance-measuring systems. It makes possible a new class of system output-to-distance relations, and in particular, it is possible to have an average output which is an unambiguous function of distance. Systems having an output which is an arbitrary function of distance are not possible because of physical realizability and practicality considerations. System characteristics can easily be changed by switching filters for the modulating noise. Using simple filters, it is possible to make a system which has its highest sensitivity at the shortest ranges, and this makes such a system particularly suited for blind landing applications.

These systems do not depend upon the detailed relationship between transmitted and returning signals, hence some time averaging must be used for accuracy. Usually, however, the averaging time required is not large.

Many variations of the basic system are possible, for example a servo can be used to control some system parameter and thus maintain constant spectrum into the receiver. The simplest noise-modulated distance measuring system is frequency modulated, and is no more complex than a conventional FM altimeter.

VLF Propagation Measurements for the Radux-Omega Navigation System*

C. J. CASSELMAN†, D. P. HERITAGE†, AND M. L. TIBBALS‡

Summary—This paper describes special VLF propagation measurements in connection with a feasibility study of a long range navigation system. Round-trip single-frequency measurement of phase stability was made between Hawaii and San Diego on frequencies from 10.2 kc to 18.2 kc in 1 kc increments.

During January 15 to 23, 1958, the standard deviation of phase stability on 12.2 kc was 4 μ sec daytime and 5 μ sec night time.

One-way two-frequency transmissions were monitored in San Diego and Washington, D. C., to determine the phase stability of a 1 kc difference frequency for pairs of frequencies from 10.2 kc to 18.2 kc. Data analyzed at time of submission of this paper (10.2–16.2 kc) indicate limitations of the two-frequency system for lane identification (resolution of cyclic ambiguities corresponding to one period of the carrier frequency).

The techniques used to instrument these tests are considered somewhat unique. Data reported herein are general and applicable to any propagation study. The data being collected are leading to a better understanding of the mechanism of VLF propagation.

INTRODUCTION

SINCE the end of World War II, the range and accuracy requirements for a military navigation system for ships and aircraft have steadily increased. Specifically, a system is desired which will provide navigation day or night, under all weather conditions, to a range of approximately 5000 miles with a fix accuracy of 1 mile or better. Until recently, such a system was believed to be well beyond the "state of the art" for radio navigation techniques.

It has been only within the past few years that the stability of LF-VLF propagation over long paths was found to approach that required for such precise navigation. This was demonstrated by work on 40 kc Radux by the U. S. Navy Electronics Laboratory, San Diego, and Naval Research Laboratory, Washington, D. C., and by monitoring 16-kc transmissions from GBR Rugby, England, by Pierce.¹ The first demonstration of the phase stable nature of LF signals over long distances was performed at NEL in 1952. Measurements of Hawaiian transmissions on 40 kc at NEL indicated that the received carrier frequency was sufficiently phase stable day or night to provide a line-of-position (l-o-p) accuracy of less than one nautical mile. This was confirmed in the same year for transmissions from San Diego to Cruft Laboratory, Cambridge, Mass.

The predecessor of Radux-Omega, 40 kc Radux proposed by J. A. Pierce² in 1947, was a hyperbolic CW

phase comparison system with an expected range of 2500–3000 nautical miles and a fix accuracy of ± 5 miles. NEL and NRL began development of this system in 1950. An experimental system was demonstrated during 1954–1957. 40 kc Radux was shelved, however, because of 1) requirements for greater accuracy, and 2) excessive cost of a world-wide system since many stations (approximately 24) would be required.

Radux-Omega³ was proposed based on experience gained on the 40-kc system and is also a hyperbolic CW phase comparison system. It is intended to operate in the VLF spectrum below 20 kc, and a world-wide system is envisaged with as few as eight stations (predicated on 5000 nautical mile baselines). A network of eight stations would provide 15 probable sets of lines-of-position over the earth's surface, with at least two sets of l-o-p's being available at all times to the navigator for a fix.

A compromise of factors such as signal-to-atmospheric noise ratio, lane width (distance between equi-phase lines-of-position on a hyperbolic navigation chart), and propagation stability, suggest a frequency band of 10–15 kc for a long-range navigation system. At these frequencies, it is not possible to transmit sufficiently narrow pulses to allow positive separation of the various propagation modes, *i.e.*, ground wave, 1 hop, 2 hop, 3 hop, etc. The reason for this is that physically realizable antennas have extremely high Q resulting in antenna bandwidths of 20–50 cps. Therefore, it is necessary to utilize a CW phase measuring system whereby the signal received from each station is a composite of contributions from all the transmission modes. Even so, measurements show this composite signal to be surprisingly stable and should allow l-o-p measurements to better than 1 mile in the coverage area.

The stability and predictability of ionospheric effects on radio propagation are found to be limiting factors in providing a long range navigation service with the requirements previously stated. A rather comprehensive propagation study is necessary to establish the required parameters.

This paper is concerned with the instrumentation and propagation data obtained in the study of VLF phase stability as applied to the Radux-Omega system. Particular attention was directed toward the problem of resolving cyclic ambiguities which at 12 kc are only 6.7 miles apart. The data were collected from stations operating over the Hawaii-San Diego path (2270 nautical miles or 4200 km) and the Hawaii-Washington, D. C.

³ C. J. Casselman and M. L. Tibbals, "The Radux-Omega long range navigation system," *Proc. Second Natl. Conv. on Military Electronics*, pp. 385–389; June, 1958.

* Original manuscript received by the IRE, December 3, 1958; revised manuscript received, February 16, 1959.

† U. S. Navy Electronics Lab., San Diego, Calif.

¹ J. A. Pierce, "The diurnal carrier-phase variation of a 16-kilo-cycle transatlantic signal," *Proc. IRE*, vol. 43, pp. 584–588; May, 1955.

² J. A. Pierce, "Radux," Cruft Lab., Harvard University, Cambridge, Mass., Tech. Rep. No. 17; July, 1947.

path (4200 nautical miles or 7780 km).

In the past, theoretical analysis of the propagation mechanism⁴ at VLF has been compared with empirical amplitude data.⁵ In one paper,⁶ phase data obtained on a single frequency and over a fixed distance are discussed in connection with the mode theory.

It is believed that the instrumentation methods for phase measurements reported herein should provide a powerful tool for further investigation of the propagation mechanism.

THE RADUX-OMEGA SYSTEM

The reasons for performing the specific propagation measurements will be evident from the following brief description of the Radux-Omega system.

This navigation system can be considered as one of a family of hyperbolic systems, the best known of which are Loran (American) and Decca (British). The basic principle involves measurement of time-difference between signals received from two or more fixed stations. The signals are transmitted with a fixed time or phase relationship. The loci of constant phase or time-difference in Fig. 1 are represented by a family of hyperbolas whose foci correspond to the locations of the two stations.

As mentioned in the introduction, Radux-Omega is a CW phase comparison system. The time-difference is actually obtained by a measurement of the phase-difference between the two radio frequency carriers (common frequency).

In such a system, a practical problem arises because ambiguities (lines of identical phase) occur at intervals corresponding to half periods of the radio frequency carrier. For example, at 12 kc, ambiguities occur every 6.7 miles along the baseline (somewhat greater distances as coverage limits are approached because of divergence of hyperbolic lines). In the present Radux-Omega system, an attempt is made to solve this problem by transmitting two frequencies from each station (e.g. 12 kc and 13 kc). A coarse reading is obtained by measuring the phase-difference of the master and slave difference-frequency (1.0 kc in the example). This results in new ambiguities spaced 81 miles which present no serious problem to today's navigator.

Fig. 2 illustrates the principles by which the fine-phase and coarse-phase l-o-p's are furnished the navigator.

The master station transmits alternately on frequencies f_1 and f_2 . At the slave station, transmissions on f_1 and f_2 are individually phase-locked to signals of cor-

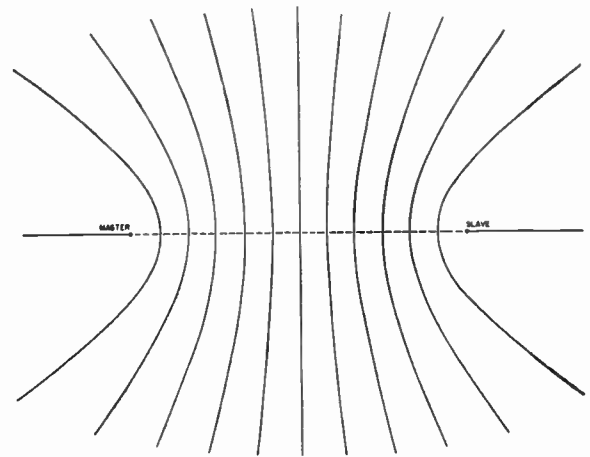


Fig. 1—Hyperbolic equi-phase lines.

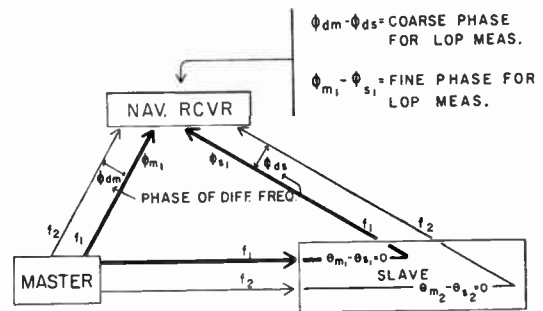


Fig. 2—Phase relationships for single l-o-p. Two station system.

responding frequency received from the master. The condition of zero phase relationship between received master and transmitted slave signals, is indicated by

$$\theta_{m_1} - \theta_{s_1} = 0$$

and

$$\theta_{m_2} - \theta_{s_2} = 0,$$

where subscripts m_1 and s_1 designate master and slave signals on frequency f_1 , and m_2 and s_2 master and slave signals on f_2 .

At the navigator's position, a discreet phase relationship between the f_1 and f_2 signals received from both stations is assured by the synchronism maintained at the slave station.

The navigator's equipment measures $\phi_{m_1} - \phi_{s_1}$ for determination of the fine-phase line of position, where ϕ , with specific subscripts defined above, represents the relative phase of each signal at the navigator's position. Discreet values of $\phi_{m_1} - \phi_{s_1}$ occur throughout the hyperbolic pattern at differential transmission time intervals of $1/f_1$, where

$$\frac{1}{f_1} = (t_m - t_s)$$

and

- t_m = transmission time of master signal
- t_s = transmission time of slave signal.

⁴ J. R. Wait, "The mode theory of VLF ionospheric propagation for finite ground conductivity," *Proc. IRE*, vol. 45, pp. 760-767; June, 1957.

⁵ J. L. Heritage, S. Weisbrod, and J. E. Bickel, "A study of signal vs distance data at VLF," *NBS Boulder VLF Symposium*, vol. 4, paper No. 29, pp. 77-81; January, 1957.

⁶ D. D. Crombie, A. H. Allan, and M. Newman, "Phase variations of 16 kc/s transmissions from Rugby as received in New Zealand," *Proc. IEE*, vol. 105, pt. B, pp. 301-304; May, 1958.

A frequency of 12 kc has cyclic ambiguities every 83.3 μ sec. Referred to the baseline, the ambiguities are spaced 6.7 nautical miles.

The navigation equipment also measures the phase of the difference frequency ($f_1 - f_2$) received from each station. Phase comparison of the two difference frequency components establishes the coarse-phase l-o-p.

A 1000-cps difference frequency predicated on f_1 and f_2 signals of 12.0 and 13.0 kc, respectively, provides discreet phase values every 1000 μ sec in the coverage area. Referred to the baseline, the coarse-phase ambiguities would occur every 81 nautical miles.

The final system requires measurement of the coarse-phase to an accuracy of less than one-half cycle of the fine-phase period in order to resolve the latter's ambiguity. If this condition is satisfied, the inherent precision of the fine-phase portion of the system can yield an accurate and unambiguous fix.

For a 12- and 13-kc system, the standard deviation of the phase of the difference frequency must be less than 10 μ sec. Phase recordings obtained in the 12- to 14-kc band indicated that the difference-frequency phase stability was this good or better during the daytime only.

As previously stated, transmissions occur from two or more locations on each frequency necessitating use of time sharing techniques. At each transmitting and receiving location, a commutator (sequence programmer) is employed to control the time and duration of transmissions and associated switching functions. The commutator consists of segments of various lengths arranged in a unique pattern permitting use of correlation techniques to effect synchronization.

In describing the basic system, only a single master and slave station combination is considered since all additional slave stations are identical. Each slave station is assigned in fixed sequence to its respective commutator segment. The equipment necessary for a master station is shown in the simplified block diagram (Fig. 3). Stable reference frequencies are derived from an oscillator having a stability of a few parts in 10^9 per day. The frequency divider following the oscillator produces the two carrier frequencies f_1 and f_2 , and a 10-cps reference for the electronic commutator. The commutator produces the necessary keying information for the transmitter, segment A for keying f_1 , segment B for keying f_2 , and a long segment centered about A for tuning the antenna from f_2 to f_1 . The unused transmission time between segments A and A on f_1 , and B and B on f_2 , are used to accommodate up to seven slave stations.

The slave station synchronizer and transmitter are shown in the simplified block diagram (Fig. 4). Information obtained by the commutation-sync detector is used to adjust the timing of the slave commutator to coincide with that of the received signals. Once commutation synchronism is obtained, very little correction in time will be necessary to maintain it. The commutator switching information is then fed to channels 1 and 2 to perform their necessary functions.

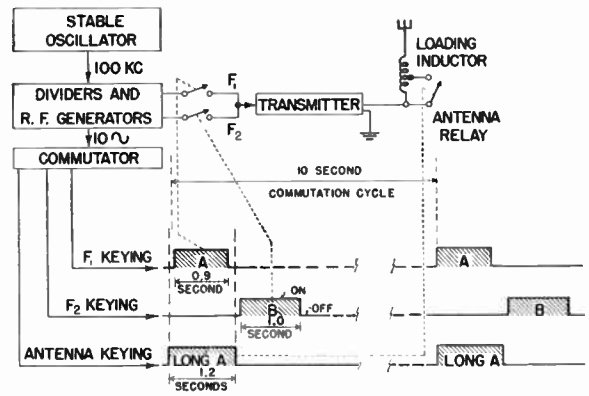


Fig. 3—Master station block diagram.

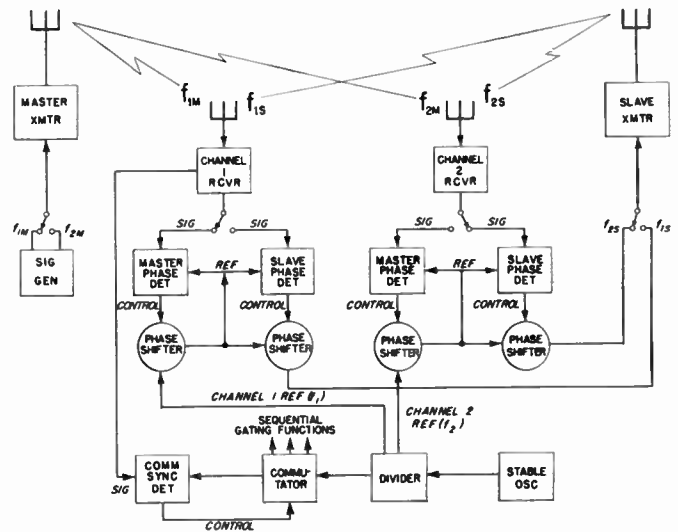


Fig. 4—Slave station synchronizer block diagram.

Consider first the left side or channel 1 portion of Fig. 4. The master station transmission of f_1 during segment A is received by the channel 1 receiver and switched to the signal input of the master phase detector. A reference signal is derived from the frequency divider and passed through the master phase shifter to the reference input of the phase detector. The error voltage developed is fed to a servomechanism which adjusts the master phase shifter until the master reference is in quadrature with the received master signal.

During transmission time B, the slave transmission of f_1 is received by the same receiver as above and is switched to the slave phase detector. It is compared with the master reference, and a servo controls the slave phase shifter until the slave transmissions are in quadrature with the common reference. Since the received master transmissions and slave transmissions are both in quadrature with a common reference, they must be in phase with each other.

Phase measurement and control for the f_2 transmissions is accomplished in an identical manner in the channel 2 portion of the synchronizer.

Since master and slave transmissions are held at zero phase on both f_1 and f_2 , the slave difference frequency is

held at zero phase with the master difference frequency.

A simplified form of a Navigation Receiver (single l-o-p) is shown in Fig. 5. Its operation is similar to that of the slave station synchronizer.

After commutation synchronism is established, the master station and slave station transmissions on each frequency are compared with local references derived from a stable oscillator in the receiver. The error voltages developed control their respective servos and phase shifters to hold the phase detector references in quadrature with the signals. Differentials, between the master and slave phase shifters in each channel, indicate the relative phase of the master and slave transmissions for f_1 and f_2 . Either of these indications may be used as the fine-phase reading for position fixing. A differential between the differentials in channels 1 and 2 indicates the relative phase of the master and slave difference frequencies. This is the "coarse" reading used for lane identification. The "coarse" reading will indicate one cycle of difference frequency every $f_1/(f_1-f_2)$ cycles of f_1 throughout the fine-phase hyperbolic pattern.

INSTRUMENTATION FOR DATA COLLECTION

In the Radux-Omega system, we are concerned with the stability of two quantities: 1) phase of the carrier frequency (fine-phase or vernier indication), and 2) phase of the difference frequency of the two carrier frequencies (coarse-phase for resolving ambiguities). The difference-frequency phase stability is dependent on the relative stability of the two carriers.

A series of propagation tests was initiated to investigate these quantities as a function of frequency throughout the range of 10 kc to 18 kc. One group of tests involved measurement of the round-trip carrier phase from Hawaii to San Diego to Hawaii (Fig. 6), on each of the following frequencies: 10.2, 11.2, 12.2, 13.2, 14.2, 15.2, 16.2, 17.2 and 18.2 kc. A navigation receiver was located at a monitoring site 12 miles from the master station. Proximity to the master station assures constant propagation time of the ground wave signal. Therefore, round-trip measurement yields a phase quantity containing only propagation time variations (master to slave to monitor) and slave synchronization errors. The round-trip instrumentation thus removes the effect of unstable frequency standards in the system.

Round-trip carrier phase tests were instrumented as follows:

- a) The master station in Hawaii was used to transmit on a single frequency f_1 using commutator segment .1 (Fig. 3).
- b) The slave station at San Diego transmitted on the same frequency during the master station off-time (segment B) and was held by the synchronizer to zero degrees carrier phase difference referred to the antenna site. Additional commutator segments (not indicated in Fig. 3) were utilized to provide maximum emission time from each station.

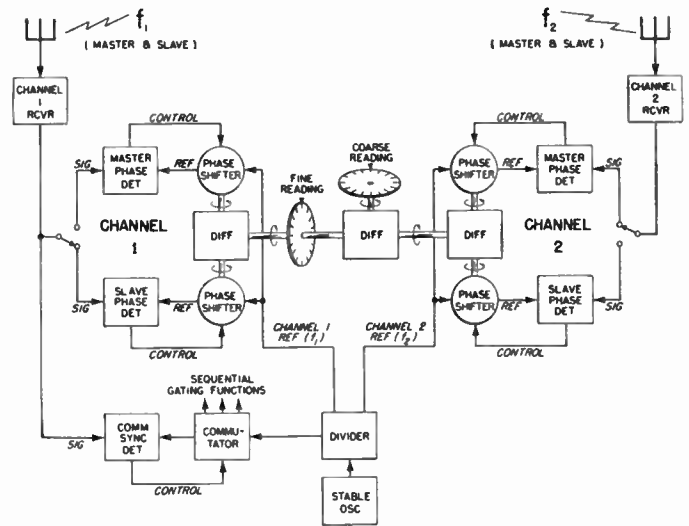


Fig. 5—Navigation receiver block diagram.

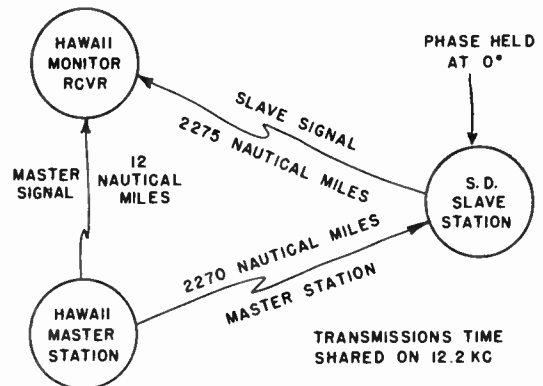


Fig. 6—Round trip propagation path for data collection.

c) Synchronization was accomplished using the channel 1 portion of the Radux-Omega synchronizer (Fig. 4).

d) A monitor receiver was installed at the Hawaiian site 12 nautical miles from the master station to record the phase difference of the master and slave transmissions. This receiver is equivalent to the channel 1 portion of the navigation receiver (Fig. 5). In the Radux-Omega system, this would correspond to the fine-phase reading used for position fixing.

e) A similar monitor receiver operating independently of the synchronizer was installed at the slave station to record the phase difference of the received master and slave transmissions. This provided a continuous check of the degree of system synchronization. Recordings indicated the synchronizer and/or monitor errors to be less than $\pm 1 \mu\text{sec}$.

f) A recorder coupled to the master phase shifter of the synchronizer (Fig. 4) indicated phase variations due to propagation, plus differential drifts in the frequency standards used at each end of the system. A gradual drift in phase is present since it was difficult to hold the reference oscillators at the master and slave locations within a few parts in $10^9/\text{day}$.

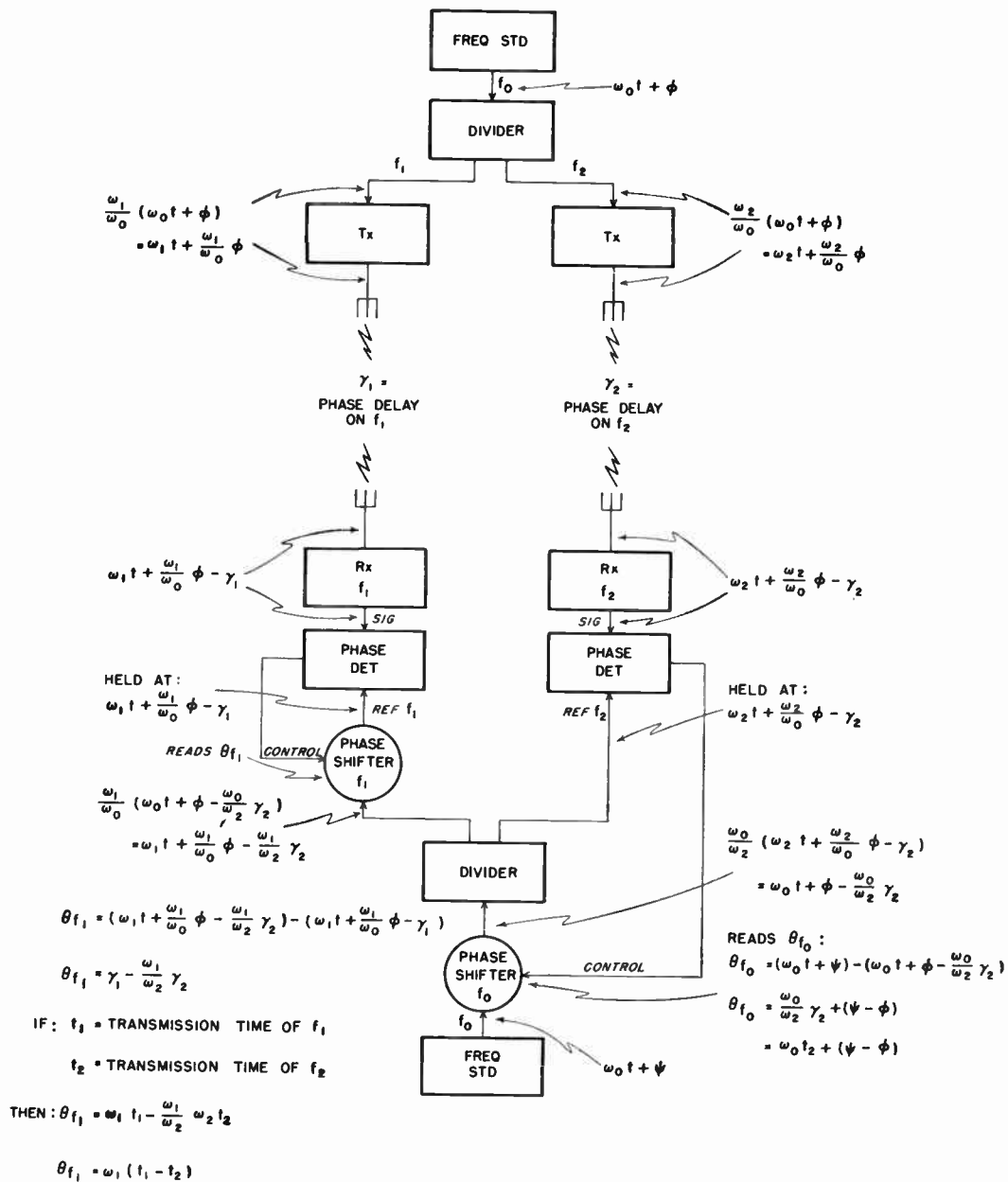


Fig. 7—Instrumentation for two-frequency one-way phase measurements.

g) The master reference in the slave station synchronizer (Fig. 4) was held in quadrature with the received signal. By shifting the master reference 90° and comparing it with the received signal in an auxiliary phase detector, a dc output voltage was obtained which varied proportionally with the signal amplitude. This dc output voltage was recorded to provide indications of variations in signal strength. Since the dynamic range of the receiver was limited, there was some suppression of the signal strength indication due to noise. This relative signal strength reading has not been evaluated in terms of $\mu\text{V}/\text{meter}$; however, the general trend is indicated.

h) An AN/URM-41 field intensity meter was used to record the atmospheric noise level on a frequency offset by 500 cycles from the carrier frequency. These data

have not yet been analyzed but will be used with above signal amplitude data.

A second group of tests involved measuring, at NEL and NRL, the phase of the difference frequency produced by two sequentially keyed carriers spaced 1 kc apart and transmitted from the master station location in Hawaii. The pairs of frequencies investigated were 10.2–11.2 kc, 11.2–12.2 kc, 12.2–13.2 kc, 13.2–14.2 kc, 14.2–15.2 kc, 15.2–16.2 kc, 16.2–17.2 kc, and 17.2–18.2 kc.

Fig. 7 illustrates the instrumentation and signal phase relationships throughout the system, employed in measurement of the desired quantities. This approach was used to 1) remove effects of the phase difference or drift between the two reference frequency standards,

and 2) remove the phase shifts due to equal changes in effective transmission time of each carrier. This is necessary since it is desired to determine the instability of the propagation medium in terms of the relative transmission time of two frequencies. If these variations can be measured, the difference frequency phase may be indicated with respect to the phase of one of the carriers.

In Fig. 7, right-hand channel,

$\omega_2 t + (\omega_2/\omega_0)\phi$ = the radiated phase of f_2

γ_2 = the phase delay of f_2 due to transmission time t_2 , defined by γ_2/ω_2

$\omega_2 t + (\omega_2/\omega_0)\phi - \gamma_2$ = the received phase of f_2 .

For simplicity, zero phase shifts are considered for the fixed delays introduced by equipments.

The received carrier is compared with f_2 reference in the phase detector. Control information is applied to phase shifter f_0 to hold the f_2 reference at zero phase with respect to the received signal. This phase shifter indicates the sum of two quantities:

$$\theta_{f_0} = \frac{\omega_0}{\omega_2} \gamma_2 + (\psi - \phi),$$

where

$(\omega_0/\omega_2)\gamma_2$ = phase delay of f_2 due to propagation, in terms of f_0 phase,

and

$(\psi - \phi)$ = phase difference of the two standard frequency sources.

If

$$(\psi - \phi) = 0, \quad \gamma_2 = \frac{\omega_2}{\omega_0} \theta_{f_0}.$$

Therefore, if the standard frequency sources are sufficiently stable, the recordings of θ_{f_0} indicate the propagation stability of the medium.

In the left-hand channel of Fig. 7,

$\omega_1 t + (\omega_1/\omega_0)\phi$ = the radiated phase of f_1

γ_1 = the phase delay of f_1 due to transmission time t_1 , defined by γ_1/ω_1

$\omega_1 t + (\omega_1/\omega_0)\phi - \gamma_1$ = the received phase of f_1 .

The latter signal is compared with the f_1 reference in the associated phase detector. Control information is applied to phase shifter f_1 to hold the f_1 reference at zero phase with respect to the received signal.

The phase shifter f_1 indicates

$$\theta_{f_1} = \gamma_1 - \frac{\omega_1}{\omega_2} \gamma_2.$$

As the reference f_0 is shifted to correct the f_2 reference, the f_1 reference is shifted in phase equal to

$$\frac{\omega_1}{\omega_2} \gamma_2.$$

Then

$$\gamma_1 = \frac{\omega_1}{\omega_2} \gamma_2, \quad \text{if } t_1 = t_2.$$

Therefore, when t_1 equals t_2 , the phase detector f_1 reference is always held at zero phase with the received signal without additional adjustment of phase shifter f_1 .

$$\theta_{f_1} = \gamma_1 - \frac{\omega_1}{\omega_2} \gamma_2 = 0, \quad \text{if } t_1 = t_2,$$

but in reality $t_1 \neq t_2$, and

$\theta_{f_1} = \omega_1(t_1 - t_2)$ = the effective differential time of transmission of f_1 and f_2 .

This reading also indicates directly the effect of this differential time on the phase of the 1-kc difference frequency.

One-way recordings of two-frequency transmissions were instrumented as follows:

- a) The master station in Hawaii was used to alternately transmit on the two carrier frequencies (f_1 and f_2) to be measured (Fig. 3). Alternate transmissions on f_1 and f_2 were used throughout the full commutation period to increase the duty cycle.
- b) A monitor receiver was installed at the Hawaiian monitor site to measure the phase of the difference frequency (Fig. 7). This monitor provided a check on the local transmissions, free of propagation changes. It indicated the stability of the master station timer, dividers, transmitter, and antenna tuning. Any phase variations recorded at this location were subtracted from the remote monitors to give a true indication of propagational effects. These recordings to date indicate the master station stability of the difference frequency to be constant within $\pm 4^\circ$. This error would be cancelled out in the complete Omega system.
- c) A monitor receiver was installed at NEL, San Diego (Fig. 7), to record the following data:
 - 1) Phase variations of the difference frequency $f_1 - f_2$ in degrees, due to unequal transmission time, $\theta_{f_1} = \omega_1(t_1 - t_2)$. This recording is free of errors due to variations in frequency of the stable oscillators at each end of the path and variations in transmission time which affect both carriers identically.
 - 2) Phase variations of the high frequency carrier f_2 due to propagation (θ_{f_0}). This also includes a drift due to the necessity of using imperfect frequency standards at each end of the path ($\psi - \phi$).
 - 3) Amplitude of both the high and low frequency carriers as instrumented and described for the round trip measurements. These data for the 12.2-13.2-kc test were evaluated in terms of $\mu\text{V/m}$.

- d) An AN/URM-41 field intensity meter was used to record the atmospheric noise level on a frequency midway between f_1 and f_2 .
- e) The NRL designed and constructed a monitor receiver to record the one-way two-frequency phase measurements. They operated this receiver during the tests to provide the following concurrent data over the greater path length.
 - 1) Independent phase variations of the high- and low-frequency carriers (f_1 and f_2).
 - 2) Relative transmission time in microseconds, $t_1 - t_2$.

This one-way propagation test provided a direct comparison of the effects of the propagation medium on two closely spaced frequencies transmitted during the same period. It demonstrated the degree of capability of the system to resolve lane ambiguities by 1-kc difference phase measurements. The propagation data obtained are useful in determining a suitable operating frequency and carrier spacing. A record of signal strength variations on two closely spaced frequencies during the same period of operation was also obtained. This provided a comparison of propagation effects over two paths at the same time, *i.e.*, Hawaii to San Diego and Hawaii to Washington, D. C.

COLLECTED DATA

All data presented herein were collected on strip chart recorders and plotted as shown. At the time of submission of this paper data were available only for the frequencies between 10.2 and 16.2 kc.

Round-Trip Single-Frequency Transmissions

The round-trip data collection program was inaugurated to provide a measure of carrier phase stability as one factor in determination of optimum frequencies for use in the Radux-Omega system. The complete picture is not revealed by data taken over a particular propagation path. However, the unstable phase conditions experienced in these tests should occur in varying magnitude in other portions of the coverage area of a long-range navigation system.

The round-trip data were instrumented as degrees of phase delay (Fig. 5, channel 1). In plotting, the phase delays were converted to equivalent microseconds in order to compare various frequencies on a common basis for relative stability and magnitude of diurnal change.

Figs. 8–14 illustrate round-trip transmission between Hawaii and San Diego. Fig. 8, plotted from data over a 6-day period in January, 1958, shows half-hour mean and standard deviation of 6-minute readings of the phase delay on 12.2 kc. The increase in standard deviation between 2100 and 2200 GMT was due to a sudden ionospheric disturbance (S.I.D.) at 2113 GMT on January 16, 1958, and recorded by our daily phase measurement.

Solar-Geophysical Data⁷ corroborates this. The ordinate represents relative values of phase difference between the transmitted signal from Hawaii and received signal over the round-trip path via San Diego. All readings between 1800 and 2300 GMT produced a standard deviation of 4 μ sec; between 0430 and 1430 GMT, 5 μ sec. These results imply a good day-to-day repeatability of a 1- σ to within $\frac{1}{2}$ nautical mile for this particular path and transmission frequency.

More recent data obtained over the same propagation path, represented by Figs. 9–14, are in most cases for a single 24-hour operation. Although operations were conducted over a continuous 48-hour or longer period, equipment failures in the Hawaiian monitor station usually precluded using but one uninterrupted 24-hour period of data. The ordinates in Figs. 9–14 are on a relative phase shift scale. Unfortunately, purely from oversight, the monitor in Hawaii was not zeroed prior to round-trip propagation tests on each frequency. For this reason, the aforementioned plots can be compared only on a relative basis. These comparisons are valid for magnitude of diurnal change as well as for phase variations over corresponding hours.

In subsequent monitoring, a common zero phase reference setting will be established to provide recordings on an absolute delay scale. Data will thus be recorded on each frequency as an integral number of carrier cycles plus a fraction.

It is noted that greater night-time stability is experienced on the lower values of frequency. For example, the peak-to-peak night time phase variation is approximately 8 μ sec for 10.2 kc and 50 μ sec for 15.2 kc. The variations are attributed to phase interference between the various propagation modes. Additional information will be extracted from the round-trip phase recordings of the individual frequencies noted above. Subtraction of phase variations between pairs of frequencies will aid in prediction of the effects of propagation on a difference frequency in multiples of 1-kc separations.

During the entire data-taking program, fairly frequent occurrences of S.I.D.'s were recorded during daylight on the propagation path. Three S.I.D.'s appear on Fig. 9, occurring at 2041, 2131, and 2147 GMT. On Fig. 10, an S.I.D. occurred at 1951 and 2322 GMT. The unique character of an S.I.D. makes its occurrence easily recognized by graphic recording. Allowances can, therefore, be made.

One-Way Two-Frequency Transmissions

Figs. 15–21 show how the effective differential transmission times of the carrier frequencies influence the phase of the 1-kc difference frequency. Data were recorded at NEL and NRL throughout 72-hour periods in all cases.

⁷ Natl. Bur. of Standards, "Solar-Geophysical Data," CRPL-F 162, pt. B, p. IVa; February, 1958.

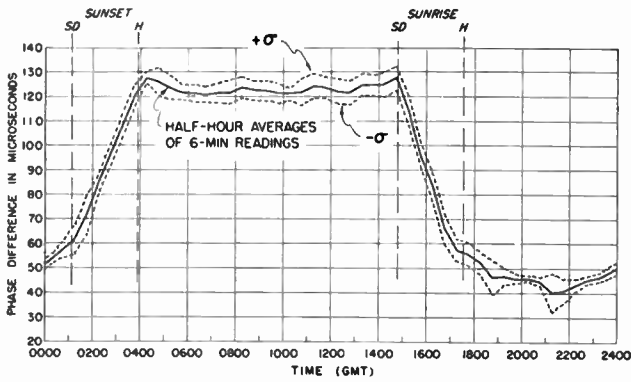


Fig. 8—Round-trip phase difference measured in Hawaii for 6 days between January 15 and 23, 1958, San Diego-Hawaii pair.

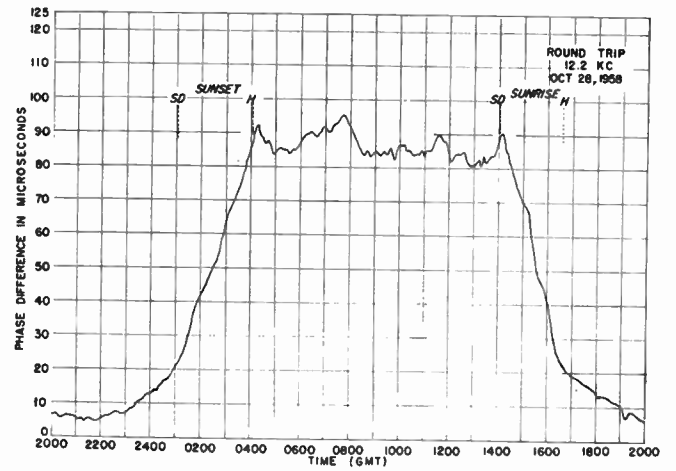


Fig. 11—Round-trip phase, 12.2 kc, Hawaii-San Diego.

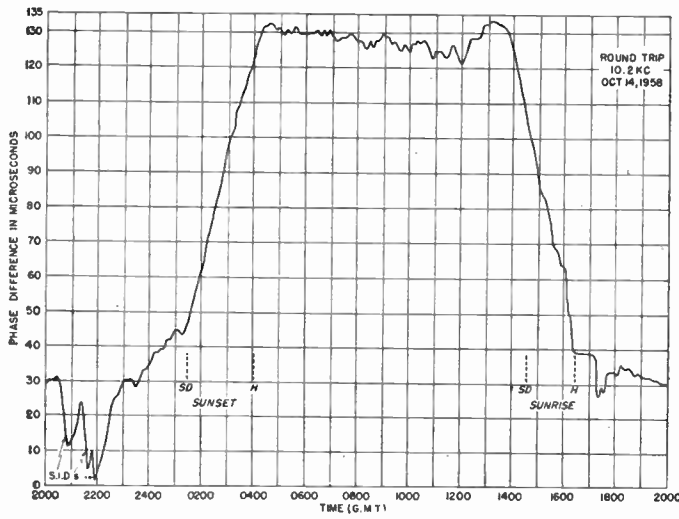


Fig. 9—Round-trip phase, 10.2 kc, Hawaii-San Diego.

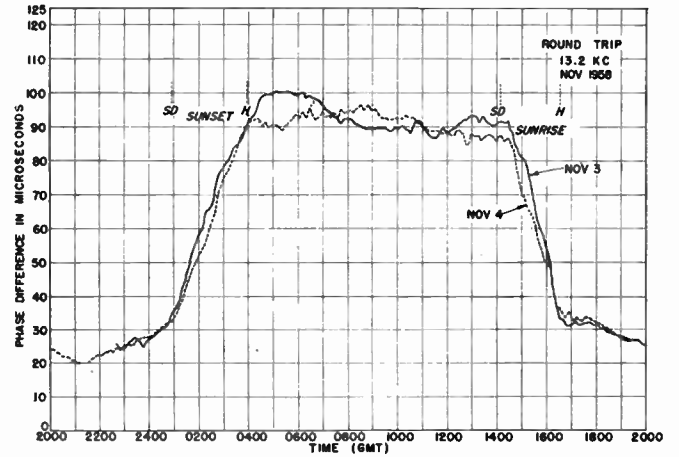


Fig. 12—Round-trip phase, 13.2 kc, Hawaii-San Diego.

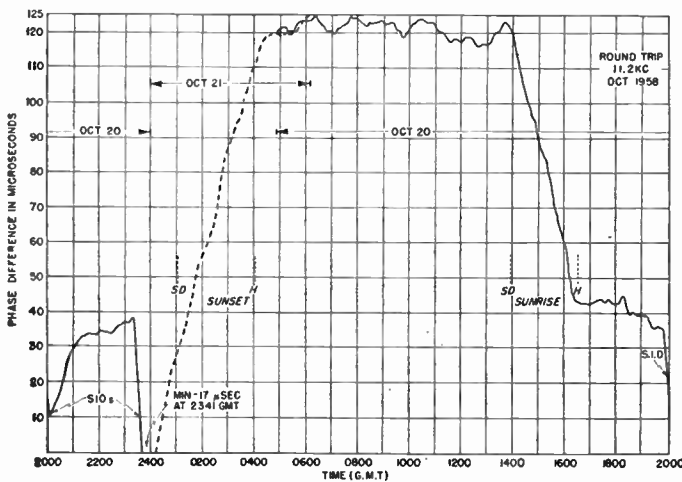


Fig. 10—Round-trip phase, 11.2 kc, Hawaii-San Diego.

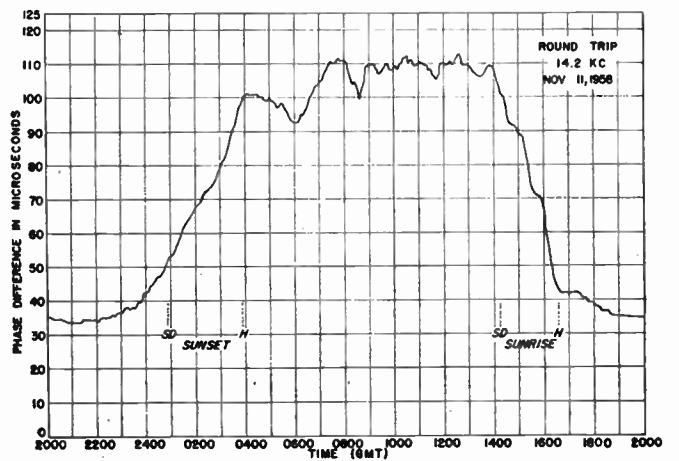


Fig. 13—Round-trip phase, 14.2 kc, Hawaii-San Diego.

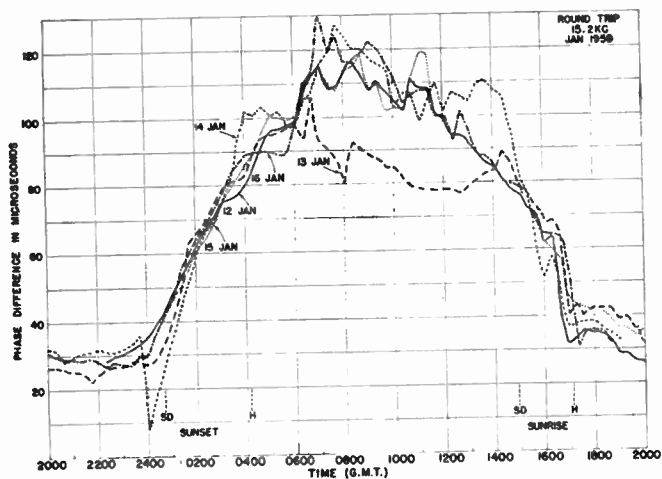


Fig. 14—Round-trip phase, 15.2 kc, Hawaii-San Diego.

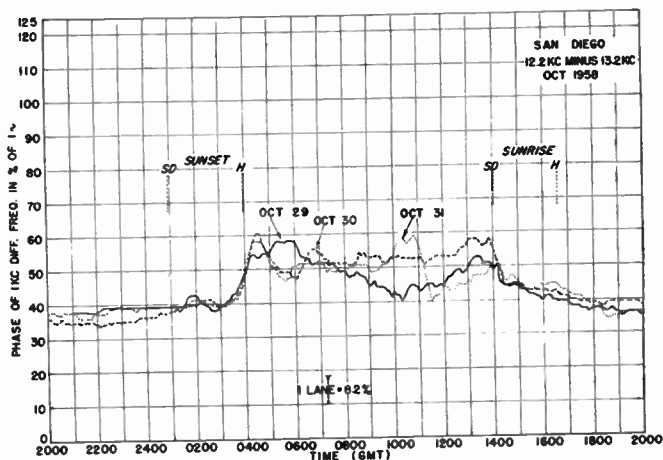


Fig. 17—One-way difference phase, 12.2 kc-13.2 kc, Hawaii-San Diego.

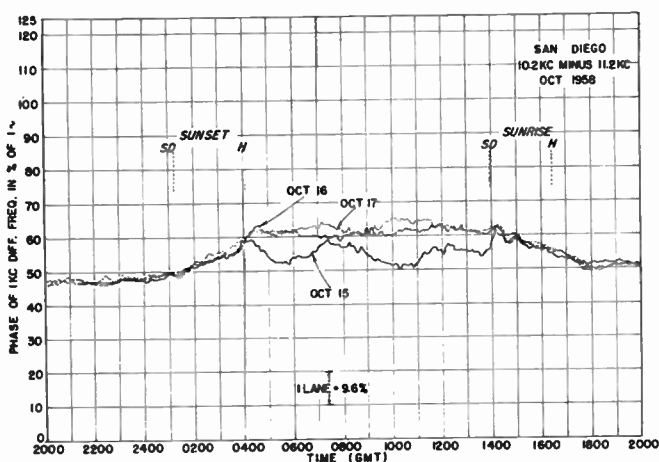


Fig. 15—One-way difference phase, 10.2 kc-11.2 kc, Hawaii-San Diego.

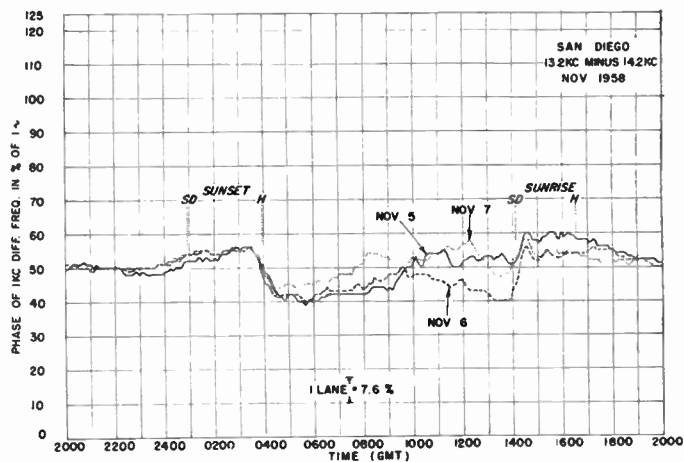


Fig. 18—One-way difference phase, 13.2-14.2 kc, Hawaii-San Diego.

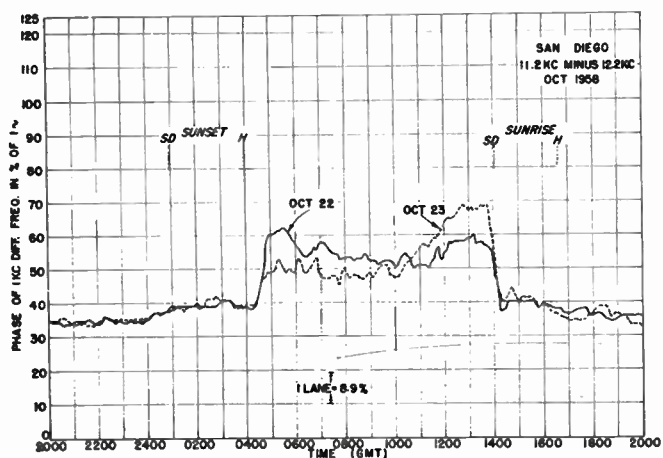


Fig. 16—One-way difference phase, 11.2 kc-12.2 kc, Hawaii-San Diego.

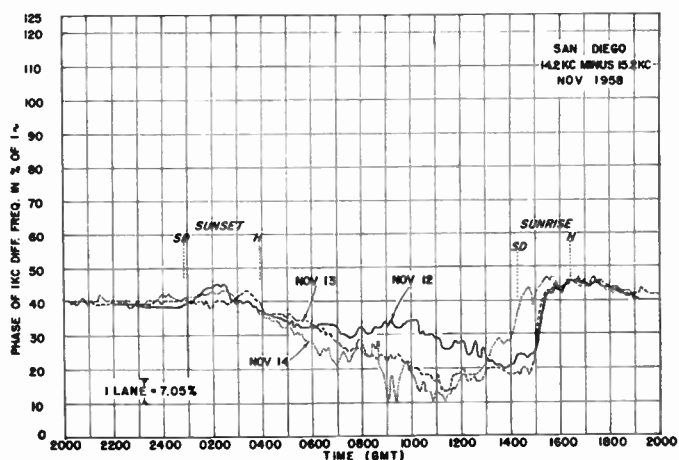


Fig. 19—One-way difference phase, 14.2-15.2 kc, Hawaii-San Diego.

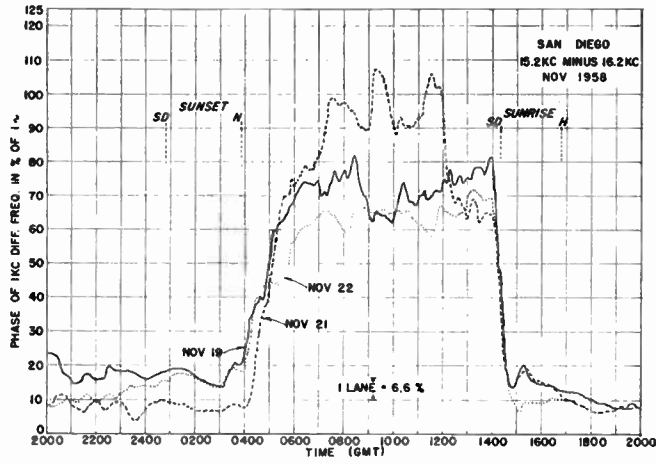


Fig. 20—One-way difference phase, 15.2 kc-16.2 kc, Hawaii-San Diego.

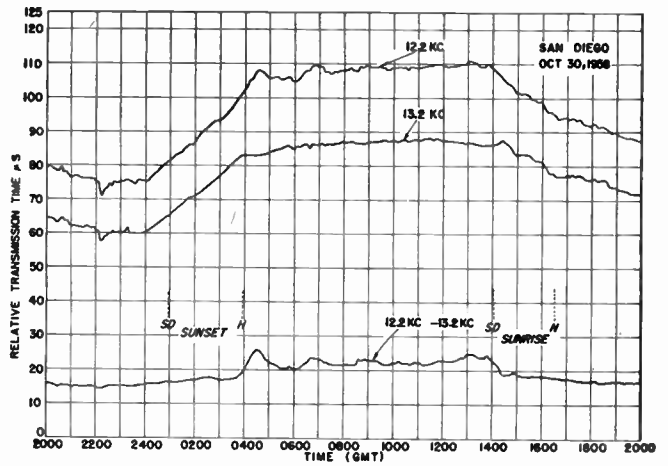


Fig. 22—One-way relative transmission time, 12.2 kc, 13.2 kc, and 12.2 kc-13.2 kc, Hawaii-San Diego.

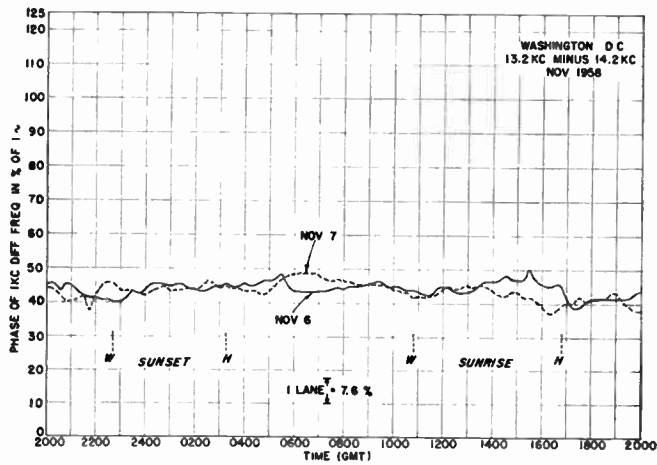


Fig. 21—One-way difference phase, 13.2 kc-14.2 kc, Hawaii-Washington, D. C.

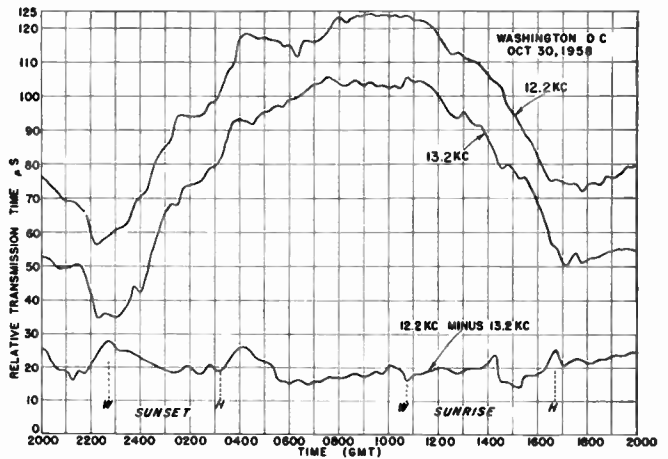


Fig. 23—One-way relative transmission time, 12.2 kc, 13.2 kc, and 12.2 kc-13.2 kc, Hawaii-Washington, D. C.

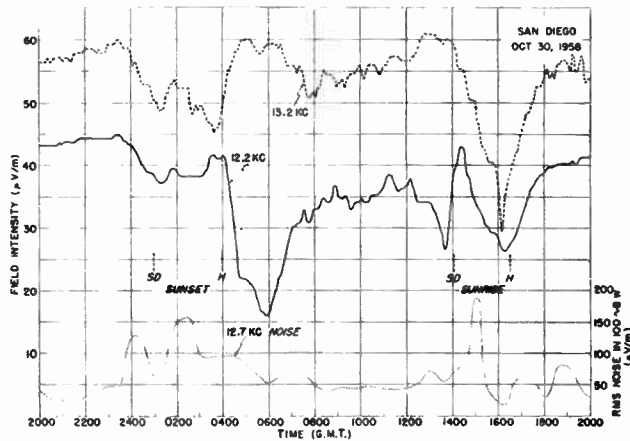


Fig. 24—Carrier field intensity 12.2 kc and 13.2 kc, Hawaii-San Diego, rms noise 12.7 kc.

The following is observed by visual inspection:

- 1) Standard deviation of the NEL recordings are lower in value during the daylight hours.
- 2) Lane resolution should be possible during daylight hours according to NEL data for frequencies not greater than 14.2 kc.
- 3) In some cases the diurnal change does not affect the mean values until after sunset time in Hawaii.
- 4) The diurnal change is of the opposite sense in Figs. 18 and 19, as compared with other photos.
- 5) By comparing NRL data (Fig. 21) and NEL data (Fig. 18) taken during the same period, it can be seen that there is more randomness and the diurnal change is less pronounced in the NRL data.
- 6) Difference frequency phase repeats reliably from day to day during the daylight hours.

Figs. 22-24 should be compared. Fig. 22 shows the received phase of the 12.2-kc and 13.2-kc carrier at San Diego, for one 24-hour period of operation. Phase is shown as relative transmission time in microseconds for easier comparison.

The 13.2-kc phase record was obtained by recording θ_{f_0} as described previously and contains both propagation phase γ_2 and phase difference of the two frequency standards ($\psi - \phi$). This 24-hour period was chosen since the frequency standards were unusually stable for this time, and ($\psi - \phi$) was small.

The 12.2-kc phase record was obtained by adding (12.2 kc-13.2 kc) the difference phase to the 13.2-kc phase, and it also contains ($\psi - \phi$).

This provides a graphical presentation of the differential transmission time with which this study is concerned.

Fig. 23 shows similar information collected by NRL during the same period of operation. Their data, however, were instrumented in a slightly different manner to provide 13.2-kc carrier phase, 12.2-kc carrier phase and their difference in propagation time directly.

Since the quantity ($\psi - \phi$) is small, the diurnal change is readily seen and may be compared.

Fig. 24 presents the amplitudes ($\mu\text{v}/\text{m}$) of the 12.2-kc and 13.2-kc carriers as received in San Diego for this same period of operation. The radiated power from the Hawaiian station on 12.2 and 13.2 kc was estimated at 230 and 270 w, respectively. Rms atmospheric noise (12.7 kc) in a 100-cps bandwidth is also plotted. It is interesting to note the diurnal effect on frequencies

separated by only 1 kc. This is attributed in part to phase interference between various propagation modes.

CONCLUSIONS

Based on the data taken to date, it appears that 24-hour per day lane resolution is not probable in a two-frequency system over a distance corresponding to the Hawaii-San Diego path. It is believed that at least part of the instability observed is due to interference between various propagation modes at this distance. At longer distances, the situation is expected to improve, but present indications are that the lane resolution will be marginal.

If further tests prove lane resolutions to be impractical, consideration should be given to a single-frequency Omega system. Accurate dead reckoning traces could resolve the lane ambiguities of approximately 7-mile spacing. A single-frequency system would obviously be considerably less complex than the two-frequency system and would allow appreciable reductions in size, weight, and cost of the navigator's equipment.

Before all parameters of the Omega system can be specified, propagation data are required as a function of distance. Mobile monitors are needed to explore parts of the coverage area to obtain information on zones of interference. The width of these zones and the severity of phase instability must be known. Tests are planned for 1959 which will yield much of this information.

The propagation data presented herein and data to be obtained by the conclusion of these tests should be of value for comparison with theories of propagation in the VLF spectrum. The instrumentation techniques used in this project are considered very effective for any future data collection and study of VLF propagation.

ACKNOWLEDGMENT

A development project of this magnitude requires contributions from many individuals and organizations. The authors wish to acknowledge the work of the Naval Research Laboratory and permission to include their data in this report.

J. A. Pierce, the inventor of Radux, proposed Radux-Omega from results of LF-VLF propagation studies. His contributions as a consultant throughout this program were invaluable.

The work of the VLF Navigation Section and the Transmitter Section of Navy Electronics Laboratory is acknowledged in connection with the design and development of the instrumentation.

In-Service Improvements in Air Traffic Control*

P. T. ASTHOLZ†, SENIOR MEMBER, IRE

Summary—This paper covers several of the significant “in-service improvements,” including the introduction of simulation study techniques, new displays, and automatic data processing that have been taken to meet the air traffic control requirements. A review of the three basic simulation methods: graphical, dynamic, and fast time is provided together with a description of the dynamic simulation facilities at the Technical Development Center of the Federal Aviation Agency (FAA), Indianapolis, Ind. The requirements for daylight viewing displays and the programs conducted to obtain TV types of radar displays are included. The report concludes with the introduction of digital computers to perform certain data processing functions.

INTRODUCTION

A STEADY, but not widely known, improvement in air traffic control has been obtained by the Civil Aeronautics Administration (now the Federal Aviation Agency as of December 31, 1958) through the introduction of certain new study techniques, new equipments, and new procedures in an attempt to keep pace with the growing air traffic and the user requirements. While certain major and long-range steps are required, and programs to accomplish them have been introduced by the Airways Modernization Board and the new Federal Aviation Agency, it is believed that some of the day-to-day steps sometimes called “in-service improvements” are of interest. In the final report by the President’s special assistant, referred to as the Curtis Report,¹ a three-part program for meeting the nation’s future needs was outlined. This program, as summarized in the report, was as follows:

- Phase I—Immediate steps to augment and improve the current plans of the Civil Aeronautics Administration.
- Phase II—Immediate establishment and execution of a unified modern research and development program, addressed principally to early application of existing scientific technology.
- Phase III—Establishment of a permanent program of continuing modernization of the nation’s aviation facilities.

The in-service improvements are now generally called Phase I of this program although some overlapping into Phase II is apparent. The safety and capacity of the air traffic control (ATC) system have been increased with the implementation of a two-way VHF and UHF communication system between the pilot and controller,

with the VORTAC navigation system, and with reorganization of the airway system. This paper covers several of the other significant improvements, including the use of simulation techniques to study basic ATC characteristics, improved radar displays, and introduction of digital computers for semiautomatic data processing.

SIMULATION

In order to obtain a firmer grasp of a problem area, it is important to understand the basic characteristics and functions, together with the requirements involved. If possible, a mathematical model of the system is created to aid in defining the parts of the system. In human-machine relationships, such as those encountered in air traffic control, it is difficult to describe certain parts of the system in mathematical terms. As a result, simulation techniques are being used as a study tool to provide estimates of the level of system performance using different combinations of human controllers, data processing equipment, and information displays. Actually, the simulation of air traffic control systems, equipments, and procedures has existed for a number of years. The results of earlier work led to the establishment of simulation facilities at the Technical Development Center (TDC) of the Civil Aeronautics Administration in 1950.² This program has led to a number of improvements in the present ATC system, as well as a number of new concepts which are expected to have application in future systems. In fact, the results of earlier simulation studies, as applied to the operating ATC system, have produced sufficient field data to evaluate the effectiveness of these simulation studies. Comparison of the studies with actual field data has indicated that the results are closer to actual operations than had been anticipated and has resulted in an increased demand for expansion of this type of study activity.

Since its conception, this work at IDC has been aided by The Franklin Institute Laboratories for Research and Development, and more recently, by Courtney and Company, in consideration of the human factors involved, the proper formulation of the study problems, and the analysis of data obtained during the evaluations. A number of other organizations have conducted work in this field. Some of these are: the International Business Machines Corporation, Airborne Instruments Laboratory, and the Armour Research Foundation. The Laboratory of Aviation Psychology of Ohio State University has been playing an important and outstand-

* Original manuscript received by the IRE, December 2, 1958; revised manuscript received, February 6, 1959.

† Tech. Dev. Center, Fed. Aviation Agency, Indianapolis, Ind.

¹ E. P. Curtis, “Aviation Facilities Planning, Final Report by the President’s Special Assistant,” Supt. of Documents, Govt. Printing Office, Washington, D. C.; May, 1957.

² R. E. Baker, A. L. Grant, and T. K. Vickers, “Development of a Dynamic Air Traffic Control Simulator,” Library of Congress, Washington, D. C., TD Rep. No. 191, PB No. 127, 846; October, 1953.

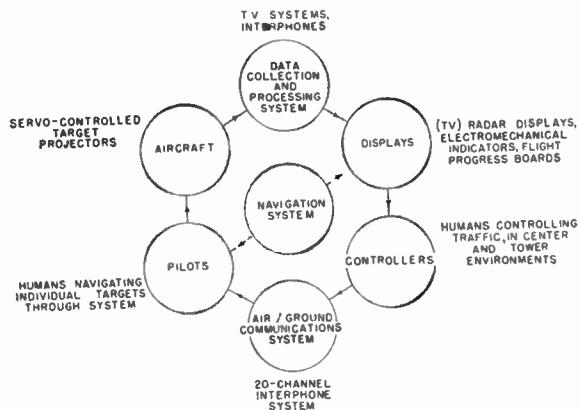


Fig. 1—Functional diagrams of ATC system with analogs employed in dynamic simulation.

ing role for a number of years in establishing psychological principles applicable to the design and operational use of future air traffic control equipment and procedures.

Simulation studies can be classified in a number of ways but generally are grouped under graphical, dynamic, and fast-time methods. Since all of these methods have been described in detail in other papers and in various reports published by the different agencies which have conducted work in this field, a detailed review of these methods is omitted in this paper.

Graphical simulation is normally considered a paper-and-pencil technique. The usual graphical plots used include altitude-time, space-time, and number-time curves depending upon the air traffic control characteristic under study. This method of simulation has been used actually to solve many of the less complex problems and is in use to determine the ideal performance of a specific traffic control system when presented with a certain air traffic load.

Dynamic simulation is the principal technique in use at the FAA's Technical Development Center to date, although graphical simulation is used to obtain the ideal performance results for comparison with dynamic simulation results. As expressed earlier, the human variables in air traffic control are large factors, difficult to determine accurately, and subject to change with the system being studied. The dynamic simulation technique in use is real-time simulation in which the system with its various parts and procedures can be tried under properly designed, controlled, and representative conditions. The environment around the human controllers is made as close to actual operations as possible, and the simulated aircraft inputs are manned separately in order to provide the real-time communications required and the aircraft characteristics affecting air traffic control. Problems given the simulator are in terms of certain arrangements and densities of aircraft, equipments, and procedures. To understand some of the principles used, Fig. 1 provides a greatly simplified functional diagram of the ATC system showing a closed-loop process involving a number of elements linked together. Because

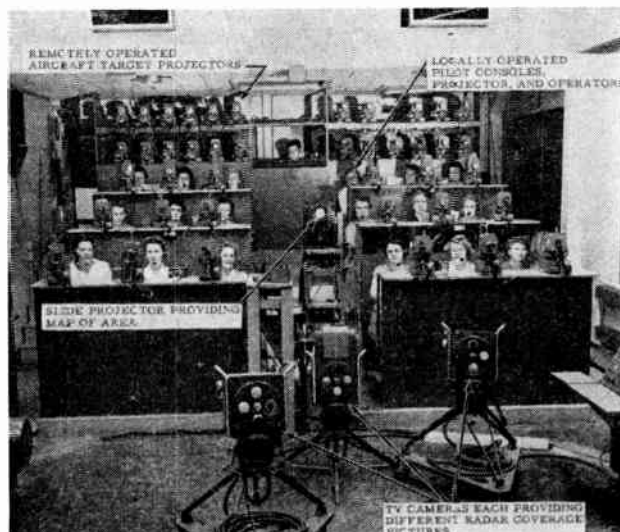


Fig. 2—Pilot consoles, projector equipment, and TV cameras used in ATC dynamic simulator.



Fig. 3—Remote controlled pilot consoles in dynamic simulator.

of interactions between these elements, a change in one variable may affect other variables in the system, sometimes in devious ways. For this reason, the analysis and evaluation of an air traffic control system can become a very complex procedure. The simplified analogs employed in dynamic simulation are indicated outside the ATC elements in Fig. 1.

Briefly, the operating dynamic simulator includes a large screen on which a map of the area or arrangement under study is projected. Spots of light, controlled by electromechanical means from remotely located pilots' consoles, are used to indicate the actual position of 42 different aircraft on the map. Figs. 2 and 3 show the target projectors, pilot console operators, screen, and other components. At the present time, the resulting picture on the screen is televised and presented as radar data on different types of displays. Interphone channels connect the pilots to the simulated control tower and en route traffic control centers. Recording means are



Fig. 4—Air route control room layout in dynamic simulator.

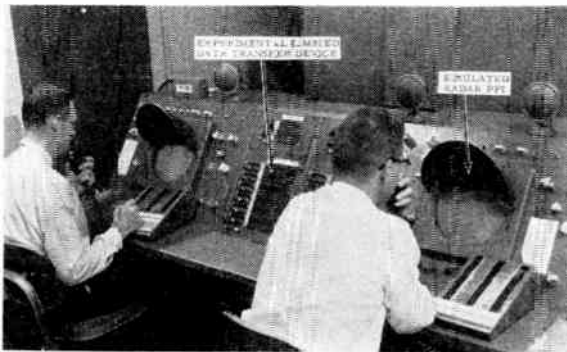


Fig. 5—Dynamic simulator terminal area control consoles.

provided to obtain data for evaluating the tests. Fig. 4 shows a typical arrangement of a new en route ATC display under evaluation, and Fig. 5 provides a view of the typical layout for terminal area tests.

The limitations of the techniques used in the existing dynamic simulator were recognized and studies³ were conducted in cooperation with The Franklin Institute and Computing Devices of Canada, Limited, to determine technical approaches and specification requirements for a new improved and more flexible dynamic ATC simulator. From this work, it was concluded that digital techniques used in the design of a large dynamic simulator offered certain advantages over that provided by analog techniques in terms of high accuracy, large capacity, great flexibility, and data processing. The digital approach, where the basic parameters throughout the simulator can be altered by relatively simple means, appears to be best for a research tool. However, for the study of specific present problem areas where cost, availability, and interchangeability with existing simulation equipments are of utmost importance, the analog approach has certain major advantages. Recently, the FAA placed an order for a 63-target dynamic air traffic control simulator system using analog techniques. In addition to simulating relationships

between aircraft and fixed ground installations, such as radars and projected ground display system, the equipment provides a more automatic means of data collection and reduction using digital techniques.

A third approach to simulation tests has been through the use of digital computers for nonreal time operation. Recent work by IBM, using computer simulation, has as its objective the determination of the relative capacities of two airway systems for the New York area, study of the feasibility of using adjustable altitude blocks at key intersections of an airway system, preparation of realistic and statistically valid air traffic samples for inputs to simulation studies, and a study of several flow control concepts. The advantages of this method are: the rapid running of the problem, larger scope of operation, low errors, and the inclusion of data-processing means to analyze the results. The disadvantages include the time and workload in developing the programming steps required and the difficulty of including unknown human functions. However, it now appears that a complete simulation study facility will use all three methods, that is, graphical, dynamic, and fast-time, since they supplement each other. For example, as more data on human characteristics in an ATC system are obtained by dynamic simulation tests, these can be applied to the graphical and fast-time study methods to gain the advantages offered by these methods. In planning complex human-machine experiments, a choice must be made between obtaining data on an idealized system or on a system with some unknown degree of degradation, such as obtained from radar clutter, communication interference and noise, equipment malfunctions, and so forth. Actually both types of experiments are required, one to provide an estimate of the upper limit of performance and the other to provide the practical status.

Prior to the introduction of simulation techniques, the difficulty of testing and analyzing air traffic control systems was a critical handicap to the development of systems improvements. This difficulty was caused not only by inherent complexity of the system, but by the fact that any system tests required the use of actual aircraft flying in the actual system environment. In terms of results, such tests are extremely slow, awkward, and expensive. Since actual development and installation are involved, system development and modification requires lengthy periods of time. Comparative performance measurements are almost impossible to achieve due to the difficulty of duplicating the initial traffic input and flight conditions on any subsequent tests. In addition, it is difficult to get the weather to "cooperate" with the test conditions; and unless the test operations are isolated from other air traffic, the safety considerations prohibit complete test conditions. Although live flight testing is necessary to provide final proof that the system operates satisfactorily as installed, a large part of the system design and operational testing during the development phase can be handled through the use of simulation techniques.

³S. M. Berkowitz and R. S. Grubmeyer, "Requirements for a new universal air traffic control simulator," IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS, vol. ANE-4, pp. 59-64; June, 1957.

The introduction of civil jet aircraft into the present ATC system puts us on the threshold of a challenging new era. Actually, a considerable amount of information and experience already are available as a result of military jet operations in many areas. One of the active and interesting areas of work is the application of simulation techniques to improve our understanding of the problems and to provide system design guidance together with the operating procedures to be used for various traffic densities. Through the use of simulation, it has been possible to fly thousands of civil jet operations⁴ through high-density traffic conditions representative of those forecast for the next five years.

RADAR DISPLAYS

It generally is agreed that the human controller will be retained in future air traffic control systems.⁵ Suitable displays will continue to be essential to provide him with information on the movement of all aircraft and to permit him to determine the control instructions required for safe and expeditious movement of air traffic. Most of the work completed to date indicates the need for both a pictorial-situation display, showing current pictures and relationships between aircraft, and a planning display, showing proposed occupancy of airspace to permit advance planning for control and flow regulation.

With the introduction of radar into the ATC system, various arrangements of displays were tested, the pictorial display being provided by a PPI and the symbolic information in the form of handwritten paper strips placed on flight progress boards. For both en route and terminal air traffic control, the association of the two types of displays and the high ambient light level encountered placed a critical problem on the radar display. Most of the radar indicators in use utilized cathode-ray tubes having P-7 or P-19 phosphors or dark-trace techniques. Although these indicators met the general storage requisite set by the scan methods of radar, they were restricted in their use to areas of relatively low ambient light levels. Various types of lighting systems were developed to cope with the problem of using handwritten data located beside the PPI, and although fair results were obtained with blue lighting, minus-green, and polarized systems, none was completely satisfactory.

Tests performed at the TDC in 1951 indicated that the scan-conversion techniques using the RCA Graphechon storage tube should provide a satisfactory approach to this problem. This technique involves the writing of radar information on a storage element within the Graphechon, then reading the stored picture out at such

a rate that a continuous display may be presented on cathode-ray tubes having bright and short-persistence phosphors, thereby providing a display which may be viewed under relatively high ambient lighting. A complete system was built using the Graphechon storage tube and an operational and technical evaluation of the equipment was conducted at TDC. The limitations of the original system basically were those of the storage tube and consisted of problems in resolution, discernibility of weak signals, and storage decay characteristics.

A number of other scan-conversion and direct-view storage tubes were tested in an attempt to obtain a daylight viewing radar display. A breakthrough was made in this field by the Compagnie Général de Telegraphie sans Fil (CSF) with their introduction of the Type 403X scan-conversion storage tube. The tube used in this conversion process has excellent storage and good resolution characteristics for this application. The general construction of the tube is shown in Fig. 6. This

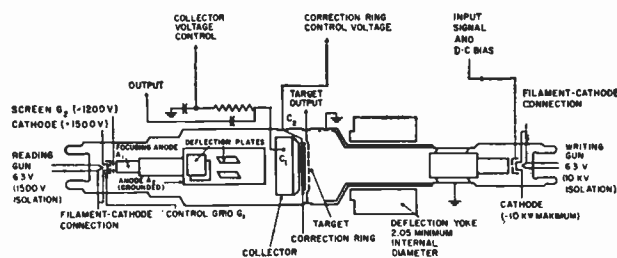


Fig. 6—TMA 403X scan conversion storage tube.

tube employs a writing gun mounted at one end, a reading gun mounted at the opposite end, a thin insulating target between the two guns, and a collector cylinder surrounding the target. The construction of this storage tube is similar to that of the Graphechon; however, the tube is larger. The writing beam is magnetically deflected and electrostatically focused, while the reading beam uses electrostatic deflection and focus. System horizontal resolution was measured in excess of 600 lines. Storage is adjustable from about 0.1 second to approximately 20 minutes. The signal output from the TMA 403X is derived from both the target back plate and the collector. When these two signals, which are 180 degrees out of phase, are applied to the grid and cathode, respectively, in a preamplifier stage, a greater output level is available than from types of tubes previously available. At the same time, improved circuitry was incorporated to cancel crosstalk signals. This tube is employed in a Type T1-440 scan-conversion equipment manufactured by CSF and distributed in this country by Intercontinental Electronics Corp. (INTEC). It is capable of halftone rendition. During image decay, the target trail exhibits several steps of grey scale. The output is shown on a TV monitor in Fig. 7. The display uses a standard of 625 lines, interlace 2:1, and a 30-cycle frame rate.

This conversion progress provides a radar signal in the form wherein the many developments of the vast tele-

⁴ P. T. Astholz and T. K. Vickers, "A Preliminary Report on the Simulation of Proposed ATC Procedures for Civil Jet Aircraft," TD Rep. No. 352, Office of Tech. Services, U. S. Dept. of Commerce, Washington, D. C.; May, 1958.

⁵ F. S. McKnight, "Operational Requirements for ATC Displays," TD Rep. No. 308, Order No. DB131387, Office of Tech. Services, U. S. Dept. of Commerce, Washington, D. C.; April, 1957.

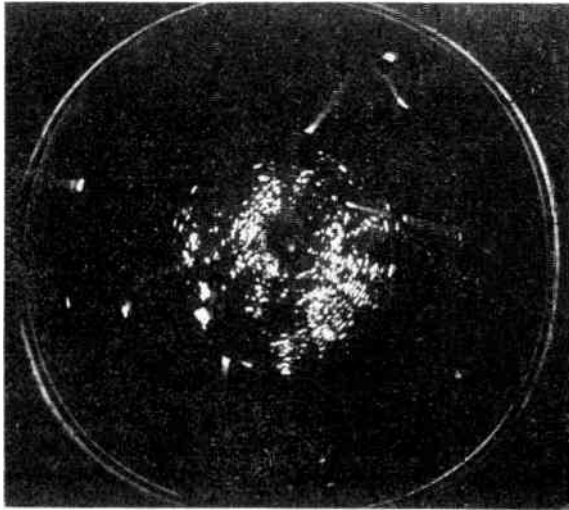


Fig. 7—T1-440 radar display.

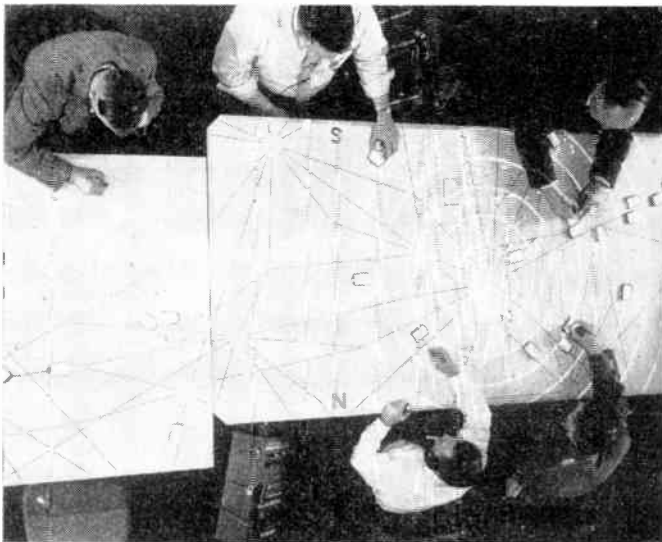


Fig. 8—Panoramic ATC display.

vision industry may be utilized to advantage for air traffic control displays. Some of these advantages include availability of highly engineered reliable equipment priced in a competitive market. Such equipment includes monitors, mixers, video switching devices, distribution equipment, and remoting systems, all of which can be used as building blocks for attaining maximum efficiency in meeting the specific requirements of individual control towers and en route centers.

Although the display of a scan-converted radar picture can be relatively simple, and since standard types of TV monitors may be used, the operational necessity of coordinating all of the known information of aircraft, the majority of which is written on flight progress strips, with specific radar targets places additional design requirements on the displays. Methods of using the scan-converted picture along with aircraft identifying markers have been explored. An early approach was to project the picture from overhead onto a large horizon-

tal plotting surface on which a map of the area was drawn. The plotting markers, sometimes called "shrimp boats," containing essential control information, such as aircraft identity, altitude, route, type, speed, destination, and so forth, were moved along by hand in such a manner that the radar return of the aircraft fell on or beside its corresponding marker. This method of display appears to have some merit for use in en route control; however, the available TV projectors lacked the brightness required for operating under sufficient ambient light since penciled notations were used on the markers. This manual tracking type of display is known as PANOP, or Panoramic Operational Radar Display, and is at the present time receiving further evaluation using a rapid process film radar projection system. Fig. 8 is an overhead picture of the display, as generated by a Kelvin-Hughes rapid process film system.

The availability of 22-inch flat-face Kinescopes for use in horizontal plotting displays solved the immediate operational requirement. As a result, this type of display, together with the scan-conversion equipment, is being procured and installed in the air traffic control system on a country-wide basis. One arrangement of the display is shown in Fig. 9. Small clear plastic markers are used, and the identification and sometimes the altitude information are noted on the marker with a grease pencil. The remaining essential information on the aircraft is carried on the flight progress strips located in the central part of the display. Additional large-size TV monitors also are being obtained for location in both vertical and horizontal positions to supplement these displays.

Another display known as SPANRAD, or Superimposed Panoramic Radar Display, evolved from efforts to utilize the scan-converted picture in a plotting technique with markers under high ambient light levels. Fig. 10 shows one of these displays in operation. A 27-inch TV monitor is mounted vertically alongside a horizontal plotting table. The plotting table is viewed from overhead by a Vidicon camera synchronized to the scan-conversion system, and the resulting picture is electronically mixed with scan-converted radar picture on the 27-inch monitor. Manual tracking then is performed by placing the markers on the horizontal board so that they are superimposed with the radar targets. This type of display is still under test in the dynamic simulator at TDC in an operational environment.

Tests have been conducted to determine the operational use of the SPANRAD if different video inputs were in color. For example, the plotting surface information was in red, the radar input in green, and the video mapping was in blue for one test. One advantage derived by the color display is the creation of greater contrast between the manual movable marker and the radar target trail. The program for the application of color to ATC displays is being continued with the evaluation of colors for displaying identification and altitude information from the ATC radar beacon system.

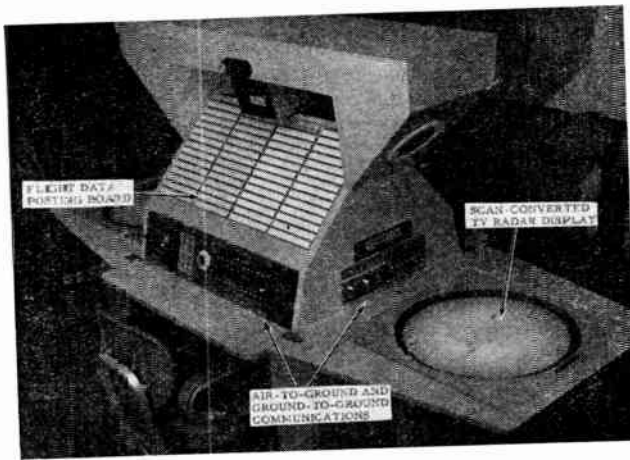


Fig. 9—22-inch Kenoscope Island display.

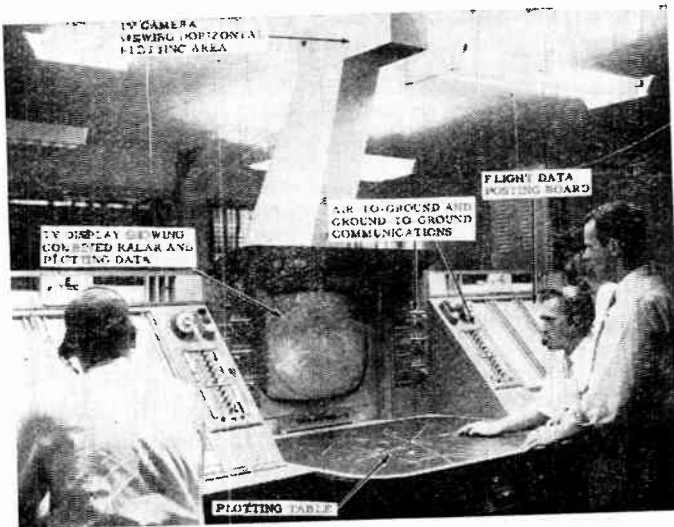


Fig. 10—SPANRAD ATC display.

The present limitations on the SPANRAD system are the horizontal and vertical resolutions which permit reading of characters no smaller than $\frac{1}{2}$ -inch on a 40-inch diameter plotting surface. The resolution of the system does not limit the use of the display when it is used as the point of control, since it is not necessary for the camera to read the information on the markers. However, this type of operation is not considered completely satisfactory. Therefore, steps are being taken to obtain higher resolution systems.

The next steps in displays are closely coupled to the development programs for the automatic processing of flight data. Techniques are being evaluated for the addition of tabular information alongside the radar target positions on the pictorial display in order to identify properly the target and to provide a more coordinated display to the controller. Various track-while-scan and computer-derived systems, together with character generating tubes and equipments are being evaluated. One of the important operational requirements which must be considered for ATC displays is the need for reliability and protection against the complete loss of information. New techniques for future displays are needed in order to meet the operational requirements. Among these re-

quirements are: high brightness, good resolution, fast scanning, accurately controlled halftones, and flexible storage characteristics. Electroluminescent and other solid state devices, together with direct-view storage tubes, appear to offer considerable promise in these areas.

AUTOMATIC DATA PROCESSING

As is true with nearly all large and complex systems, when changes are indicated, there is a tendency to ignore the immediate needs of the ATC system and the absolute necessity for an orderly transition from the present manual system to a system wherein a certain portion of the manual operations are taken over by a machine. The present manual system uses voice and teletypewriter information relayed to the controller of an area. In some cases, the information has passed through several human relays. The controller uses this information to compute time estimates for the various fixes en route and posts the information on flight strips on his manual type of symbolic display for planning and decision-making purposes. Some of the operational problems encountered with this display have been with the techniques used for gathering, processing, and distributing the flight data. It is evident that a digital computer is capable of preparing (and this includes the calculation of estimates) from 10–100 times the number of flight progress strips that average human controllers can prepare.

To confirm some of the studies made on data processing and display techniques, a study program was started in 1951, and later a series of tests was conducted in an environment of automatic tabular displays, typed flight progress strips, a magnetic storage drum, a teletypewriter network, including an automatic switching center, and a position reporting input device.⁶ Since a digital computer was not available for the project at the time of the tests, a simulated computer was used consisting of a group of highly skilled personnel in functions similar to that of an electronic computer. From this test it was concluded that immediate improvement in air traffic control could be realized from a system of encoding and processing tabular flight data. It was believed the system should include input devices to compose flight data messages in language and format compatible with data processing equipment, electronic computers to process the data and direct it to the display, programs within the computers to permit utilization of the stored data for confliction-search and flow-control routines, the capability of on-line service from the input devices to the computer, between computers, and from the computer to the display, and printing devices for the preparation of typed flight progress strips. To exploit fully the capabilities of the electronic data processors, the tabular display should be automatic.

⁶ G. E. Fenimore, "Real-time processing for CAA air traffic control operations," presented at Eastern Joint Computer Conf. Washington, D. C.; December 9–13, 1957.

As a first operational step in December, 1955, a Model 650 Digital Computer built by I.B.M. was proposed for the processing of flight plans and the preparation of fix-posting strips in the Indianapolis Air Route Traffic Control Center. Considerable time was required to prepare a detailed program for the computer, since this was the first time that machines of this type had been used for this or similar purposes. The program had to provide for an area containing over 11,000 miles of airways and for the airway and point-to-point direct route flights. (On an average day about 7000 flight progress strips must be prepared, or about 700 strips during a peak hour.) It became apparent during the initial use of this equipment that additional storage capacity was required. As a result, a Type 355 RAMAC Disk Storage device was added to the system and a new program was developed which included a simple conflict detection routine.

The use of the computer in the Center is illustrated in Fig. 11. Incoming flight plans are copied onto punched cards. The punched cards are read into the computers through an input card reader on the 407 accounting machine printer. The computer then sorts and rearranges the flight plan data for processing. Since the computer program processes either airway or direct route flights, it determines which fixes in the area the flight will pass and computes the estimated time over each of these fixes. In the case of direct flights, it provides the distance and direction from the nearest fixes that the aircraft will pass. A simple conflict detection routine compares a new flight plan with other flight plans which already have been processed, to determine if aircraft will be over the same fix at the same altitude within plus or minus ten minutes. Flight progress strips are printed out on perforated paper by the 407 printer, separated into individual strips, and inserted into strip holders. Although this separation and insertion now is a manual function, an experimental strip stuffer has been developed and tested which will eliminate this step. The strip holders then are distributed to the various posting boards in the control room via belt conveyors.

As a result of this initial program, additional computers now are being leased and installed in a number of other Centers in this country. At the present time, the Remington-Rand UNIVAC File Computers are being placed into operation in the New York and Washington Air Route Traffic Control Centers.

The existing computer equipments being used are standard off-the-shelf items. In order to permit greater utilization of automatic data processing, improvements

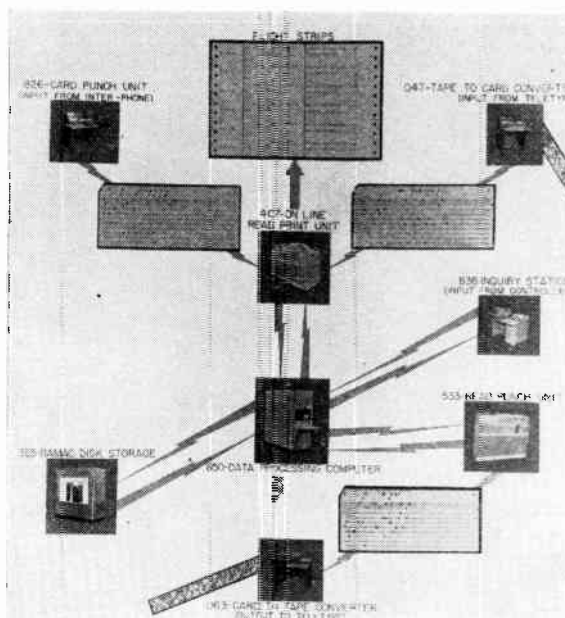


Fig. 11—Automatic flight data processing.

are required in the handling of input and output data. Developments are under way at TDC to permit direct input to the system so that the majority of flight plans will be received and read into the computer automatically without manual handling. A controller's input keyboard will permit the controller to insert flight plan changes directly from the control board. Updated flight data from the computer will be read out automatically at the control boards. Computers in adjacent centers will be interconnected in order that flight plan data will be forwarded automatically from center to center. More sophisticated conflict-detection program routines will be developed and flow control may be programmed into the computer. This work is being conducted using both the IBM-650/RAMAC and Remington-Rand UNIVAC equipments at TDC.

A more complete approach in applying computers to air traffic control is being undertaken by the Airways Modernization Board's (now the Federal Aviation Agency's) Phase II program.⁷ In this approach, a complete system including input devices, computers, and displays, is being built especially to meet all of the known operational requirements. Both the in-service and this longer range program are closely coordinated to ensure a smooth transition.

⁷ H. S. Stokes, "Applying computers to air traffic control," IRE TRANS. ON AERONAUTICAL AND NAVIGATIONAL ELECTRONICS, vol. ANE-5, pp. 152-159; September, 1958.

Pulsed Analog Computer for Simulation of Aircraft*

A. W. HERZOG†

Summary—This article describes the logical design of a pulsed analog-digital computer which could be used to solve the system of nonlinear differential equations normally encountered in the simulation of aircraft. The design calls for a magnetic drum to perform the functions of storing the program and the aerodynamic data required in the computation. Time-shared analog computing elements encompassing the normal arithmetic operations of addition, subtraction, multiplication and division are used. They are connected to a common bus through gating circuits. In addition, electronic integrators (one for each integration required) and analog-storage elements are provided. Utilizing these elements sequentially, it is possible to complete a computation cycle in each revolution of the drum. Function generation is accomplished in this computer by means of specialized circuitry which determines the position of the four surrounding break points of a given value of a function of two or more variables. The discrete values of the function at the break points are selected from digital drum storage for an interpolation procedure. A type of "floating point" system is included to scale automatically all voltages to fall within optimum levels for the operating elements. The computer appears to offer the advantages of flexibility, compactness, ease of programming, and economy commensurate with the limited accuracies required in flight-training simulation.

LOGICAL DESIGN OF A PULSED ANALOG COMPUTER

ELECTRONIC analog computation has been used for the simulation of aircraft for training since the inception of this art, except for a short period when purely mechanical means were used. In this analog computation, aerodynamic data is normally stored by linear potentiometers which are tapped and loaded to represent nonlinear aerodynamic functions. With this technique, the changing of aerodynamic functions, often necessitated by changes in the aircraft or by correction of original design data, is difficult because of the complex way in which various potentiometers contribute to the generation of these functions. In order to achieve the aim of a flexible computer, it was decided at the outset of the program¹ that this problem of function generation should be investigated. Various methods of function generation were studied. Among these were: diode function generation, storage using matrices of potentiometers, and digital storage of discrete points. Of these systems, the storage by discrete points in digital form seemed to offer the greatest possibility of fulfilling the requirements of flexibility, speed of response, and ease of set-up. When discrete values of a function of one or more variables are used to describe the function, it is necessary to extract only the pertinent data from the digital storage and to perform an interpolation to arrive at the desired function value. A method was developed

for function generation using analog computation, time-sharing arithmetic elements, for the interpolation procedure.² The system devised, however, is capable of more elaborate computation tasks than merely function generation. In fact, it is believed that a complete set of aircraft aerodynamic equations could be solved using this technique. This paper will describe a logical design of a pulsed analog system which is currently undergoing further development and study.³

The system block diagram is shown in Fig. 1. This system utilizes a magnetic drum for digital storage of aerodynamic data and program. A decoder is used to convert digital data into voltage pulses for computation. Inputs to the computer from the physical system come directly from the pilot's stick, throttle, rudder and other similar position indications. These inputs are connected to a main bus and can be sampled through gates during the computation. The computing elements are time-shared analog elements operating on 10- μ sec voltage pulses, the pulse magnitude representing the value of the function. There is a separate element for each class of arithmetic operation that is to be performed. An adder/subtractor, multiplier, divider, several storage elements and dc integrators (one for each integration to be performed) make up the arithmetic element. The gating of signals into and out of these units is under the control of the program stored on the drum. By a sequential program of gating values onto these elements and performing the arithmetic operations, a complete set of equations can be solved.

In addition to its basic pulsed analog design, this system has floating-point circuitry. With this system, all analog pulses can be kept within optimum levels and at the same time an increase in the dynamic range of the computer is realized. The system also utilizes a separate logical element to control the performance of function generation. Discrete values of the function are stored digitally on the drum. When a particular value of the function is required, the special-purpose element obtains the pertinent discrete points from the storage unit. These values are converted to voltage form and sampled by gates. The arithmetic element performs the interpolation by operating on the sampled voltage pulses. Data-access time is not a problem since the data is stored on the drum immediately following the program command to operate on the data. A complete computa-

* Original manuscript received by the IRE, December 8, 1958; revised manuscript received, February 16, 1959.

† U. S. Naval Training Device Center, Port Washington, N. Y.

¹ Conducted by the Servomechanisms Lab., M.I.T., Cambridge, Mass., under contract with the U. S. Naval Training Device Center, Port Washington, N. Y.

² C. Holmquest, "Data Storage, Level Sensing and Switching in an Analog-Digital Function Generator," Servomechanisms Lab., M.I.T., Cambridge, Mass.; August, 1957.

³ R. Lee, "Logical Design of a High Speed Analog Digital Computer for Simulation," Servomechanisms Lab., M.I.T., Cambridge, Mass.; May, 1958.

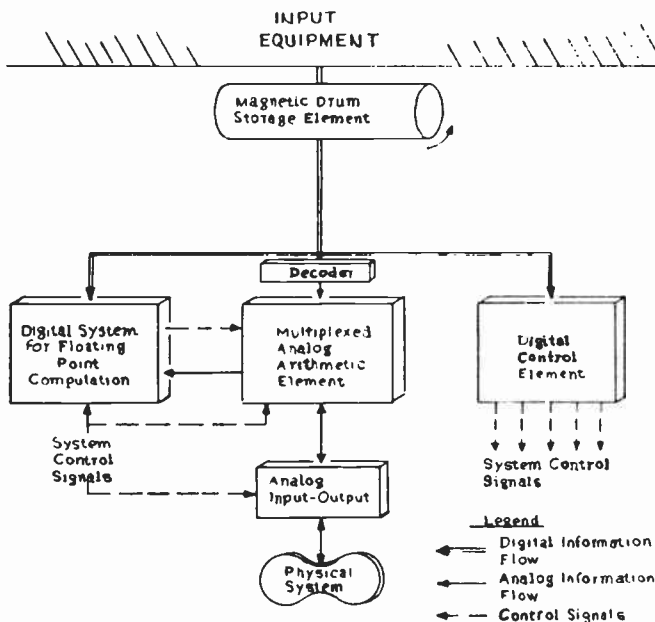


Fig. 1—System block diagram.

tion of the equations of motion of an aircraft can be accomplished in each revolution of the drum. With additional computations, such as aircraft-system operation, engine equations, and various special flight conditions (landing, take-off, etc.), it might become necessary to use several blocks of channels on the drum and to have various arithmetic units operating in parallel.

ARITHMETIC ELEMENT

The analog arithmetic element is shown in Fig. 2. The computation is under the control of digitally-stored instructions which activate a control element. In the case of multiplication, one program cycle is required to gate one of the operands to a multiplier-storage element (upper right on Fig. 2). In the next program cycle, the other operand would be gated to the other storage element. The unscaled product is available soon after both inputs have been gated to the multiplier inputs. A total of three program cycles (each of 10- μ sec duration) is required for multiplication and scaling. The adder/subtractor requires additional circuitry to compare the exponents of the operands since a floating-point system is used. The scale factor of the smaller operand must be adjusted to coincide with that of the larger operand. This is accomplished by appropriate attenuation of the smaller operand with a passive element. The digital computation on scale factors, to be discussed later, is performed at a high enough rate so that the analog addition is not slowed down while the scale factors are computed.

The adder/subtractor performs subtraction by passing one of the operands through an inverting amplifier. Normally only positive voltages are handled by the computer. A sign bit (in the accompanying digital-scale factor) indicates whether or not the operand must pass through the inverting amplifier. The output of the adder

passes through an absolute-value device to detect and correct any negative voltages. A positive voltage will always result from an addition or subtraction, and the sign bit will designate the polarity of the resultant.

Integration is performed in the computer by means of separate integrators. The voltage pulse representing the quantity to be integrated is gated to the storage element preceding the integrator, and the integration is performed. The value is corrected on every solution cycle of the computer. The output of the integrator is available continually and can be sampled through the output gate for use in the computation. In most cases, the integrator outputs are also computer outputs. These outputs are available directly from the integrator and present a relatively smooth function.

A typical system of equations programmed for this computer consists of a set of simplified aerodynamic equations. In this computation, eight integrators are required. In addition, eight temporary storage elements are provided for holding intermediate values that are needed for further computation.

INPUTS

As stated previously, analog inputs can be gated directly onto the main bus and sampled as required in the computation. Three gates provide aileron deflection, elevator deflection, and rudder deflection.

FUNCTION GENERATION

Special digital logic is used for the selection of desired values of the functions of two or more variables. This circuitry controls the flow of the digitally-stored data to the arithmetic unit via the decoder. The arithmetic element then compares the analog voltages obtained from the decoder with the analog values obtained in the computation in a manner described below, to provide the specific desired value of the function to the computation.

Fig. 3 shows a typical function of two independent variables. In a general case, shown in Fig. 4, the function Z must be obtained by interpolation on the four surrounding stored discrete values of the function $Z_{j,i}$, $Z_{j+1,i}$, $Z_{j,i+1}$, and $Z_{j+1,i+1}$.

The position of the point Z on the curve is known in terms of the values of the two independent variables, x and y , for this point. The function-generation circuitry provides each stored discrete value of x to the arithmetic element. Fig. 5 shows the method of storing the data on the drum. It can be seen that the values of the eight discrete points chosen to tabulate the independent variable x pass the read heads first. The arithmetic element compares each stored value with the given value of the independent variable arrived at in the computation. The values are compared in turn until a value of x just exceeds the given x . The position (address) of this value is retained in a three-position digital counter. This value is $i+1$. From this counter we can also get the address of the value just below the given x , called i .

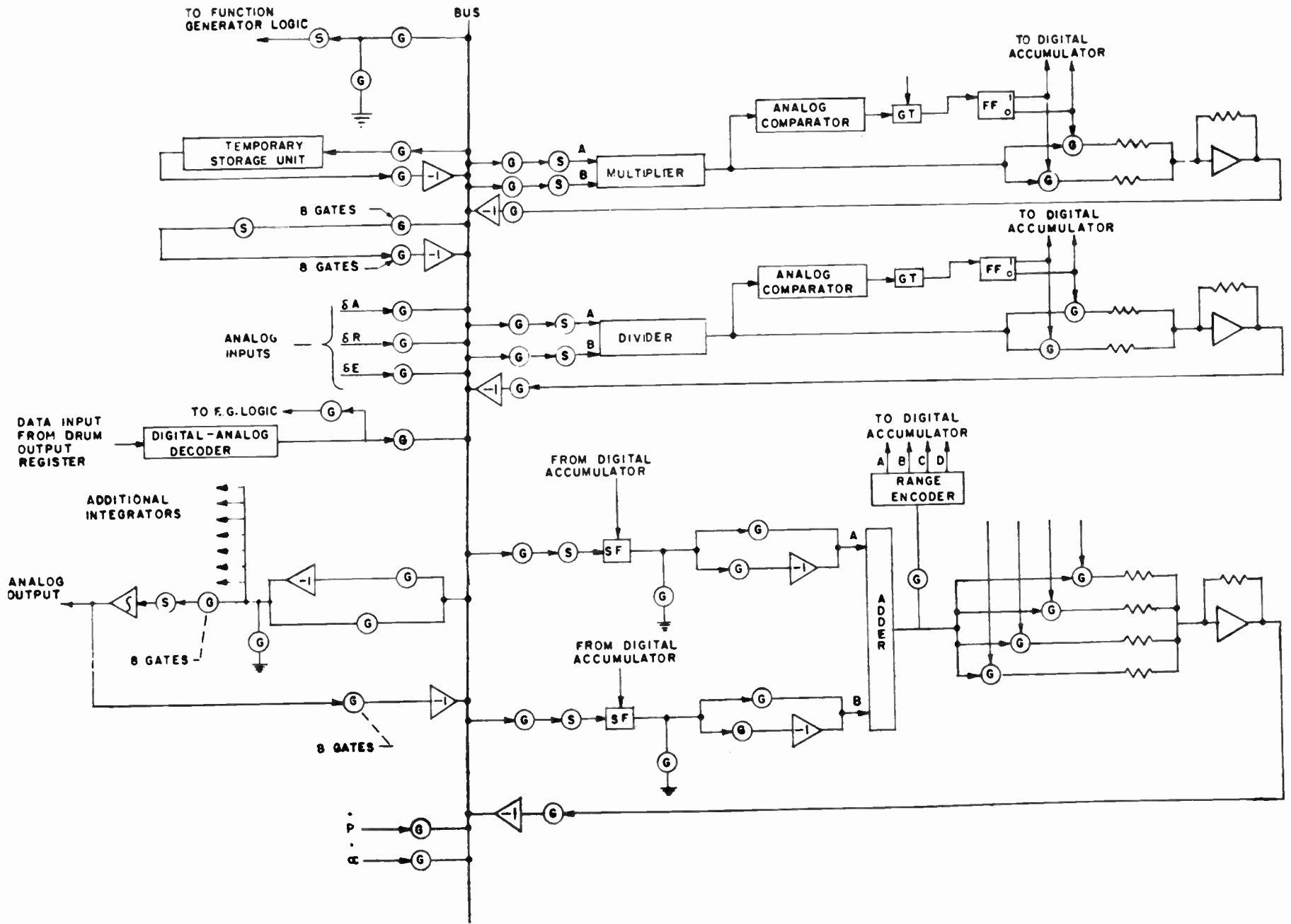


Fig. 2.—Analog arithmetic element.

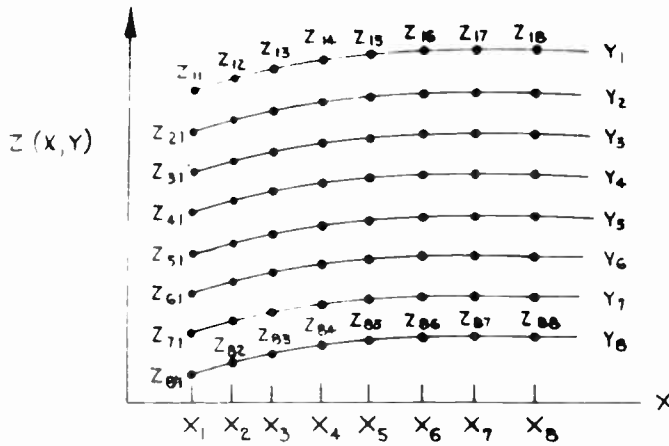


Fig. 3—Typical function of two variables.

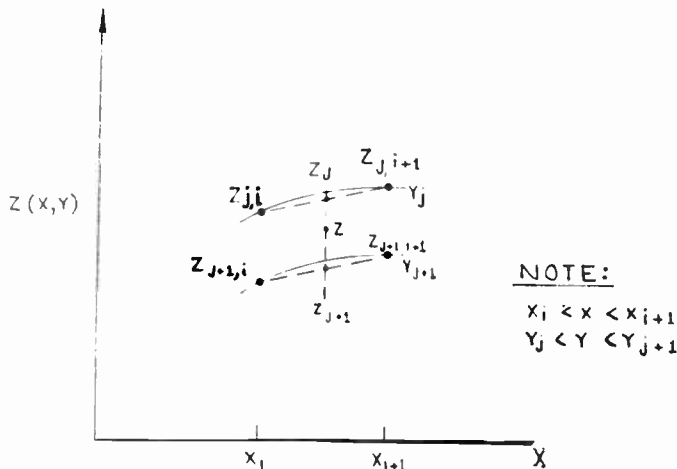


Fig. 4—Linear interpolation for generating a function of two variables.

In a similar manner the values j and $j+1$ are obtained by step-by-step comparison of the given value of the independent variable y with each discrete value of y stored digitally on the drum. This is stored as the second block of data on the drum shown in Fig. 5.

We then have available in this special digital storage the values $i, i+1, j,$ and $j+1$. Eight additional data blocks are stored on the drum. The first to pass the read heads in sequence would be all the discrete values of the dependent variable Z from $Z_{1,1}$ to $Z_{8,1}$. Then the values $Z_{1,2}$ to $Z_{8,2}$ and so forth to $Z_{1,8}$ to $Z_{8,8}$. The function-generation circuitry, by straightforward counting, provides the values of the dependent variable at the four breakpoints. The arithmetic element utilizes the converted analog-voltage pulses to obtain the value of the function Z at the required point using the following interpolation formulas:

$$Z_{j+1} = Z_{j+1,i} + (x - x_i) \frac{Z_{j+1,i+1} - Z_{j+1,i}}{x_{i+1} - x_i}$$

$$Z_j = Z_{j,i} + (x - x_i) \frac{Z_{j,i+1} - Z_{j,i}}{x_{i+1} - x_i}$$

$$Z = Z_j + (y - y_j) \frac{Z_{j+1} - Z_j}{y_{j+1} - y_j}$$

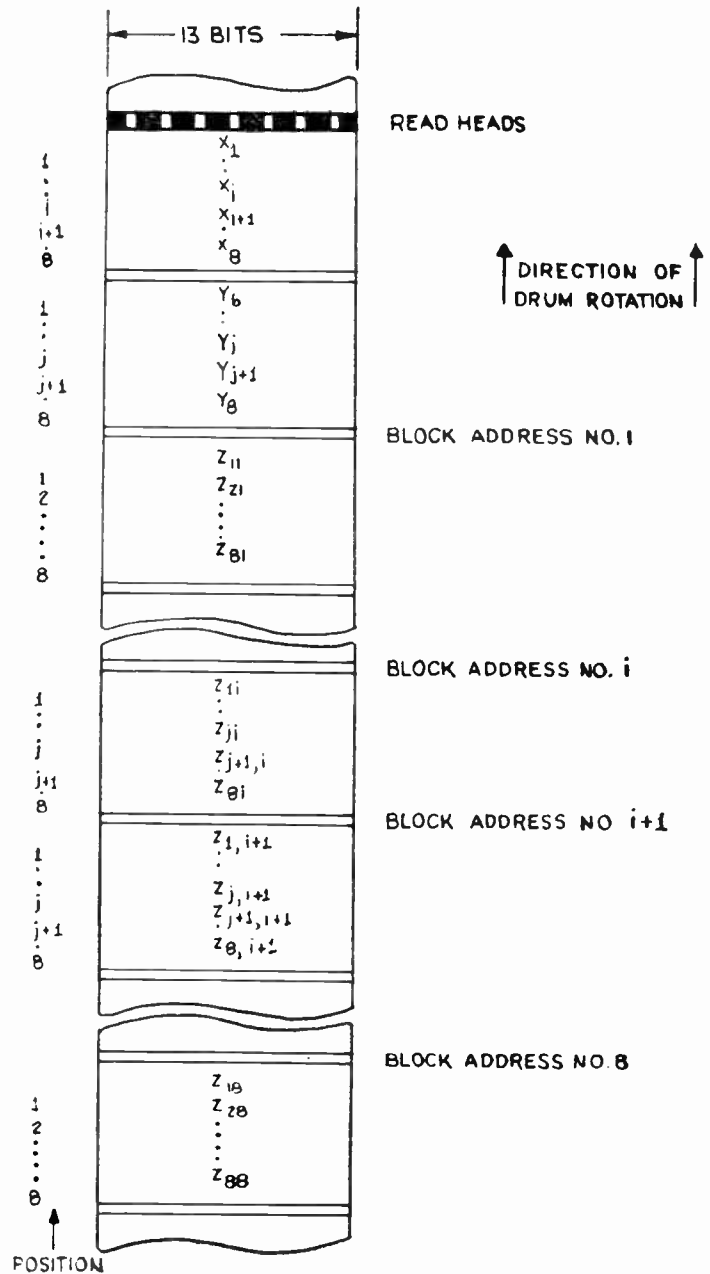


Fig. 5—Data storage on magnetic-drum surface for function generation.

FLOATING POINT

The floating-point system automatically scales the analog-machine voltage to an optimum range. The digital exponent and sign associated with each machine variable must be stored by the digital floating-point system. Instructions for the floating-point system are stored on the magnetic drum. A small-core memory is used to store the exponent values during the computation. A digital accumulator is required to operate on the exponent values. This accumulator can be kept small since the words to be operated on are to be three or four bits in length. The pulse-repetition frequency is approximately one mc, and a system of dc logic is suitable.

The digital system designed for this computer operates in a synchronous manner. Information is trans-

ferred over a single bus "set" in binary form. The main bus consists of five lines, four for the exponent and one for the sign. It should be remembered that the operations being performed on this digital system are occurring at the same time that operations are being performed on the analog-machine voltages. Therefore, although there is no direct connection between the analog and digital equipment, there is a flow of information exchanged between the systems.

COMPONENTS

The development of components to satisfy the requirements of the logical design described in this article are being completed. A variety of components have been breadboarded and tested. The requirements for the system indicate that rise times in the order of fractions of microseconds, and operating speeds of the order of 10 μ sec, are necessary. One of the elements developed is a quarter-square multiplier using solid-state components.⁴

⁴ H. Dumas, "A Pulsed Analog Multiplier-Divider," Servomechanisms Lab., M.I.T., Cambridge, Mass.; May, 1958.

The storage element and its associated gate is being developed. This element will charge a capacitor in approximately 10 μ sec, maintain this charge to an accuracy of ± 1 per cent for 20,000 μ sec, and discharge the capacitor in less than 10 μ sec. Such an element is under study using four-layer diodes, silicon diodes, and thyrite as components.

A simple computation involving addition, multiplication, several integrations, inversion, data storage, and data transfer will be activated using the breadboarded components in a fixed program. The equations to be used represent a set of equations for the longitudinal motion of an aircraft. Data will be inserted by potentiometers for this test, and a recording of the output will be made on a graph recorder.

The program will be run repetitively to determine the accuracy of the computation and the deterioration of computer values (if any) over a long time period. An over-all accuracy criterion of ± 5 per cent will be considered to be satisfactory for the computation of aircraft equations of motion for the training situation.

Determination of Satellite Orbits from Radio Tracking Data*

I. HARRIS†, R. JASTROW† AND W. F. CAHILL‡

Summary—A computer program has been developed which permits an approximate determination of a satellite orbit from a minimal amount of tracking data.

INTRODUCTION

THE program described here is designed for the rapid production of an approximate satellite ephemeris from a small number of radio tracking observations. The input data consist of 1) a subsatellite fix, *i.e.*, latitude and longitude at a given time; 2) initial approximations to the orbital inclination, planar orbit elements and drag coefficient; 3) three or more radio observations. The observations may be in any of the forms normally obtained from tracking stations, such as latitude or longitude crossing times for interferometer

stations, or bearing, elevation, and slant range for radar stations. A program prepared for the IBM 704 computer first integrates the equations of motion with atmospheric drag included, using initial approximations to the orbit elements, and computes the position of the satellite relative to each of the observing stations. The program then corrects the orbit elements and drag coefficient to give agreement with the original observations.

Our procedure differs to some degree from other orbit programs in its emphasis on the use of a minimal amount of data. It also makes use of certain dynamical approximations in order to shorten the computations. In spite of these limitations, the program has been found to yield results of surprising accuracy in its application to the orbit of Satellite 1958 Beta.

OUTLINE OF THE CALCULATION

We simplify the numerical integration of the equations of motion by assuming initially that the mass distribution of the Earth is spherically symmetric and

* Original manuscript received by the IRE, December 23, 1958; revised manuscript received, March 2, 1959. A report on this program was presented at the Vth CSAGI Assembly in Moscow (1958).

† Theoretical Div., Natl. Aero. and Space Admin., Washington, D. C. Formerly with U. S. Naval Res. Lab., Washington, D. C.

‡ Theoretical Div., Natl. Aero. and Space Admin., Washington, D. C. Formerly with Natl. Bur. Stand., Washington, D. C.

the orbit therefore confined to a plane. The mass distribution actually contains a quadrupole term describing the bulge at the equator, which causes the plane of the orbit to precess about the axis of rotation of the Earth, and the line of apsides to rotate in the orbital plane. Both effects, the precession of the plane and the rotation of the line of apsides, are included in the program phenomenologically, as modifications of the transformation from coordinates in the orbital plane to coordinates fixed in and rotating with the Earth. The equations of the complete transformation are given below in this section.

Calculations based on probable values of the equatorial bulge indicate that the precessional period is of the order of months vs 2 hours for the period of revolution of typical satellite orbits. Thus our description of the orbit is a good approximation to the exact motion.

For the application of the program, zero order approximations to the orbital inclination and planar orbit elements—period, perigee argument, and perigee radius—must first be deduced from tracking data. A satellite fix, *i.e.*, subsatellite latitude and longitude at a known time t_0 , must also be given. Approximate values for the planar elements may be obtained directly from interferometer data by hand calculations, using the method described below.

Initial Determination of the Orbit Elements from Interferometer Data

Let D_y and D_x be the separations of the interferometer antennas on the E-W and N-S baselines, respectively, measured in wavelengths. Let n_y be the phase difference, in wavelengths, between the antennas on the E-W baseline when the subsatellite crosses that baseline, and let n_x be the corresponding quantity for the N-S baseline. Then the angle between the subsatellite path and the E-W baseline is given to order h/R_e by

$$\tan i = \sqrt{\frac{D_x^2/n_x^2 - 1 + h/R_e}{D_y^2/n_y^2 - 1 + h/R_e}}, \quad (1)$$

where h is the mean altitude of the satellite during the pass, and R_e is the radius of the Earth. The altitude is given by the formulas

$$h = v\Delta t \cos i \sqrt{D_x^2/n_x^2 - 1 + h/R_e}, \quad (2a)$$

or

$$h = v\Delta t \sin i \sqrt{D_y^2/n_y^2 - 1 + h/R_e}, \quad (2b)$$

in which v is the mean velocity of the satellite during the pass, and Δt is the time difference between the subsatellite crossings of the E-W and N-S baselines.

These equations are conveniently solved by an iterative procedure. Let $i^{(0)}$ and $h^{(0)}$ be the zeroth order approximations to i and h , obtained by setting $h/R = 0$ on

the rhs in (1) and (2), and replacing v in (2) by its mean value. If three independent Minitrack observations are available,¹ the three values of $h^{(0)}$ deduced from these now determine first order approximations to the orbit elements by insertion in the ellipse equation

$$R_e + h = \frac{\bar{R}(1 - e^2)}{1 + e \cos(\gamma - \theta)}. \quad (3)$$

in which \bar{R} , e , and θ are, respectively, the mean radius, eccentricity, and argument of perigee. The velocity at each station may now be calculated from the orbit elements by the formula,

$$v = \bar{v} \sqrt{\frac{1 + e^2 + 2e \cos(\gamma - \theta)}{1 - e^2}} \quad (4)$$

and substituted on the rhs in (2) together with $h^{(0)}$, to give improved values of h .

During one revolution, the velocity departs from its mean value by a fractional amount, e , hence the maximum error in $h^{(0)}$ is also of order e . After the first iteration on (1)–(3), the altitudes and orbit elements are accurate to order e^2 , or 1 per cent for typical satellite orbits.

Equations of Motion

The equations which determine the orbit are

$$\begin{aligned} \ddot{X} &= -MGX(X^2 + Y^2)^{-3/2} - \frac{1}{2} C_D \rho \frac{A}{m} \dot{X}(\dot{X}^2 + \dot{Y}^2)^{1/2} \\ \ddot{Y} &= -MGY(X^2 + Y^2)^{-3/2} - \frac{1}{2} C_D \rho \frac{A}{m} \dot{Y}(\dot{X}^2 + \dot{Y}^2)^{1/2}, \quad (5) \end{aligned}$$

in which G denotes the gravitational constant, M and m the masses of Earth and satellite, respectively, A the effective cross-sectional area of the satellite, C_D the drag coefficient, and ρ the atmospheric density.² The density and scale height of the model atmosphere used in (5) are adjustable at all altitudes and form a part of the input data. Eq. (5) is integrated numerically in the program.

Transformation to Earth Coordinates

We transform from XY to a coordinate system $X'Y'$ in which the X' axis is located on the line of nodes [Fig. 1(a)]. Let γ_0 be the angle between the X and X' axes at time t_0 , and let ω_a be the angular velocity of the line of apsides. Then $\gamma = \gamma_0 + (t - t_0)\omega_a$ is the angle between

¹ The observations are preferably taken at stations distributed around the orbit, but under favorable circumstances a single station will produce sufficient data, *viz.*, the altitudes of the ascending and descending passes, and the orbital period obtained from the times of successive latitude crossings at the station.

² The rotation of the earth's atmosphere produces a drag component normal to the plane of the orbit. The magnitude of this component is ≤ 7 per cent of the total drag, and it has, therefore, been neglected in (5).

the X axis and the line of nodes at time t , and we have

$$\begin{aligned} X' &= X \cos \gamma - Y \sin \gamma, \\ Y' &= X \sin \gamma + Y \cos \gamma. \end{aligned} \tag{6}$$

From the coordinates $(X'Y')$ of the satellite point, we obtain coordinates (xyz) with respect to a cartesian system fixed in the Earth, by the equations

$$\begin{aligned} x &= X' \cos (\omega - \omega_n)t + Y' \sin (\omega - \omega_n)t \sin \alpha, \\ y &= -X' \sin (\omega - \omega_n)t + Y' \cos (\omega - \omega_n)t \sin \alpha, \\ z &= Y' \cos \alpha, \end{aligned} \tag{7}$$

where ω and ω_n are the angular velocities of the Earth and the orbital plane, respectively, and α is the orbital inclination with respect to the pole [Fig. 1(b)]. It should be noted that in this system the z axis coincides with the axis of rotation of the Earth, and the x axis with the initial position of the line of ascending nodes.

longitude (λ_0) at time t_0 ; 2) the orbital inclination (α); and 3) the planar elements, *i.e.*, period (P), eccentricity (e) and perigee argument (ψ). Pertinent quantities, *e.g.*, latitude and longitude crossing times, are obtained by linear interpolation from the results of the numerical integration.

CORRECTION PROGRAM

The remainder of the program applies a linear differential correction scheme to the orbit elements and drag coefficient. The program accepts three observations and corrects a corresponding number of orbit elements. There is no provision for a redundancy of data.

We denote by $a_1, a_2,$ and a_3 the orbit parameters which are to be corrected, and by $c_1, c_2,$ and c_3 the triplet of observed quantities which are to be used as the basis for the correction. The triplet c_i may be any combination of altitudes or times of meridian transit, not necessarily at one station nor in one pass.

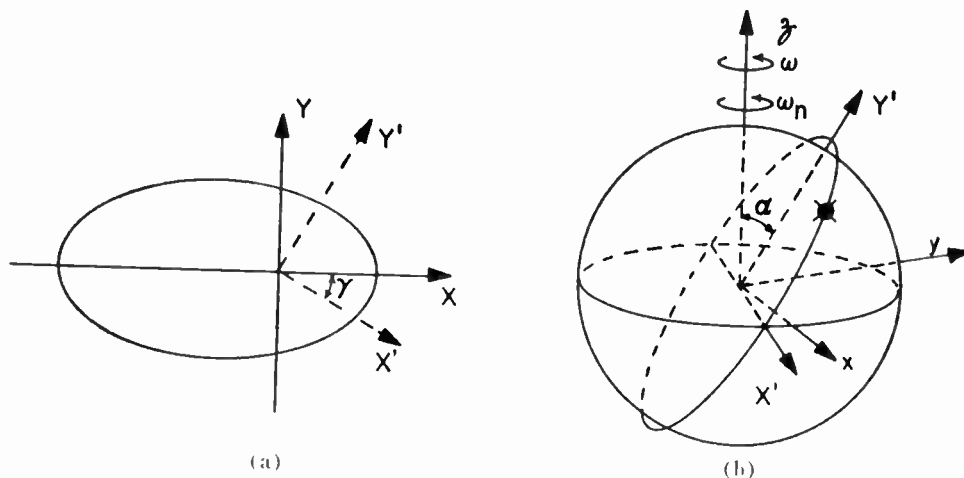


Fig. 1.

Finally, we transform the cartesian coordinates of the satellite to latitude (ϕ), longitude (λ), and altitude (h):

$$\begin{aligned} r &= (x^2 + y^2 + z^2)^{1/2}, \\ h &= r - R_m - \delta \cos 2\phi, \\ \phi &= \sin^{-1} (z/r), \\ \lambda &= \tan^{-1} y/x + L, \end{aligned} \tag{8}$$

in which R_m is the mean of the equatorial and polar radii, δ is the difference between these radii, and L is the longitude of the line of ascending nodes at time t_0 . L is taken positive or negative when the line of nodes lies initially east or west, respectively, from Greenwich.

The angles θ and L are calculated from input data which specify: 1) the subsatellite latitude (ϕ_0) and

Let $a_i^{(0)}$ be the initial values assumed for the planar orbit elements, and let $c_i^{(0)}$ be the values of the quantities c_i calculated from the ephemeris program for that assignment of orbit elements. Then the corrections Δa_i required to reproduce the observed values c_i are given, within the linear approximation, by the solutions to the equations:

$$c_i - c_i^{(0)} = \frac{\partial c_i}{\partial a_1} \Delta a_1 + \frac{\partial c_i}{\partial a_2} \Delta a_2 + \frac{\partial c_i}{\partial a_3} \Delta a_3; \quad i = 1, 2, 3. \tag{9}$$

The derivatives $\partial c_i / \partial a_j$ are approximated by difference quotients $\Delta c_i / \Delta a_j$, obtained by numerical variation in the ephemeris program. A single code produces the corrections Δa_j from (9), using c_i and $a_i^{(0)}$ as input data.

The determination of Δa_i requires 20 minutes of computing time on the IBM 704, for a representative case in which the data c_i are spread over an interval of 24 hours in the satellite history.

A separate program is available which will utilize radar observations, *i.e.*, bearing, elevation and slant range, for the correction procedure.

Application to Satellite 1958 Beta

Satellite 1958 Beta (Vanguard I) has a stable orbit which provides an opportunity for testing the accuracy of the correction procedure. As input data we use the standard lists of observed meridian crossing times for Minitrack stations. These lists, which are distributed by the Vanguard Control Center of the National Aeronautics and Space Administration, represent a selected list of overhead passes near the stations in the Minitrack network. We have arbitrarily chosen several trios of observations for the initial orbit elements, and used these observations to correct our initial assumptions for the period, perigee radius, and perigee argument. Two iterations were found to suffice for the convergence of the correction program, with further iterations producing no significant improvement.

The results of the correction program are shown in Table I for several runs representing different choices of data in the same time interval. The dispersion in the elements is seen to be 0.003 minute in period, 0.2° in perigee argument, and 9 miles in perigee radius.

TABLE I
DEPENDENCE OF ORBIT ELEMENTS ON THE DATA USED IN THE CORRECTION PROGRAM, FOR SATELLITE 1958 BETA†

Run	1	2	3
Period (minutes)	134.260	134.262	134.259
Perigee Radius (miles)	4367.12	4370.01	4376.22
Perigee Argument (degrees)	231.40	231.52	231.59

† Runs 1 and 2 are based on distinct sets of data selected from the same 8-hour interval. Run 3 corresponds to data covering a 24-hour interval.

Table II compares the observed and computed crossing times for the stations in the Minitrack network, using the orbit elements obtained in runs 2 and 3, Table I. These runs are based on input data covering intervals of 8 and 24 hours, respectively. By inspection of Table II, we see that the differences between the observed and computed times increase at the rates of approximately 3 sec/day in run 2, and 5 sec/day in run 3, indicating that the orbital periods obtained in runs 2 and 3 in Table I are in error by 0.005 minute and 0.01 minute, respectively.

TABLE II

SATELLITE 1958 BETA: COMPARISON OF OBSERVED AND COMPUTED TIMES OF MERIDIAN CROSSING FOR STATIONS IN THE MINITRACK SYSTEM, USING THE ORBIT ELEMENTS LISTED IN TABLE I†

Station	Observed Meridian Crossing Time	Calculated—Observed	
		Run 2	Run 3
Fort Stewart	June 07 ^d 11 ^h 22 ^m 41 ^s *	- 1*	
San Diego	07 13 36 26	0*	0
Fort Stewart	07 13 44 47	0	0*
San Diego	07 15 58 35	0	- 1
Lima	07 18 34 59	0*	+ 2
Santiago	07 21 05 35	-11	- 1*
San Diego	08 11 51 56	- 3	- 5
Fort Stewart	08 12 01 30	- 3	- 5
Fort Stewart	08 14 23 42	- 2	- 6
Antofagasta	08 19 18 18	- 3	- 1*
Antofagasta	09 00 14 34	- 8	+ 5
San Diego	09 12 31 21	- 6	-11
Fort Stewart	09 12 40 17	- 5	-11
San Diego	09 14 54 04	- 5	-11
Quito	09 17 28 37	- 4	-10
Antofagasta	09 22 25 48	0	+ 9
Havana	09 07 51 13		-11
Antigua	10 06 12 08	- 9	-15
Havana	10 08 31 51	-10	-16
San Diego	10 10 45 45	-10	-16
Fort Stewart	10 10 56 42	- 9	-16
Fort Stewart	10 13 19 08	- 8	-16
San Diego	10 15 32 55	- 8	-17
Antofagasta	11 01 37 26	-17	
Havana	11 09 11 45	-12	-20
Fort Stewart	11 09 12 05	-12	-21
Fort Stewart	11 11 35 38	-14	-24
San Diego	11 13 49 20	-11	-21
Antigua	11 14 03 40	-10	-22
Antofagasta	11 18 52 19	-12	-18
Havana	12 07 24 42	-14	-24
Fort Stewart	12 09 51 28	-15	-26
Fort Stewart	12 12 14 27	-13	-26
San Deigo	12 14 28 13	-13	-27
Havana	12 14 37 12	-12	-26
Antofagasta	12 19 32 18	-16	-22
Antofagasta	13 00 27 42	-23	-21

† The last two columns list the differences between the observations and the times computed with the orbit elements of runs 2 and 3. The asterisks indicate observations used as input data for the correction program. In both runs, the starting position was obtained from a pass over Antigua at 07^d 06^h 38^m 26^s. Observed times are rounded to the nearest second.

ACKNOWLEDGMENT

The authors wish to thank J. T. Mengel and the members of the Tracking and Guidance Branch of Project Vanguard for many helpful conversations on the interpretation of Minitrack data; and Dr. J. A. O'Keefe of the Army Map Service, Prof. P. Herget of the University of Cincinnati, and Drs. J. W. Siry and R. H. Wilson of Project Vanguard for a number of informative discussions on orbit theory. We also wish to express our sincere appreciation to J. H. Wegstein of the Mathematics Division of the National Bureau of Standards for his collaboration in the early phases of the IBM 704 program, and to C. L. Wade for his assistance in the execution of the computations.

Phenomena of Scintillation Noise in Radar-Tracking Systems*

J. H. DUNN†, D. D. HOWARD†, SENIOR MEMBER, IRE, AND A. M. KING†, MEMBER, IRE

Summary—In 1947, an investigation was begun at the U. S. Naval Research Laboratory to study noise in tracking radars. Of particular interest was the noise which was found to be caused by a target when it is of finite size and complex shape; this suggested a new basis for optimum radar design. With the increasing emphasis on precision in radar guidance in the space and missile age, complete instrumentation was prepared to permit investigation of target-scintillation or target-noise phenomena. While extensive theoretical studies were being carried out, the instrumentation was built to include a unique simulator of multiple, finite-size targets with a complete closed-loop-tracking system, an instrumentation radar composed of a dual-channel-tracking radar for measurement of components of target noise under actual tracking conditions, and equipment designed to provide automatic statistical analysis of noise data. The results of these studies showed several ways in which different components of the target noise affect the radar-tracking performance and how these components determine the choice of tracking systems and optimum design of the system. The studies revealed phenomena such as the two-reflector target, which can cause tracking errors of many target spans outside the physical extent, and a new concept of target noise in terms of a tilting of the phase front of the echo signal from a target.

INTRODUCTION

A USEFUL understanding of complex physical phenomena is generally found to lag considerably behind the observation of their effects. In the case of radar-target scintillation, the first observations and the beginning of an understanding helpful to the radar engineer were nearly ten years apart, and the incomplete concepts of target behavior which prevailed during this period led to an unfortunate amount of fruitless effort on the part of systems designers.

At the time the target-noise or target-scintillation studies were initiated at the U. S. Naval Research Laboratory in 1947, the tracking accuracy of fire-control radar systems was comparable to the dispersion of projectiles fired from conventional weapons. For this reason, very little attention had been given to the determination of the source or elimination of such errors. The advent of the guided missile has made the consideration of these errors much more important, because of the increased range and the controlled flight path. The high speeds of targets and the resulting shortening of smoothing time and computer time for prediction has also made it necessary to obtain the most noise-free data possible. Future radars to meet the challenge of the space age will be required to determine the absolute position of vehicles in space with information obtained in a short

time interval and will therefore be required to limit data smoothing as much as possible.

BACKGROUND

The term target scintillation as used in this paper refers to the whole class of wave-interference phenomena which occur whenever the physical dimensions of an illuminated object of complex shape represent a number of wavelengths of the radiation in use. In the case of fire-control tracking radars operating in the kilo-mega-cycle region, target spans may be hundreds of wavelengths, and the interference patterns will be extremely complex and will be changed appreciably by very small rotations of the target in space. Search-radar sets operating at lower frequencies will generate somewhat less complex patterns, but scintillation of some sort will exist over the entire spectrum which has been found useful in the radar art. It follows immediately from classical wave theory that the amplitudes of radar echoes from moving targets of practical interest such as aircraft will fluctuate widely as target aspect changes. In general, there will be rapid fluctuations caused by pitching and yawing of the airframe about the line of flight, and slow variations caused by changes in the average aspect of the airframe relative to the radar. Target-fading phenomena, observed during some of the earliest prewar experiments with search radar equipment, were found to decrease the effectiveness of the search sets appreciably by introducing an element of uncertainty into the maximum ranges of detection. The effects were, however, felt to be well understood, and it was possible to allow, at least approximately, for the loss of effective range in the original design of a system.

The effects of target scintillation on the performance of tracking radars were also observed during prewar development work and for some time were considered to be easily explained as a direct result of amplitude fading. Angle-tracking information was derived in all prewar and wartime tracking radars by a system of mechanical scanning, or lobe switching, which shifted the position of the antenna beam about the tracking axis by a small angle, generally half the two-way half-power beamwidth. The echoes from off-axis targets were thus amplitude modulated at the lobing rate, and the magnitude and phase of this modulation indicated the magnitude and direction of the pointing error. Mechanical design problems led to the choice of 30 cps as the maximum practical lobing rate in large fire-control radars, with smaller systems using rates up to 100 cps. It was evident

* Original manuscript received by the IRE, November 27, 1958; revised manuscript received, February 13, 1959.

† U. S. Naval Res. Lab., Washington, D. C.

that any amplitude modulation of the target echo other than that due to lobing could cause pointing errors if frequency components existed at or near the lobing rate. Examination of the pulse-to-pulse fading caused by target scintillation showed that such components did exist. In the case of the periodic scintillation caused by the propellers of piston-engined aircraft, the relationship between the periodic fluctuation of the echo and the false angle-error information produced by the radar was very clear, and the pointing errors produced were often large enough to cause the beam of an automatic tracking radar to be deflected completely off the target. Smaller random variations in the tracking data were also observed under virtually all conditions, and war-time workers in the radar field assumed that this tracking jitter or noise was a result of random fading of the target echo, which in turn was caused by scintillation effects as the airframe was flexed, rotated, or vibrated during flight. The dramatic illustration provided by propeller-modulation effects undoubtedly strengthened this belief, and in the absence of any contrary experimental evidence it was generally concluded that the translation of amplitude scintillations into pointing errors through the use of lobing was the sole important external cause of tracking noise. It was clear, in any case, that scintillation phenomena were currently setting a limit on the short- and medium-range performance of precision tracking radars and that no over-all gains could be expected from servo-system and receiver refinement until target effects were reduced.

Because the importance of target scintillation was well recognized, tracking-radar work at NRL in the late war years and early post-war period was very largely concerned with systems designed to minimize the effects of this scintillation on tracking accuracy. Since random fading of echo amplitude was caused by mechanical motion of the target, it seemed likely that the spectrum of these variations would show decreasing power density with increasing frequency and that a sufficient increase in lobing rate would therefore reduce external tracking noise, while a system that could function without any form of sequential lobing would eliminate the effect entirely, leaving only servo noise and receiver noise to limit tracking performance. With these considerations in mind, two experimental radar systems were devised, a high-speed-lobing radar in 1943,¹ and an amplitude-comparison simultaneous-lobing system² in 1944. The high-speed-lobing system was based on the use of a multiple-feed antenna and electronic-lobe switching with a PRF of 1000 to 4000 cps and a lobing rate of one-quarter of the PRF (one pulse transmitted on each lobe).

Both the high-speed-lobing radar and the monopulse radar were constructed in experimental form and evalu-

ated during 1946 and 1947 at NRL by measuring tracking performance against aircraft targets. The performance of both systems proved to be excellent by current engineering standards but disappointing to the experimenters, because it was not significantly better than an equally well-designed conventional conical scanning system. The results of this work indicated that under the conditions of the tracking tests, the tracking noise generated by echo-amplitude fluctuation made a relatively small contribution to the total noise; this suggested that there might be another mechanism by which scintillation could affect tracking performance.

In order to resolve the exact role of target scintillation in radar tracking, theoretical studies of the detailed nature of target scintillation were started. The target model used in this work was a rigid rod having point reflectors distributed along its length. This rod target was assumed to have a mean position normal to the illuminating beam of the radar and to generate a scintillating echo by oscillating in a random fashion about an axis perpendicular to both rod and tracking axis. An analysis based on this model indicated that there was indeed another mechanism by which scintillation could cause tracking errors. This effect, termed angle noise, is due to a distortion or tilt of the reflected phase front at the target which changed the apparent angle of arrival at the radar antenna. Theory predicted that the magnitude of the angle-noise component was proportional to the angle subtended by the target and was thus inversely proportional to range, and it was further predicted that the apparent location of the target could be shifted outside the physical span of the reflecting array. A preliminary experimental check of the target-noise theory was made by setting up a radar target consisting of just two movable reflectors. The initial experiment consisted of two corner reflectors at approximately 2000-yard range and sufficiently close together so that they were not resolved by the radar. The relative phase of the echo signals from the reflectors was controlled by moving one reflector through a small distance toward or away from the radar. When the experimental monopulse radar was used to indicate the bearing of this pair of reflectors, it was found that pointing errors amounting to several times the physical separation could be introduced by shifting one reflector in range relative to the other.

Following the work with reflectors, a carefully controlled series of tracking experiments was carried out on aircraft targets. The internal noise of the tracking system was measured, and the contribution of amplitude scintillation was calculated from amplitude-fading spectra of the targets. Independent measurements of angle noise were made with a dual-gate-monopulse radar in which a beacon return was used to allow automatic tracking, free of target noise. The second gate was used to track the target skin echo, and pulse-by-pulse subtraction of the outputs of the corresponding angle error detectors gave a continuous angle-noise function which could be recorded for analysis. The results of this work

¹O. F. Foing, Jr. and P. J. Allen, "An Electronically-Lobed Tracking and Guidance Radar," U. S. Naval Res. Lab., Washington, D. C., Rep. No. 3116; June, 1947.

²R. M. Page, "Monopulse radar," 1955 IRE CONVENTION RECORD, pt. 8, pp. 132-134.

showed good agreement with theory, considering the difficulty of obtaining accurate target parameters for the various aircraft tracking runs, and a formal report was published in November, 1950.³ The findings showed that angle noise was by far the largest component of tracking noise at short ranges, with tracking jitter falling to a minimum set by amplitude scintillation and servo noise in the middle ranges and increasing rapidly again as receiver noise became appreciable at long ranges. Since the angle-noise effects are the same in all the tracking systems considered, these facts explained the failure of monopulse and high-speed-lobing radars to outperform conventional-scanning radars during the earlier tracking tests. All of these tests had been made at relatively short ranges to avoid any contribution by receiver noise; thus angle noise, which was independent of the lobing technique used, had predominated in all cases and had obscured the reduction in amplitude scintillation effects which undoubtedly occurred when tracking with the experimental systems. Later tracking comparisons in which the various noise characteristics of the target were isolated have shown the superiority of monopulse systems very clearly.

During the first tracking-noise work with aircraft targets, it became evident that although actual data on aircraft were of considerable military interest, the method was unsatisfactory for a rigorous investigation of radar-system performance. Even aside from the expense of operating aircraft as targets and the delays caused by weather and maintenance, the difficulty of controlling the experimental variables made the technique relatively inaccurate, and it appeared that use of a target and radar simulator would offer many advantages. Since no available equipment was suitable, a development project was started at NRL based on the concept of the two-dimensional oscillating-rod target which had been used in the theoretical studies. The simulator which resulted⁴ was then used to carry forward the investigation of radar principles, while work with aircraft targets was largely limited to the measurement of the characteristics of the targets themselves. The only important source of error involved in the use of the simulator is the discrepancy between the characteristics of the oscillating-rod reflector and a real target such as an aircraft, and there is a considerable amount of experimental evidence that, within the limits of the two-dimensional simulator, the approximation of the real target can be very close. The use of the simulator has made possible detailed investigations of radar performance that would have been impractical with real targets and has led to specifications of optimum-design parameters in a number of situations of interest in military operations.

³ J. E. Meade, A. E. Hastings, and H. L. Gerwin, "Noise in Tracking Radars," U. S. Naval Res. Lab., Washington, D. C., Rep. No. 3759; November, 1950.

⁴ A. J. Stecca, N. V. O'Neal, and J. J. Freeman, "A Target Simulator," U. S. Naval Res. Lab., Washington, D. C., Rep. No. 4694; February, 1956.

RESULTS OF TARGET-NOISE STUDIES

Some of the major results of target-noise studies are described below but first, the different terms for noise components in tracking radars should be defined.⁵ The resultant of all noise sources affecting a tracking system is tracking noise, defined as any deviation of the tracking point from the center of reflectivity of a target. A target of finite size and complex shape also causes jitter in range tracking; however, this discussion will include only the noise in angle tracking, and range-tracking jitter is described at the end of this section. The noise sources which contribute to tracking noise may be separated into four components, namely servo noise, receiver noise, angle noise, and amplitude noise. The first two of these components originate in the radar itself. The second two components are called target noise, which is generated by the target when it is of finite size and complex structure. These four noise terms have been defined as follows:

- 1) Servo noise is the hunting action of the tracking servomechanism which results from backlash and compliance in the gears, shafts, and structures of the mount. The magnitude of this noise is essentially independent of the target and will thus be independent of range.
- 2) Receiver noise is the effect on the tracking accuracy of the radar of thermal noise generated in the input impedance of the receiver and any spurious hum which may be picked up by the circuitry.
- 3) Angle noise (angle scintillation or glint) is the tracking error introduced into the radar by variations in the apparent angle of arrival of the echo from a complex target of finite size. This effect is caused by variations in the phase front of the radiation from a multiple-point target as the target changes its aspect. The magnitude of angle noise is inversely proportional to the range of the target.
- 4) Amplitude noise (amplitude scintillation) is the effect on the radar accuracy of the fluctuations in the amplitude of the signal returned by the target. These fluctuations are caused by any change in aspect of the target and must be taken to include propeller rotation and skin vibration.

The NRL target-noise studies have been concerned mainly with angle noise and amplitude noise. However, as described later in this section, even the internal noise under certain conditions of AGC operation can affect the target-noise contribution to radar-tracking noise.

To give an idea of the value of these noise components, their relative amplitude is plotted in Fig. 1 vs relative range. The relative position of the curves for each component is dependent upon many characteristics of the particular radar or target, but the curves represent typical values.

⁵ B. L. Lewis, A. J. Stecca, and D. D. Howard, "The Effect of an Automatic Gain Control on the Tracking Performance of a Monopulse Radar," U. S. Naval Res. Lab., Washington, D. C., Rep. No. 4796; July, 1956.

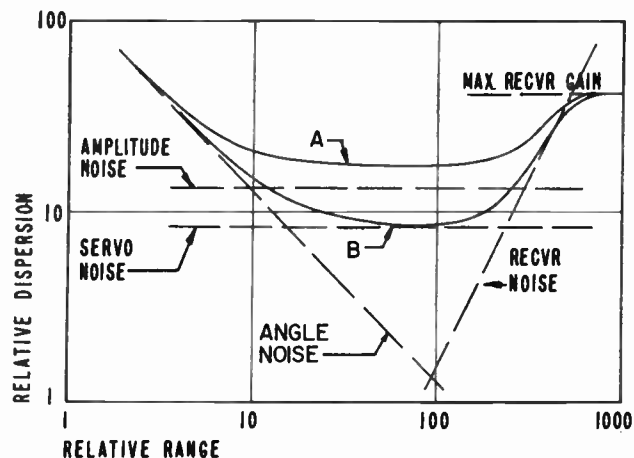


Fig. 1—Relative dispersion caused by radar-noise components as a function of range plus total dispersion for the scanning or lobeing radars, curve A, and for the monopulse radar, curve B.

Only two of these noise factors are functions of target range. Servo noise is obviously independent of range. Receiver noise increases with the square of the range (due to the inverse fourth-power law for echo power received) up to the point at which maximum receiver gain is reached. Amplitude noise, since it is interpreted by the radar as amplitude modulation of the mean signal level, is independent of range if an AGC system is used. Angle noise is a function of the linear dimension of the target and therefore varies inversely with range, as long as the target-subtended angle is small compared with the beamwidth of the antenna. Since these factors are uncorrelated, the total output noise in a given tracking system adds in mean-square fashion. Curve A is the total output noise for a sequential-lobing system. Curve B is the output noise for a monopulse system, assuming that amplitude noise is larger than servo noise. If the reverse is true, the improvement of monopulse over sequential will be negligible.

The cause of the additional noise component for sequentially-lobed or any scanning-type radar is the amplitude noise which falls in the vicinity of the lobing or scanning rate. The lobing or scanning radars obtain tracking information from the modulation of the echo signal caused by the scanning or lobing; consequently, any noise modulation contributes a noise output of the error detectors which has an rms value³ of

$$\sigma_{amp} = 0.85BA_{amp}\sqrt{\beta}, \tag{1}$$

where

B is the two-way-antenna beamwidth (the units of B determine the units of σ_{amp})

A_{amp} is fractional-noise-modulation density in the vicinity of the lobing rate in units of fractional modulation per $\sqrt{\text{cps}}$

β is the servo bandwidth in cps.

The constant has units of fractional beamwidth per fractional modulation.

In addition to the noise caused in lobing and scanning radars by high-frequency amplitude noise, there is angle noise that is a function of target size and shape plus a noise component caused by the low-frequency amplitude fluctuations which are not removed from the radar receiver output when using a long time constant AGC.⁵ The noise caused by the low-frequency amplitude fluctuations is excluded from Fig. 1 because it is not a definite function of range. All types of radars are subject to the angle noise plus the noise caused by low-frequency amplitude fluctuations which together are called target noise; and, for simplicity in this portion of the discussion, target noise will not include σ_{amp} . However, for lobing and scanning radars the noise power σ_{amp}^2 must be included to determine total target noise power.

The value of target noise has been computed for the condition of a slow-acting AGC by using the voltage function e of target noise computed in⁵ for an n -reflector target

$$e = 2k \sum_{i=1}^n \sum_{j=1}^n A_i A_j (\theta_0 - \theta_i) \cdot \cos [(\omega_i - \omega_j)t + (\phi_i - \phi_j)], \tag{2}$$

where

k is a constant determined by the antenna pattern

A_i, A_j are the amplitudes of the i th and j th reflectors

θ_0 is the error between the antenna axis and target center

θ_i is the location of the i th reflector

ω_i, ω_j are the Doppler-shifted frequencies of the echo from the i th and j th reflectors

ϕ_i, ϕ_j are the phases of the i th and j th reflectors

n is the number of reflecting surfaces.

The total noise power P can be found by summing the squares of all the ac noise terms, which for n large reduces to the total noise power with slow-acting AGC

$$P = \theta_0^2 + \frac{R_0^2}{2}, \tag{3}$$

where

R_0 is the radius of gyration of the reflecting area of the target about its center (taken in the plane of the tracking axis of interest).

Whatever units are used must be the same for θ_0 and R_0 which then determine the units of P .

It is seen that the noise power with a slow-acting AGC is a function of tracking error θ_0 , as described above, causing the parabolic shape of the curves of Fig. 2 (which shows experimental data), while the angle-noise component $R_0/2$ (a constant independent of θ_0) determines the minimum of the curve. Since the tracking noise described in Fig. 2 is a function of target span, the units of noise power are normalized in terms of target span. With fast AGC, the θ_0 term in the noise power drops out,⁵ and the angle noise is theoretically infinite,

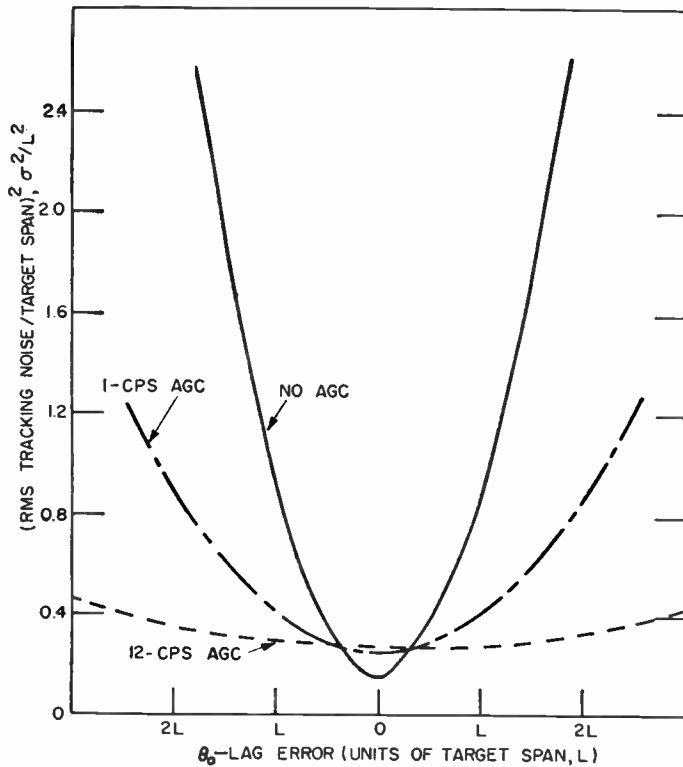


Fig. 2—Radar-tracking noise power as a function of tracking error for three AGC bandwidths.

assuming a theoretically perfect AGC. However, with a practical, fast AGC, the angle-noise term is only approximately doubled, while the total noise is essentially independent of tracking error (Fig. 2).

This increase in target noise with tracking error when using a slow-acting AGC is particularly significant at medium and long ranges, where the target subtends a small angle and tracking errors become large with respect to target span. Furthermore, the data in Fig. 2 exclude internal noise sources which, if taken into account, would increase tracking errors, thus causing an even greater contribution by amplitude noise when using the slow-acting AGC. Consequently, it is concluded that under practical tracking conditions a fast-acting AGC will give better tracking performance.

A further consideration in choosing optimum radar parameters is the servo bandwidth, which has a very definite effect on the tracking noise of a radar system. All noise in the output of the error detector appears at the input to the servo system. This includes angle noise as well as amplitude noise and internal noise. Hence the servo bandwidth of a tracking system should be limited to that necessary for the tactical requirement as determined by target range, velocity, and course.

Another point of interest is tracking radars is their ability to resolve multiple targets. This is a complex statistical problem, because of the random fluctuation of the targets, and the ability of a radar to resolve two targets, for example, can only be expressed as a probability. Fig. 3 shows the probability distribution of the tracking point of a radar looking at two targets sepa-

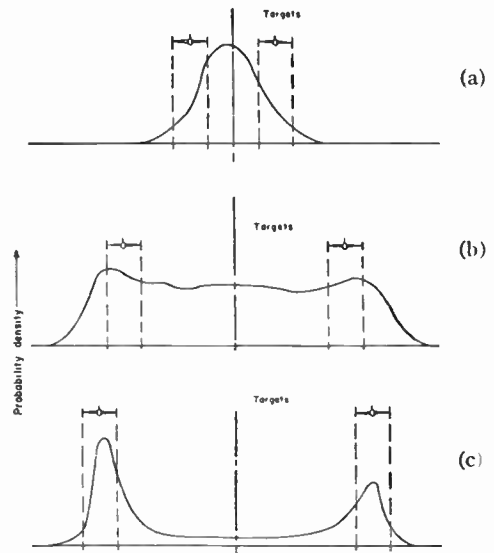


Fig. 3—Probability distributions of radar-pointing angle when tracking two targets for three different angular separations of the targets: (a) 0.50 antenna beamwidth, (b) 0.75 antenna beamwidth, (c) 0.85 antenna beamwidth.

rated by different fractions of the antenna beamwidth. The data show a strong tendency toward resolution as the target separation increases beyond 0.85 times the antenna beamwidth.⁶ The data were obtained by the target simulator and were analyzed as described in the following section.

A special condition of target noise, as discussed earlier, is the simple target composed of two reflectors. This simplified target is particularly significant because it is a close approximation of the low-angle-tracking problem. This problem occurs when the radar sees the target plus its image reflected from the surface, and serious resultant elevation-tracking errors render the elevation data almost useless. Meade analyzed the two-reflector phenomenon several years ago⁷ when he showed the point tracked by the radar to be the function

$$E = \left(\frac{L}{2}\right) \cdot \left(\frac{1 - a^2}{1 + a^2 + 2a \cos(\phi_1 - \phi_2)}\right), \quad (4)$$

where

L is the reflector spacing in either linear or angular units

$\phi_1 - \phi_2$ is the relative phase of the echoes from the two reflectors

a is the amplitude ratio of the echoes from the two reflectors.

This function is plotted in Fig. 4, which shows the apparent location of the target as seen by the radar. The two-reflector target shows, first of all, that the angle

⁶ A. J. Stecca and N. V. O'Neal, "Target Noise Simulator—Closed-Loop Tracking," U. S. Naval Res. Lab., Washington, D. C., Rep. No. 4770; July, 1956.

⁷ J. E. Meade, "Target considerations," in A. S. Locke, *et al.*, "Guidance," D. Van Nostrand Co., Inc., Princeton, N. J., ch. 11, pp. 440-442; 1955.

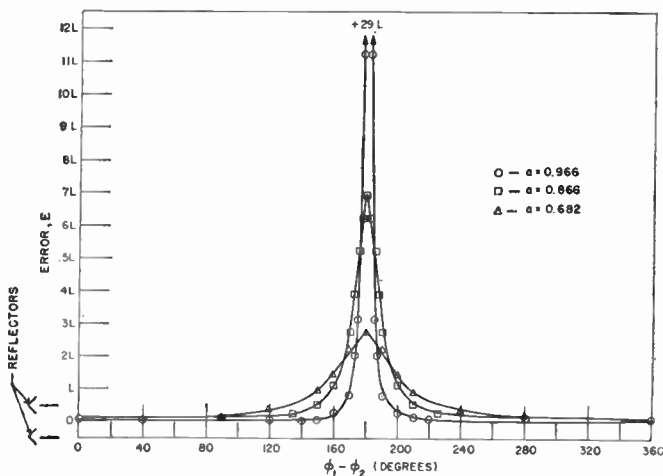


Fig. 4—Apparent location of the two-reflector target as a function of relative phase for three values of relative amplitude, a , where the reflector separation is length L .

noise can cause the apparent target to appear to be far outside the extremities of the actual target. Under certain conditions, the angle noise can even cause loss of radar track of a normal target. Secondly, it can be seen that the maximum errors occur at a 180-degree relative phase, where the echo amplitude is at a minimum (the echoes from the two reflectors subtract), so that there is a negative correlation between echo amplitude and magnitude of error. The negative correlation is significant in the effect of radar-AGC characteristics on target noise.⁵

Many of the contributions to a general understanding of the effects of target-noise sources on radar performance have come about through the ability to visualize the phenomena. The angle noise, however, which can cause the apparent target to fall completely outside the extremities of the target, is a phenomenon which is difficult to visualize, even though it has been demonstrated both theoretically and experimentally. To aid in visualization of angle noise, it can be shown that the motion of the apparent position of a target with respect to its physical location is a result of a tilting of the phase front of the echo signal received from any complex finite-size target. The phase of the signal reflected from a complex finite-size surface of n reflectors may be computed for any point in space. The slope of this function of phase with respect to the angle from the target as compared to that of a point source of reflection is then determined. This angular slope, or tilt, can be shown to be identical to the angular errors caused in a radar by target-angle noise. The phase-front-tilt concept of target-angle noise shows why this phenomenon is universally effective on any target-locating device.

Another component of noise caused by a target of finite length and complex shape is range noise which causes jitter in range tracking. The range noise⁸ is

⁸ D. D. Howard and B. L. Lewis, "Tracking Radar External Range Noise Measurements and Analysis," U. S. Naval Res. Lab., Washington, D. C., Rep. No. 4602; August, 1955.

shown to be a function of target length and shape and the initial investigation of measured range noise indicates an rms value of approximately 0.8 of the target length when tracking with a split-video-type range-error detector.

TECHNIQUES OF TARGET-NOISE MEASUREMENT AND ANALYSIS

Many of the original theoretical studies of target noise were substantiated by measurements made with a modified Mk 50 monopulse tracking radar. The Mk 50 radar, except for the antenna and pedestal, was duplicated to provide a closed-loop channel which tracks a beacon, in addition to an open-loop channel which tracks the skin echo. Thus, subtracting the two error-detector outputs, a voltage is obtained which is proportional to the error between the skin tracking and the reference beacon. A range potentiometer is also used to remove the range dependence of angle noise, and calibration techniques are used to determine the sensitivity of the radar and associated circuits so that error is measured in yards-error at the target, essentially independent of radar characteristics.

Amplitude noise is measured while tracking with the Mk 50 radar by recording the instantaneous echo level about the long-time average level. Closed-loop-tracking data are taken using boresight cameras. The recording and analyzing of the data are described later in this section.

The initial investigation of target-caused noise in tracking radars showed that several important aspects of the target-noise problem could not be conveniently studied by theoretical means or by actual tracking runs with the Mk 50 radar. For example, theoretical studies are limited in problems involving nonlinear characteristics of the radar or regenerative conditions that can arise in closed-tracking loops on finite size targets. Also when radar tracking multiple targets, for example, it is difficult to maintain target parameters such as target spacing. Consequently, accurate simulation techniques are needed to extend target-noise studies beyond the limitations of theory and radar-tracking runs.

To aid in the study of these special problems and to provide substantiating data in support of results from other means of analysis, a unique, finite-size target simulator was designed to work with a simulated tracking radar. The simulated radar is capable of complete closed-loop tracking.^{4,6} The basic simulator provides multiple targets, each of which may be composed of up to eleven reflecting surfaces, which may be arranged as a rigid body similar to an aircraft of adjustable distribution in space and having random motion to simulate the roll, pitch, and yaw of an aircraft. The simulated targets can have either independent or identical random motion and may be adjusted to any size or spacing desired. The target signals are processed by a simulated radar antenna where the individual reflectors of a target have the desired finite spacing so that they each appear at different bearings in the simulated space.

Closed-loop tracking is accomplished through a simulated tracking-radar system, with angle-error detector and servo control of the simulated antenna. The servo system is adjustable in bandwidth and, although it is a tight, noiseless system, simulated servo noise may be introduced.

The simulator provides accurate representation of targets which are not limited by weather conditions, as are real aircraft, and which can be maintained under a fixed set of simulated flight conditions. The various external and internal noise sources can be analyzed individually or in combination, as desired.

Several studies, performed mainly by use of the target simulator, are described below.

Analysis of the Ultra-Low-Frequency Target-Noise Components

The ultra-low-frequency components of target noise are difficult to obtain from experimental data because of the long data samples required for reliable statistical data. The simulator, however, has an adjustable time base so that, essentially, time may be speeded up to obtain considerable data on ultra-low-frequency noise components in a relatively short time. Furthermore, if necessary, the simulated target can be maintained under the same flight conditions for as long as desired.

Closed-loop Studies such as the Investigation of Resolution of Multiple Targets

The simulator provides two distinct advantages for the closed-loop-tracking studies. First, the antenna position is in the form of a voltage function which is much more readily analyzed than boresight film data taken during experimental tracking runs. Second, multiple-target noise data can be obtained with the target held at fixed angular spacing, which is of considerable aid for target-resolution studies. Sample results of the two-target-resolution studies are shown in Fig. 3, as previously described, showing the probability density of antenna-pointing position for three different target separations. The technique of analyzing the data to obtain amplitude probability distribution is described later in this section.

Study of the Effects of the Ratio of Servo Bandwidth to Target-Noise Bandwidth on Radar-Tracking Performance

This study was readily performed by the simulator because although the target noise is a random function, its spectral shape could be adjusted and maintained indefinitely as desired, and servo bandwidth is readily controlled.

Studies of the Effects of AGC Characteristics on the Performance of Tracking Radars

This study was particularly appropriate with the simulator for closed-loop tracking studies, because of the convenient form of closed-loop data and of the ease of introducing different target-tracking lags which are significant to the problem (as previously described).

The accurate determination of the statistical quantities involved in target-scintillation analysis required the processing of very large volumes of data, and it became evident early in the tracking-noise program that some type of automatic instrumentation was a necessity. The two quantities of particular interest in the radar problems were the spectral-energy distribution and the amplitude-distribution density. The frequency spectrum of a complex nonperiodic function can be obtained manually only by an extremely time-consuming analysis based on the Fourier integral, and the manual derivation of the amplitude distribution is only slightly less laborious. In both cases, a great number of points on the function must be obtained by graphical construction or numerical tabulation, and the handling of the desired quantity of tracking data in this way was virtually impossible. The automation of data handling was started in 1948, with the use of an electrical wave analyzer to obtain frequency spectrums,⁹ but preparation of records for analysis was still a very time-consuming manual operation, and as the study program was enlarged, it became clear that the use of automatic processing must be extended further. Two particular needs were the elimination of all manual replotting and transcribing of data and the provision of an automatic amplitude-distribution analyzer to complement the frequency-spectrum analyzer already in use. It was found possible to set up a complete system for automatic data handling through the use of magnetic-tape recording and playback of the target-noise function.¹⁰

To avoid the problems of direct magnetic recording of subaudio signals, an amplitude-modulated carrier was used. The choice of an AM system rather than FM led to much less severe tape-transport design requirements, and the dynamic range, although less than that of the best FM systems, was quite satisfactory for this application. A block diagram of the magnetic-tape equipment and associated units is given in Fig. 5. The magnetic-tape recorder is a commercial AF unit which is converted for low-speed recording by the simple substitution of a low-speed gear-box motor for the original standard-speed motor. The modulator and detector are special units developed for the system, but the use of the carrier technique with frequency translation during playback allowed both of them to be designed around standard AF components.

In the case of a radar-tracking analysis, the source of subaudio noise shown in the figure would be the tracking-error function from either an actual radar set tracking an aircraft or from the target simulator. A representative sample of tape from the low-speed recording would then be formed into a loop for high-speed playback, yielding a detector output suitable for spectral energy⁹ and amplitude-distribution analyses. The

⁹ A. E. Hastings, "Methods of Obtaining Amplitude-Frequency Spectra," U. S. Naval Res. Lab., Washington, D. C., Rep. No. 3466; May, 1949. Also, *Rev. Sci. Instr.*, vol. 23, pp. 344-346; July, 1952.

¹⁰ D. D. Howard, "Instrumentation for recording and analysis of audio and sub-audio noise," 1958 IRE NATIONAL CONVENTION RECORD, pt. 5, pp. 176-182.

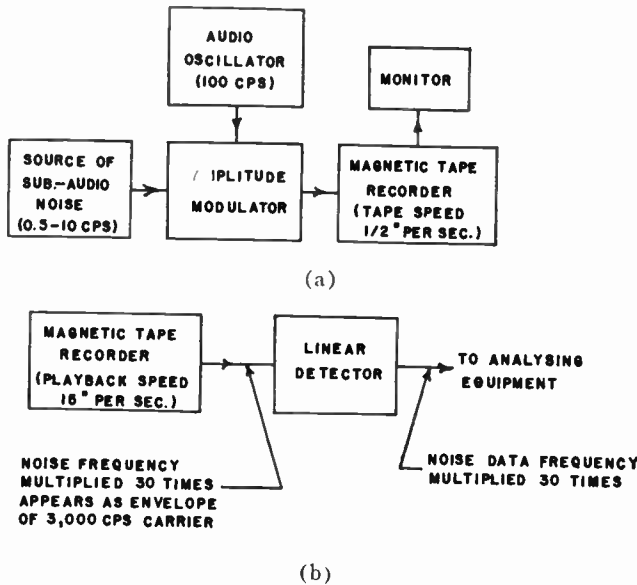


Fig. 5—Block diagram of equipment for: (a) recording, (b) playback of subaudio noise.

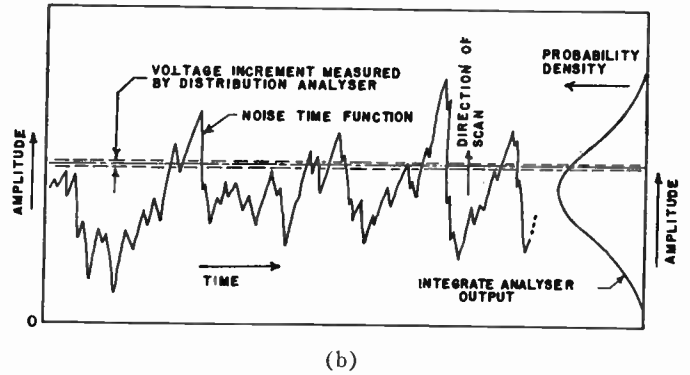
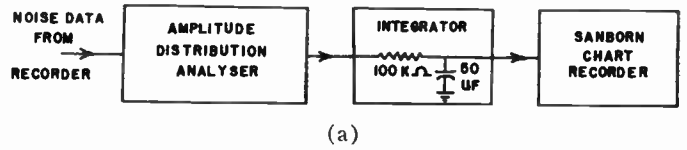


Fig. 7—Amplitude distribution analyzer: (a) block diagram of equipment, (b) operation.

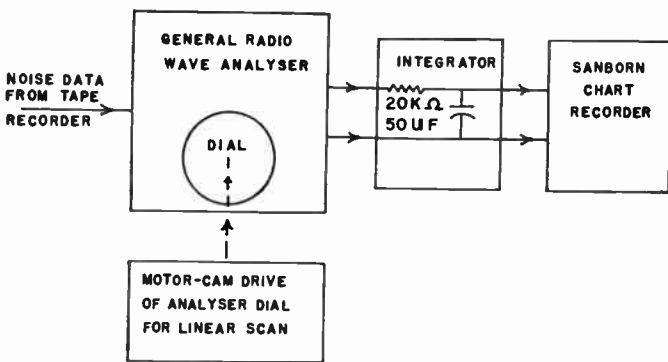


Fig. 6—Block diagram of equipment for spectral-energy-distribution analysis.

equipment for plotting the spectral-energy distribution from a loop playback is shown in Fig. 6. The wave analyzer is a standard commercial unit, except for the motor-cam adaptation and a special output connection for the integrator. The amplitude-distribution plotter is diagrammed in Fig. 7. The distribution analyzer¹¹ is a special unit developed as a companion instrument to the spectrum analyzer and performs an amplitude scan in about the same period of time required for frequency-spectrum analysis. The function plotted is the amplitude-distribution density, or second-probability density, a quantity which is of more direct significance than the amplitude distribution itself.

A sample of recording of a target-angle noise spectrum is given in Fig. 8. Since a frequency expansion of 30 was used, true noise frequency is obtained by dividing the abscissa scale by 30. The fine structure is caused by incomplete integration of the sample repetition rate and can be ignored in observing the spectrum. Longer inte-

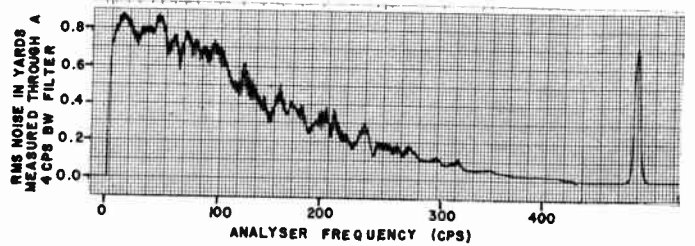


Fig. 8—Sanborn recording of the voltage spectrum of a typical sample of angle noise from an aircraft where the true noise frequency is the analyzer frequency divided by 30 as indicated in Fig. 5 (the peak at the end of the recording is a 1-volt rms signal used in scale calibration).

gration times would give a smoother plot, at the risk of some distortion of the spectrum shape. The spike at the end of the record is a 1-volt rms 60-cps-calibration signal used to determine the ordinate scale, which indicates the spectral density. By calibration of the radar, the error-detector voltage may be converted from volts to yards-error at the target. The ordinate units are rms noise in yards measured through the 4-cps analyzer bandwidth when the band is centered at the value indicated on the abscissa. The root-mean-square process must be used to obtain total rms noise. Therefore the data are generally replotted as a power spectrum and normalized to power density in yards squared per cps of true noise frequency. The rms noise is then determined by integrating the power density spectrum and taking the square root.

A second example showing a spectrum analysis of amplitude noise from a propeller-driven aircraft is given in Fig. 9, which demonstrates the care which is necessary in interpreting some types of spectrum plots. The plot of a continuous spectrum such as that of Fig. 8 may evidently be converted to a power scale and the area integrated to obtain total noise power. However, because of the limited resolving power of the analyzer, the lines of a discrete-frequency spectrum will have finite width even though the area under the peaks does not

¹¹ A. M. King, "An Amplitude-Distribution Analyzer," U. S. Naval Res. Lab., Washington, D. C., Rep. No. 3890; December, 1951.

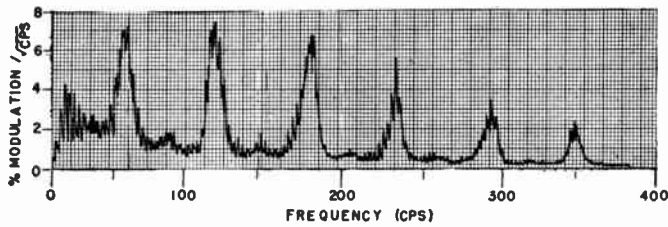


Fig. 9—Sample amplitude-noise spectrum, showing spectral lines caused by target-propeller modulation of the echo signal.

represent total power. Thus, in Fig. 9, the area under the propeller-modulation peaks must be ignored and the total noise power taken as the area lying under a line cutting through the bases of the spikes. The interpretation of the amplitude-density plots is relatively straightforward. An example of a plot for an aircraft target is given in Fig. 10. The possibility of angle-noise peaks outside the target span is shown clearly here.

OTHER WORK ON TARGET CHARACTERISTICS

The study of target noise described in this paper is only a part of the over-all investigation at NRL of target characteristics. Other work being carried on at NRL at the present time includes measurement of radar back-scattering cross section, analog simulation of targets, and recording and automatic data-reduction processes for analysis of statistical data on target characteristics.

In addition to the work described in this paper, other unclassified published work on target characteristics is listed in the literature.¹²⁻¹⁶

CONCLUSIONS

Target-noise studies at the Naval Research Laboratory have, through analysis, measurement, and simulation techniques, shown how the different components of the noise affect tracking systems and how these effects may be minimized. The conclusions from the noise studies are listed below:

- 1) Amplitude noise (amplitude scintillation) causes noise independent of target range in lobing or scanning radars because of the high-frequency components which fall in the vicinity of the lobing or scanning rates. This effect is eliminated by proper use of the monopulse-tracking techniques.
- 2) Angle noise (angle scintillation or glint) is a function of the target size and shape; therefore its magnitude varies inversely with range.

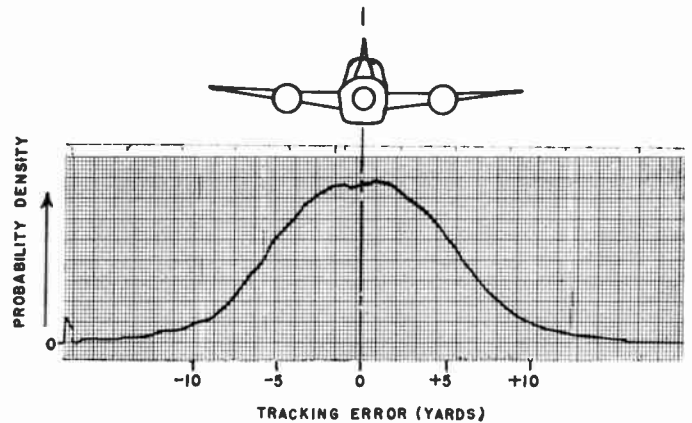


Fig. 10—Noise-amplitude-probability distribution of a sample of target-angle noise.

- 3) The AGC time constant or bandwidth affects the magnitude of angle noise and also determines the degree to which the low-frequency-amplitude noise can modulate any true tracking-error signals; thus when using a narrow-bandwidth AGC, additional noise results which is proportional to true tracking errors. Consequently, under practical tracking conditions, a wide-bandwidth short-time-constant AGC is recommended.
- 4) Servo bandwidth should be kept to the minimum, consistent with tactical requirements, to minimize target-noise effects.
- 5) The two-reflector target phenomenon shows that the angle noise can cause the target to appear many target spans outside the physical extent of the target. This phenomenon also demonstrates how the low-angle target causes severe tracking errors.
- 6) The angle noise caused by a target is shown to be caused by phase-front tilting of the echo signal, and the angular tilt is identical to the angle-noise errors.
- 7) A target of finite length and complex shape can cause jitter in the range-tracking circuit. The range noise has an rms value of approximately 0.8 times the target length when using a split-video-type detector.
- 8) It is demonstrated that a radar can resolve multiple targets when the target separation is in the vicinity of 0.8 times the antenna beamwidth or greater.
- 9) Target-noise data can be analyzed conveniently by analog techniques.

ACKNOWLEDGMENT

The authors wish to acknowledge the contributions of present and past members of the Tracking Branch of the NRL Radar Division, all of whom have taken some part in the radar-target-scintillation studies. The authors also wish to thank A. H. Schooley, the NRL Associate Director of Research for Electronics, Dr. R. C. Guthrie, the Superintendent of the Radar Division for their helpful criticisms, and E. Lauria, who typed the manuscript.

¹² R. H. DeLano, "A theory of target glint or angular scintillation in radar tracking," *Proc. IRE*, vol. 41, pp. 1778-1784; December, 1953.

¹³ R. H. DeLano and I. Pfeffer, "The effect of AGC on radar tracking noise," *Proc. IRE*, vol. 44, pp. 801-810; June, 1956.

¹⁴ J. J. Freeman, "Principles of Noise," John Wiley and Sons, Inc., New York, N. Y., ch. 10, pp. 275-292; 1958.

¹⁵ R. M. Ashby, F. W. Martin, and J. L. Lawson, "Modulation of Radar Signals from Airplanes," M.I.T. Rad. Lab., Cambridge, Mass., Rep. No. 914; March 28, 1946.

¹⁶ C. E. Brockner, "Angular jitter in conventional conical-scanning, automatic-tracking radar systems," *Proc. IRE*, vol. 39, pp. 51-55; January, 1951.

Development of Underwater Acoustic Arrays for Passive Detection of Sound Sources*

J. M. IDE†, SENIOR MEMBER, IRE

Summary—This article contains a review of a ten-year program of underwater acoustic research which has led to a ten-fold improvement in the capability of submarines to detect and track target vessels by means of passive sonar. Included in the article are a brief historical review of submarine sonar techniques utilized by the United States and other countries and a discussion of experimental sonar arrays developed by the Underwater Sound Laboratory. Technical information obtained from these experiments was incorporated into specifications for improved sonars now installed in submarines of the United States Navy.

I. INTRODUCTION

IN this article a ten-year program of researches in underwater acoustics is described; it has provided the scientific foundation for a major gain in the U. S. Navy's capability for effective action. It is seldom possible to point to so dramatic an instance of increased military equipment performance resulting directly from scientific studies.

The program referred to is one concerned with improving the ability of U. S. submarines, while underway completely submerged, to detect and track target vessels. The magnitude of the advance is stated most simply in terms of target range; here an improvement factor of at least ten has been achieved. We have gained much more than range, for we have vastly increased our ability to resolve multiple targets in the field at the same time. The certainty of target determination has also been improved, and we now get a printed record of the target's bearing, instead of a fading memory in an operator's ears.

II. REVIEW OF PASSIVE SONAR TECHNIQUES

To what, on the scientific side, do we owe this breakthrough in our capability for target detection from submerged submarines? Let us review briefly the complex of studies, experiments, and sea trials which brought us step by step to the present position. This program provides a good illustration of the fitting together of results from many areas of research to form the background for a development of great value to the Navy.

Between the first and second World Wars, underwater-sound developments in Great Britain and the United States were concentrated in the ultrasonic-frequency range. In Germany, however, as early as the late twenties, submarine sonar systems were operating in the audio-frequency range from 1 kc to 2 kc. Large multi-

element arrays were developed in order to provide directionality at the lower frequencies, and streamlined hydrophone enclosures were located below the bow of the submarine in order to decrease self-noise in the listening equipment. A bulbous appendage known as a "balcon," after the German word for "balcony," was often welded below the keel, and spot hydrophones were mounted in its surface to form an array shaped like a horse shoe. During the early years of World War II, units of the German navy and, later, the Japanese navy were equipped with these low-frequency passive systems while the U. S. and British forces were using ultrasonic equipments.

The JT system, which utilized a five-foot-line hydrophone arranged for mechanical training in bearing, was developed during World War II by the Columbia University Division of War Research at New London. This system provided U. S. submarines with a passive directional sonar equipment which furnished bearings on surface targets and made use of the target sounds in the upper end of the audio-frequency range. The lack of a streamlined enclosure exposed the system to direct water flow over the hydrophone and thus limited its operation to the lower submarine speeds. Since surface targets provided a good signal at the higher audio frequencies, where the system was most directional, the lower audio frequencies (below 1 kc) were usually filtered out, and some of the self-noise problem was thus minimized.

Toward the end of the war, with the advent of the snorkel submarine and the need for long-range detection at higher submarine operating speeds, it became desirable to exploit the ideas expressed in the German low-frequency systems.

In the period following World War II, three evolutionary phases of the development of low-frequency, passive sonar systems may be distinguished. In the first phase, covering the period from 1946 to 1950, basic information regarding system parameters in the low-frequency range was obtained. This information concerned 1) signals radiated by target submarines as related to operating conditions, 2) interference related to environmental factors, and 3) propagation phenomena related to oceanographic conditions.

Valuable pioneering in low-frequency listening and in submarine-vs-submarine tactics was carried out by submarine forces in the Atlantic and Pacific. Concurrently, a number of agencies, including the Woods Hole Oceanographic Institution, the Naval Research Laboratory, the Navy Electronics Laboratory, the Under-

* Original manuscript received by the IRE, January 15, 1959; revised manuscript received, February 18, 1959.

† U. S. N. Underwater Sound Lab., Fort Trumbull, New London, Conn.

water Sound Laboratory, and several universities under Navy contract, made important surveys of underwater-sound propagation, ambient noise, and noise-reduction procedures. The low-frequency, passive sonar program received a stimulus by the publication of the Hartwell Report in 1950.

The second phase, covering the period from 1950 to 1953, involved system design and field work with experimental systems which exploited some of the relationships the earlier work had established. The basic information was also increased by additional data on signals, interference, and propagation phenomena. Improved signal-processing methods were investigated, and a clipper correlator was designed to emphasize the correlation between signals received in the two halves of a split array. The use of a chemical recorder to record target bearing as a function of time was explored. These devices were incorporated into submarine, airborne, and ocean-bottomed fixed array systems, and the operational utility of these systems for antisubmarine warfare was experimentally evaluated.

The third phase, extending from 1953 to the present, has been concerned with the assimilation and integration of systems and tactics into specific operational plans. The need for improved location and classification information and for efficient long-range communication links to coordinate search-attack units is an important aspect of this phase. Associated with the work on systems and tactics is the need for establishing effective analytical guidelines such as figures of merit in order to minimize the time and cost of making operational evaluations of these equipments and techniques. The rapid evolution of new weapons and submarine-propulsion systems in recent years makes the fulfillment of this requirement extremely important. Not all the equipments derived from these studies, nor all the information gathered, have been fully exploited, but each of the studies has served to improve our basic understanding of sonar signals and the interfering noise background.

III. EXPERIMENTAL SUBMARINE SONAR ARRAYS

The factors promising performance improvement emerged from studies conducted with a series of experimental arrays on submarines. These arrays were constructed to be research tools for the collection of pertinent data. They were not intended to be models of operational equipment.

The arrays included a 6-foot circular array in the *U. S. S. Quillback* utilizing individual hydrophone elements having essentially no vertical directivity, an 11-foot circular array in *U. S. S. Threadfin* utilizing 3-foot-long hydrophone elements to achieve some vertical directivity, and finally, in 1951, the elliptical and linear arrays in the *U. S. S. Flying Fish*. The circular array in the *U. S. S. Threadfin* contained 96 barium titanate line hydrophones and was mounted on deck, forward of the conning tower. This array may be seen

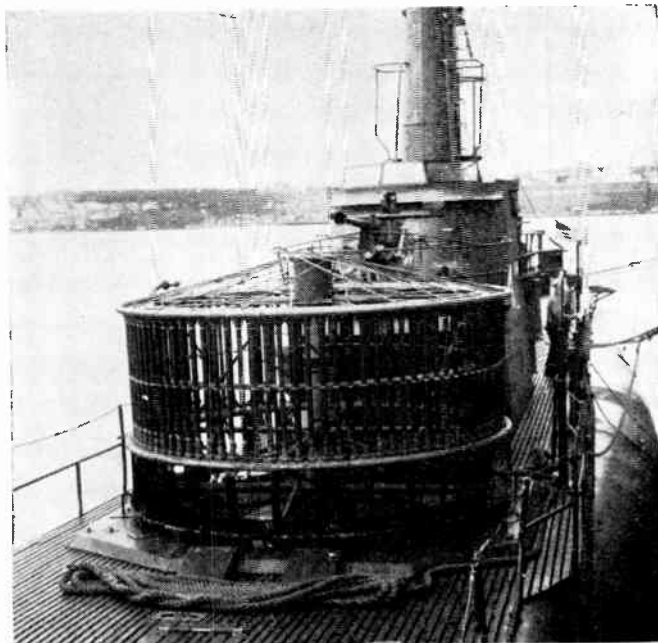


Fig. 1—Circular array of line hydrophones installed on deck of *U. S. S. Threadfin*.

in Fig. 1. The linear array in the *U. S. S. Flying Fish*, shown in Fig. 2, consisted of nearly continuous series of 2-foot-line elements having a total length of nearly 200 feet. The elliptical array in the *U. S. S. Flying Fish* had a major axis of 48 feet and a minor axis of 24 feet and utilized line elements 10 feet long. This large conformal array is shown in Fig. 3.

Extensive studies of signal processing have been carried out as part of this program. The relations existing between signals, interference, and propagation are complex and not readily amenable to theoretical analysis. An effective processing method was arrived at, however, after a study of correlation techniques applicable to sonar. In this method, the array is split, and the output from each half of the array is fed independently through separate amplifiers to a compensating switch. The compensator outputs, consisting of the summed output of each half of the array, are processed in a clipper correlator. The normalized cross-correlation function appears on a bearing chart to give a permanent record of target bearing vs time.

The current availability of data on the nature and magnitude of signals and of interference and on propagation makes it possible to determine prior to the construction of a given system what its optimum operating frequency is likely to be. The reliability with which design parameters may be predetermined in this manner has been well established by actual operating experience.

The results of a ten-year program of applied research in underwater acoustics have been reviewed. Several Navy laboratories and contractors have made significant contributions to the background information on propagation, noise, and signal processing. The specific applications to submarine sonar were accomplished

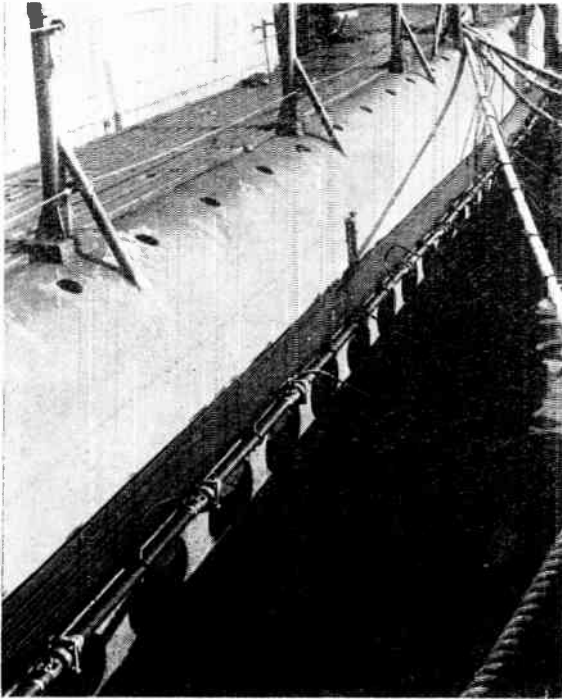


Fig. 2—Linear array of hydrophones installed in *U. S. S. Flying Fish*.

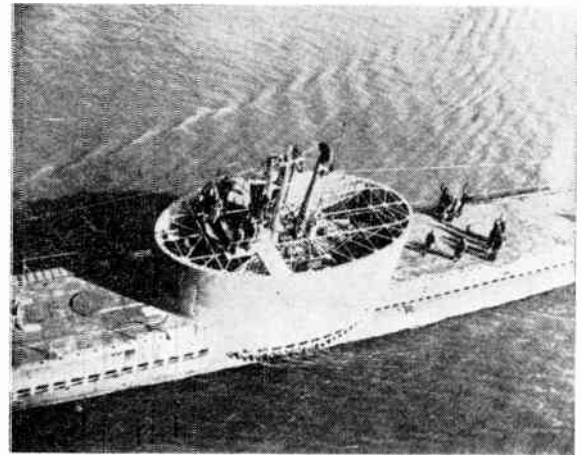


Fig. 3—Large elliptical array of hydrophones installed for experiments with *U. S. S. Flying Fish*.

chiefly at the Underwater Sound Laboratory and were carried out in large part under the supervision of C. J. Loda. The invention of the clipper correlator for sonar signal processing should be credited to J. T. Kroenert. The special transducers utilized in the experimental arrays built at the Laboratory required an immense amount of design, construction, and quality-control testing. The conformal array in the *U. S. S. Flying Fish* alone employed more than 1000 linear feet of sensitive

elements. The results achieved would not have been possible without these carefully matched transducers.

Costs charged directly to the long-range sonar detection project at the Underwater Sound Laboratory amounted to about two million dollars. The construction of the experimental arrays accounted for a large fraction of the direct costs. A tremendous amount of applied research, carried out in many places, actually contributed to the successful completion of the project. The more important areas of collateral investigation include work on propagation of sound in the sea, noise interference and noise reduction, signal processing and indicators, and transducer design. The project is an excellent example of progress made by the application of results obtained from researches in a number of related areas of endeavor to a specific purpose.

High-Fidelity Underwater Sound Transducers*

CLAUDE C. SIMS†

Summary—The problems involved in designing an underwater sound transducer for the audio-frequency range are discussed, and a new transducer for the range from about 40 cps to about 20 kc is described.

INTRODUCTION

MOST sound projectors for underwater use are resonant high-efficiency systems for radiating high intensities over a limited-frequency range. The Navy, however, also needs broad-band high-fidelity underwater sound sources that produce constant sound

fields over wide-frequency ranges for calibrating standard transducers and for other measurement functions.

The audio range has always been the troublesome part of the spectrum in underwater transducer calibrations because of the limitations of the sound sources in this range.

Existing transducers have several undesirable features such as excessive weight and size, limited frequency range, instability, irregular response characteristics, and maintenance difficulties. A new type audio-range projector has been developed at the Navy Underwater Sound Reference Laboratory to eliminate these features and thus allow more rapid and accurate calibrations. Some

* Original manuscript received by the IRE, November 27, 1958; revised manuscript received, February 16, 1959.

† Navy Underwater Sound Reference Lab., Orlando, Fla.

of the essential considerations in the design will be discussed.

FUNDAMENTAL ACOUSTIC DESIGN FORMULA

The fundamental theory of transducers for use in water is the same as that for transducers in air, aside from the problems of waterproofing, corrosion, and static pressure. The primary differences in the design result from the difference in specific acoustic impedance of water and air, and the extreme variation of static pressure to which underwater projectors are subjected.

Of the known methods of transduction, the electrodynamic method is best suited for wide-band audio-frequency sound sources in water as well as in air, since a flat response above the resonant frequency is at least theoretically possible with this method.¹

The equation for the acoustic power P_A radiated from a small source is

$$P_A = u^2 R_A$$

where u is the rms particle velocity and R_A is the radiation (mechanical) resistance. For a piston in a small housing, radiating at low frequencies into the solid angle 4π ,

$$R_A = \pi\omega^2\rho_0 a^4/4c$$

where a is the piston radius, ρ_0 is the density of the water, c is the speed of sound in water, and ω is the angular frequency. In the mass-controlled frequency range above resonance

$$u = F/\omega m_t$$

where m_t is the total mass of the piston and water load, and F is the rms driving force. At low frequencies where the projector is omnidirectional, the pressure p at distance d is related to the acoustic power radiated by

$$P_A = (p^2/\rho_0 c)4\pi d^2.$$

Therefore

$$p = F\rho_0 a^2/4m_t d,$$

which is independent of frequency.

For an electrodynamic system,

$$F = Bl i$$

where i is the rms current, B is the flux density of the polarizing field in the projector coil, and l is the length of conductor that interacts with B . The transmitting current response is then

$$\frac{p}{i} = \frac{Bl\rho_0 a^2}{4m_t d}.$$

The mass m_t consists of the diaphragm mass m_d and the water load. At low frequencies, the water load on an

unbaffled piston radiator is $1.93\rho_0 a^3$ so that²

$$\frac{p}{i} = \frac{\rho_0 Bl}{4d} \left(\frac{a^2}{m_d + 1.93\rho_0 a^3} \right).$$

Constant driving current produces constant sound pressure. It can also be seen that, for a fixed diaphragm mass, there is an optimum radius for maximum response.³ This radius is found by taking the derivative of the above expression with the radius a as the independent variable, the response p/i as the dependent variable, and the mass m_d as a constant. The response is maximum, or the derivative is zero, when

$$m_d = (1.93/2)\rho_0 a^3 \simeq \rho_0 a^3.$$

A family of curves for the response as a function of the radius for several values of m_d is shown in Fig. 1. Thus, for a piston having a mass of 7.5 grams, the response is a maximum when the radius is about 2 cm.

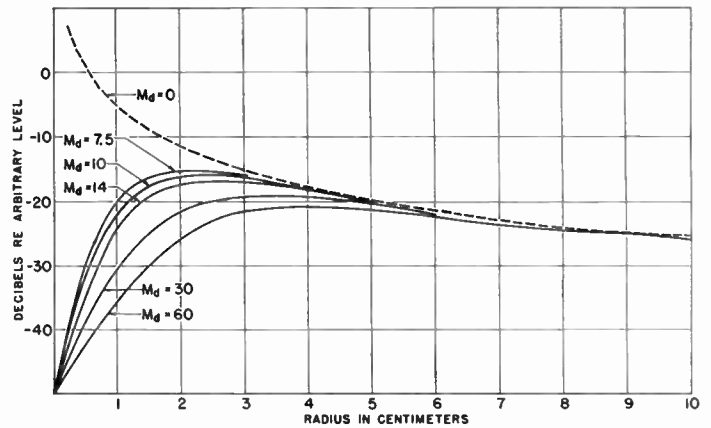


Fig. 1—Transmitting current response of a transducer as a function of piston radius for several values of diaphragm mass.

The ratio of the density of air to that of water is 0.0012, so that the radius for optimum performance in air is 10 times that for operation in water for the same diaphragm mass. This gives a decided advantage to underwater transducers in achieving high-frequency response without running into trouble from flexural resonance of the diaphragm.

Other parameters such as magnet size, coil size, weight of transducer, and frequency range are affected by the mass and radius of the diaphragm and must be considered also, so that it usually is not practical to use the proper ratio for optimum sensitivity. The curves in Fig. 1 show that considerable variation in diaphragm mass and radius is possible before there is appreciable loss in response.

² L. L. Beranek, "Acoustics," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 123-128; 1954.

³ Bell Telephone Labs., "A Subaqueous Projector for Hydrophone Calibrations in the Audible Frequency Range," OSRD-NIDRC Div. C, Sec. 4, Rep. Serial No. C4-sr212-103, OSRD No. 705; 1942.

¹ C. W. Rice and E. D. Kellogg, "Notes on the development of a new type of hornless loud speaker," *Trans. AIEE*, p. 461, April, 1925.

The complete or more general equation for the response, including directivity, of an electrodynamic projector above the resonant frequency at one meter is

$$\frac{p}{i} = \left[\frac{(Bl)^2 \rho_0 c R_\theta R_A}{4\pi [R_A^2 + (\omega m_d + X_{MR})^2]} \right]^{1/2}$$

where R_θ is the directivity factor, R_A is the series radiation resistance, and X_{MR} is the series radiation reactance.

In general, the sensitivity is limited by the size of the transducer, but it is desirable for a variety of reasons to have the transducer as small as possible in keeping with adequate sensitivity. Design should be aimed at smooth response and maximum frequency range.

The volume of the magnetic material and the size of the coil determine the Bl factor; the radius and mass of the diaphragm are determined by the frequency range required. The low-frequency limit is related to the frequency at which the total mass m_t resonates with the stiffness of the suspension system and the Q at this resonance. This frequency limit is given by

$$\omega_{01} = (K/m_t)^{1/2}$$

where K is the stiffness. The upper-frequency limit of smooth response, neglecting case radiation, is generally the frequency at which the diaphragm fails to move as a piston. For a circular plate, this occurs at the fundamental flexural-resonance frequency

$$\omega_{02} = \frac{kl}{a^2} \left[\frac{Y}{\rho(1-\sigma^2)} \right]^{1/2}$$

where k is a constant determined by the position of the driving point and the nature of the diaphragm support, l is the thickness, Y is Young's modulus, and σ is Poisson's ratio. This frequency will be modified only slightly by the water load for a relatively thick disk.

The power is limited by the maximum displacement allowed at low frequencies, and by the electrical power-handling capabilities of the coil. The displacement is given by

$$\zeta = Bli/\omega^2 m_t$$

in the mass-controlled region, and is determined by a , m_d , and the low-frequency limit.

The coil limitations are imposed by the size of wire, magnetic gap, and length of wire. The magnet size and material control the flux density B . The ratio of length to cross-sectional area of the magnet is found from the following formulas:

$$\Gamma_1 B_\theta A_g = B_m A_m$$

$$\Gamma_2 B_\theta l_g = H_m l_m$$

where A_g and l_g are the area and length of the gap; A_m and l_m are magnet dimensions; Γ_1 and Γ_2 are leakage constants; B_m is the flux density and H_m is the field strength of the magnet; B_θ is the flux density in the gap.

The magnet is designed to operate where the energy product $B_m H_m$ is a maximum.

PROBLEMS IN DESIGN OF UNDERWATER TRANSDUCERS

The major problems peculiar to electrodynamic audio-range transducers working in water result from case vibrations, the compliant water seal around the diaphragm, and the changing impedance of air behind the diaphragm due to changes in static pressure.³

Case vibrations, particularly flexural resonances, are troublesome because the acoustic impedances of the metals used in housings do not differ greatly from the impedance of the water; consequently, these vibrations radiate into the medium very well, and this spurious sound energy interferes with the diaphragm radiation. To further complicate the problem, the dimensions of the transducer are such that the natural resonant frequencies of the parts of the transducer case generally occur in the audio range.

It is also necessary that the piston be mounted with a compliant suspension system capable of operating under high static pressure. This requires compensation of the static pressure of the water by increasing the pressure of the air behind the diaphragm, so the changes of acoustic impedance of the air must be considered.

Probably the most serious defect in existing audio-range transducers is in the water seal around the piston. The radiation impedance on the piston due to the water is quite high, and to avoid shunting the radiation of the piston, there must be no areas of low acoustic impedance in the vicinity of the diaphragm. This places opposing requirements on the water seal and suspension system which are: the system must be mechanically compliant to allow large excursions and resonate at a low frequency with the relatively light piston; the system must present a high acoustic impedance to the water medium. In the past, attempts have been made to solve this problem by means of a spider-type suspension with a thin, concave beryllium-copper water seal and viscoloid dampers; with corrugated monel suspension and water seal; and with a spider suspension and rubber water seal. None of these arrangements has been satisfactory.

THE USRL TYPE J9 TRANSDUCER

The difficulties described have been overcome in the transducer shown in Fig. 2. Resonances in the magnet structure have been reduced by filling the magnet assembly ④ with lead ⑤. The cylinder ⑩ holding the piston diaphragm has been mechanically decoupled from the magnet by use of rubber O rings ⑨ which also introduce damping from the rubber in shear. The back chamber ⑪ containing the compensating air is isolated by the butyl-rubber covering ⑥, which takes advantage of the damping properties of butyl rubber.

The suspension system and water seal for the piston consist of a narrow slit ⑦ around the diaphragm which is filled with silicone oil and sealed on each end with

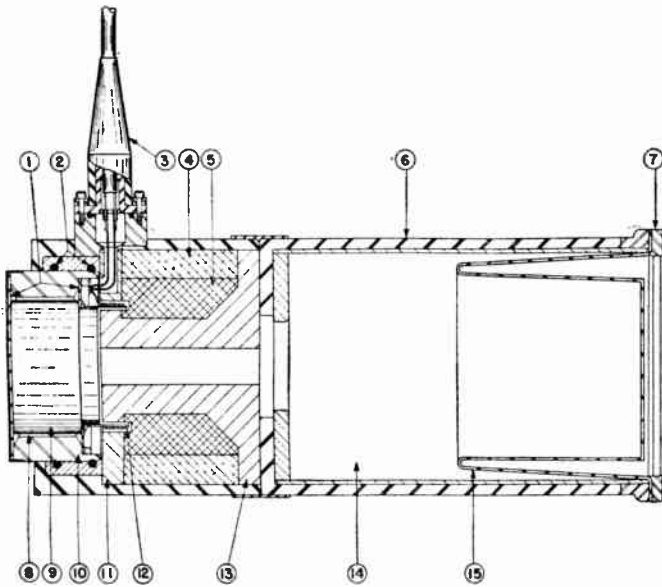


Fig. 2—Assembly drawing, USRL type J9 transducer. ① Rubber seals, ② rubber O rings, ③ cable gland, ④ magnet, ⑤ lead, ⑥ rubber jacket, ⑦ grille, ⑧ slit filled with silicone oil, ⑨ magnesium diaphragm, ⑩ diaphragm housing, ⑪ front pole piece, ⑫ coil, ⑬ back pole piece, ⑭ compensating air chamber, ⑮ rubber compensating bag.

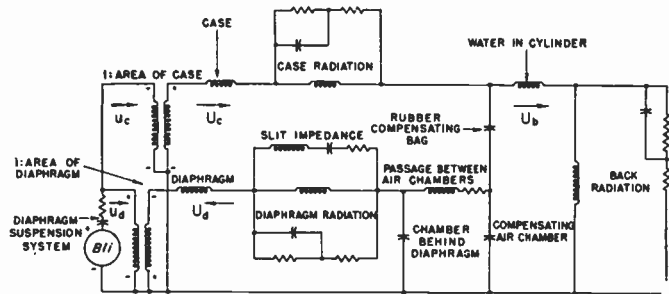


Fig. 3—Equivalent circuit of mechanical and acoustical systems of J9 transducer.

rubber ①. The rubber acts as the suspension for the piston. The mechanical stiffness which controls the resonance frequency is established by these rubber seals; the acoustic impedance of this arrangement is that of the oil-filled slit ⑥ and is controlled by the thickness and length of the slit and the viscosity of the oil. This acoustic impedance can be made sufficiently high without excessively restricting the piston motion.

The piston ⑨ is a magnesium disk with holes drilled into it to improve the stiffness-to-weight ratio.

The equivalent circuit of the mechanical and acoustical systems of the transducer, shown in Fig. 3, was derived from the various air volumes, masses, acoustic resistances, and radiation loads associated with the transducer; the electrical portion of the system is not shown. The circuit presents only a general picture of the transducer; some complicating elements are omitted for the sake of clarity. For example, the pressures that appear on various parts of the transducer because of the generated pressure in the water are not shown. The trans-

former networks necessary for coupling the two mechanical parameters at the back of the transducer (rubber compensating bag and water mass) to more than one volume velocity are also omitted; the expression for the equivalent capacitance of the rubber bag as it appears in this circuit would, therefore, be a rather complicated one.

The ideal transformers shown connect the mechanical and acoustical parameters. The masses of the diaphragm and case have been converted to acoustical units so that the driving force on the diaphragm may be represented as equal in magnitude but opposite in phase to the driving force on the case.

The currents u_d and u_c represent linear velocities; the currents U_c , U_d , and U_b represent the volume velocities associated with the case motion, the diaphragm motion into the medium and air chambers, and the motion of the back chambers in contact with the water.

The change in volume and pressure with depth will alter the values of chamber ⑭ compliance; however, the effect will not be troublesome at normal operating depths. The circuit in Fig. 3 is based on the assumption that the case and diaphragm can be treated as lumped masses (case and diaphragm-flexure modes and compliances between various components of the case are not represented).

The value for the acoustical impedance of the slit ⑥ was chosen so as to allow no loss of radiated pressure. The mechanical resistance of the diaphragm-suspension system is primarily a function of the viscosity of the oil and dimensions of the slit. The radiation impedances are the impedances of a piston in the end of a long tube, which condition is approximated in this transducer.² Except for minor variations through part of the frequency range, the impedance of this network is a fair approximation to the more complicated radiation impedance.

The circuit shows the necessity for keeping the mass of the case large compared to the mass of the diaphragm, and for keeping small the area of the case that is in a plane parallel to the diaphragm. This will keep the volume velocity due to gross motion of the case far below the volume velocity due to motion of the diaphragm, so that only the diaphragm radiation determines the sound field.

For very-low-frequency operation, the volume of air behind the diaphragm becomes important. The static pressure of the water on the diaphragm is equalized by a collapsible rubber bag ⑮ in contact with the water, as shown in Fig. 2, but using this simple method for pressure compensation introduces a variable amount of air behind the diaphragm and sufficient volume must be available at atmospheric pressure so that the air-chamber impedance will not restrict the diaphragm motion at operating depths. Also, the mass load or accession to inertia due to the back radiation, together with the mass of the water in the chamber, must resonate with the

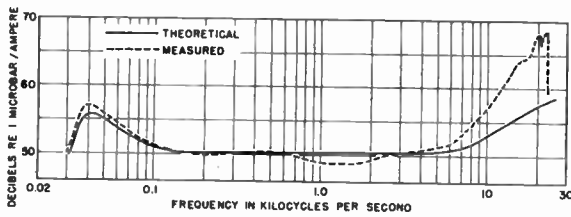


Fig. 4—Transmitting current response of J9 transducer.

compliance of the air volume below the usable frequency range. This means that the air volume and mass load must stay as large as possible.

The acoustic compliance of the compensating chamber ④ is

$$C = V/\gamma P_0$$

where V is the volume of air, γ the ratio of specific heats, and P_0 the static pressure.

The volume of air V is governed by Boyle's Law, $P_1V_1 = P_2V_2$. If V_0 is the volume of air at atmospheric pressure, V is the volume of air at the operating depth, P_0 is atmospheric pressure, and P is the static pressure at the operating depth, then

$$V = P_0V_0/P.$$

The compliance or equivalent capacitance is

$$C = P_0V_0/\gamma P^2.$$

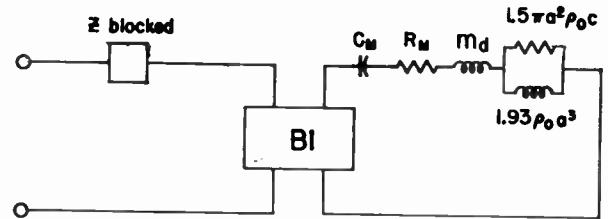
The chamber impedance increases rapidly with depth. The static pressure in the water increases by approximately one atmosphere for each 35 feet of depth, so that at a depth of 35 feet, the chamber compliance has decreased by a factor of four. The resonance frequency of the system consisting of the mass load of the water and the air volume in contact with the water does not increase as rapidly with depth, because the mass of the water in the chamber is increasing while the air volume is decreasing. This effect could be eliminated by closing the back of the transducer and allowing only a hole of high acoustic impedance to admit static pressure to the compensating bag, but it is better to leave the air chamber in the back of the transducer open so that its compliance will prevent back radiation from the case which could interfere with the diaphragm radiation.

COMPUTED CHARACTERISTICS

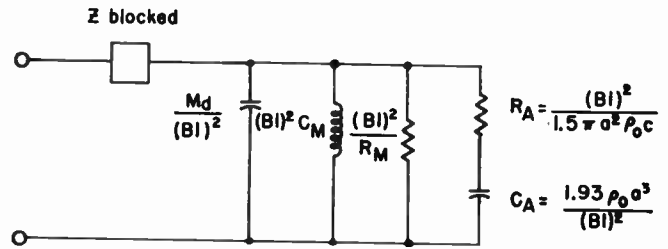
The values of the parameters for the type J9 transducer are

- $l = 16.0$ meters
- $B = 1.05$ webers/meter²
- $a = 2.85 \times 10^{-2}$ meters
- $m_d = 0.060$ kilogram.

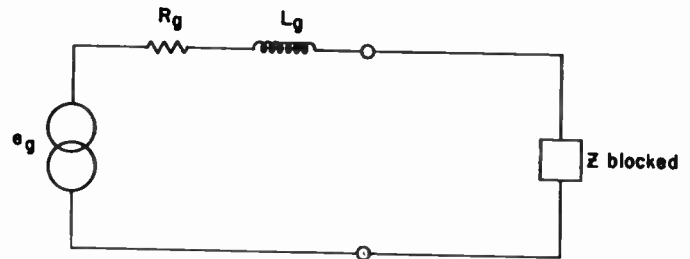
The theoretical transmitting current response above the resonance, computed from these values and the expression



(a)



(b)



(c)

Fig. 5—(a) Complete simplified circuit of an electrodynamic transducer. (b) Simplified circuit with mechanical system inverted. (c) Simplified circuit including the generator in the region above the resonance frequency.

$$\frac{b}{i} = \left[\frac{(Bl)^2 \rho_0 c R_g R_A}{4\pi [R_A^2 + (\omega m_d + X_{MR})^2]} \right]^{1/2}$$

is shown in Fig. 4. Case radiation and diaphragm break-up are neglected in this response curve. The shape of the curve near the resonance depends largely on the mechanical resistance introduced by the viscosity of the oil, the total mass, and the compliance of the diaphragm and suspension system. The transmitting current response for a resonant system with a Q of approximately 2 is shown.

When the instrument is designed so that the case velocity and the impedance of the air chamber are negligible, the simplified circuit in Fig. 5(a) may be used to represent the transducer at low frequencies. This mechanical circuit can be inverted and coupled to the electrical side with the conversion factor $(Bl)^2$ as in Fig. 5(b). The impedance of the transducer at the electrical terminals may then be found. In the mass-controlled region above resonance, the total impedance reduces to the blocked electrical impedance of the coil in the magnet gap. The blocked impedance is not a simple RL-series circuit, but has the characteristics of a coil with an iron core with hysteresis and eddy-current effects.

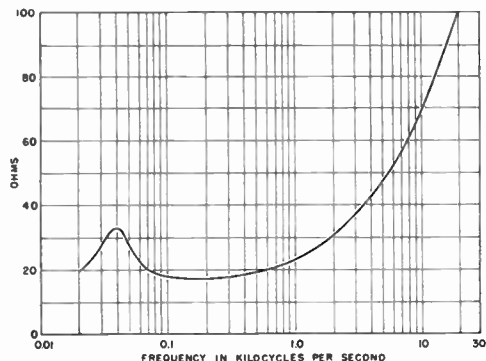


Fig. 6—Equivalent series impedance of J9 transducer.

MEASURED CHARACTERISTICS

The measured response of the J9 transducer is also shown in Fig. 4 for comparison with the computed response. The agreement is reasonable in the low-frequency region. Above 10 kc, the first flexural resonance of the piston causes a sharp rising characteristic which is a radical departure from the computed curve.

The electrical impedance is given in Fig. 6. The impedance below 100 cps is partially controlled by the viscous drag of the silicone oil and the resonance of the mass of the piston with the stiffness of the suspension system; the curve above 100 cps is simply the value of the blocked impedance.

PHYSICAL CHARACTERISTICS

Fig. 7 is an exploded view of the J9, illustrating the techniques used to simplify maintenance. A plug-in arrangement for the components avoids cable problems; each component can be removed and replaced without disturbing the other parts.

The transducer weighs approximately 15 pounds. It is 4½ inches in diameter and 10 inches long.

Fig. 8 shows the new transducer, together with the audio-range instruments previously necessary for the frequency range now covered by the USRL type J9.

UNDERWATER HI-FI

A type J9 transducer was used to provide underwater music for a water ballet as part of an AAU synchronous swim competition. The reproduction was judged as very pleasing, and indeed the experience was considered by some to be more enjoyable than listening to music in air. The reasons for this reaction are unknown, but it seems likely that it could result from the longer wavelengths in water, the different impedance ratios between the medium and the hearing organs, and different paths of sound conduction from the medium to the hear-

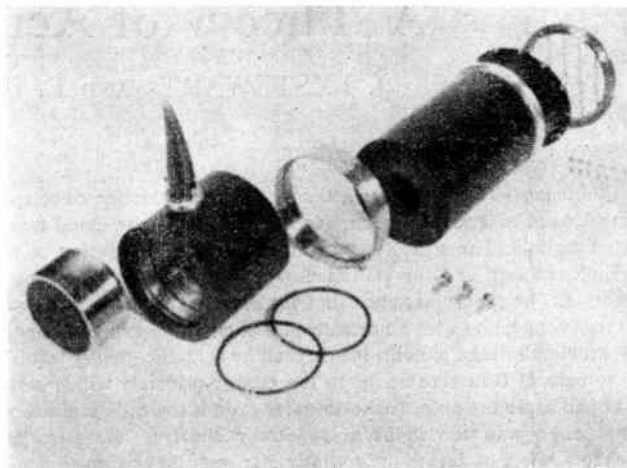


Fig. 7—Exploded view of J9 transducer.

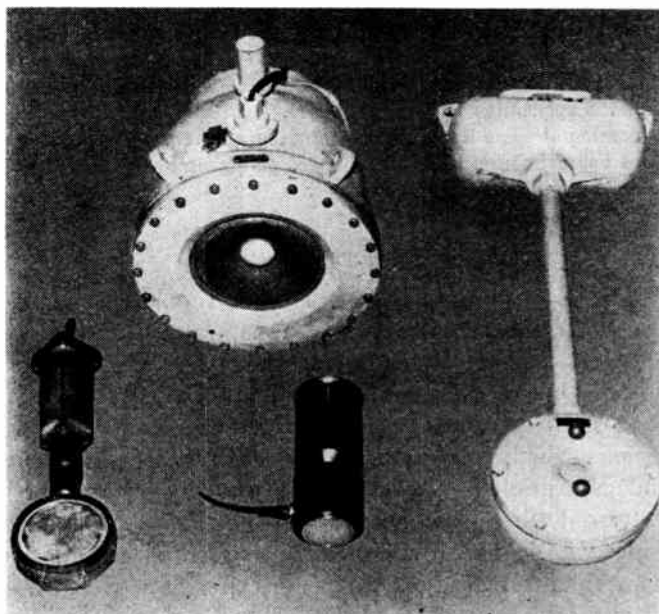


Fig. 8—Transducer type J9 in center foreground with the audio-range transducers that it replaces.

ing center. The underwater hearing process has been discussed to some extent by several investigators,⁴ but underwater music appreciation seems to be a virgin field for exploration.

ACKNOWLEDGMENT

The author wishes to acknowledge the valuable ideas and support contributed by L. J. McKay, R. J. Kieser, and I. D. Groves of the Transducer Division, Navy Underwater Sound Reference Laboratory.

⁴W. N. Wainwright, "Comparison of hearing thresholds in air and in water," *J. Acoust. Soc. Amer.*, vol. 30, pp. 1025-1029; November, 1958.

A Theory of Active Sonar Detection*

J. L. STEWART† AND E. C. WESTERFIELD†, MEMBER, IRE

Summary—The results of a theory of echo detection developed for radar are extended to sonar for the case of narrow-band transmitted signals. The exposition specifically omits discussion of the decision problem and the statistical nature of the whole process of detection. The theory assumes that the received signal is processed by cross correlation with a normalized replica of the expected signal. The problem of echo to noise is analyzed by sampling theory leading, in the case of Gaussian noise, to the energy principle which states that in all cases the output echo-to-noise ratio is the ratio of the input signal energy to the input noise power spectrum density. The threshold effect in detection is briefly discussed in conjunction with the problem of noise ambiguities. The problem of echo to reverberation (clutter) is treated by the use of the time and frequency shift correlation function leading to the ambiguity diagrams for various waveforms. The calculation of the relative reverberation power for short and long pulses of single frequency, FM and pseudorandom or noise-like waveforms for certain special cases is presented to illustrate the application of these principles. The paper concludes with a discussion of some limitations to the theory in terms of its narrow-band approximation, the instability of platforms and medium, and combined space-time ambiguities.

INTRODUCTION

THE purpose of this paper is to report the results of some recent investigations into the theory of echo detection for radar and to show the application of some of these results to sonar. Out of the mass of papers on the topic of signal detection, the three publications which offer perhaps the clearest development and exposition of the topic are those by Woodward,¹ Peterson, Birdsall and Fox;² and Siebert.³ The present paper will deal with the theory of signal processing for echo enhancement, omitting the exposition of the decision problem.²

It has long been known by the Fourier transform criterion⁴ that the optimal processing for a signal buried in additive Gaussian noise is to crosscorrelate the input with a replica of the expected signal. All of the earlier and simpler forms of signal processing for active detection denoted as ping and FM sonar may be shown to be special cases of the more general processing method called crosscorrelation or matched filtering, which can also handle the more complicated pseudorandom or noise-like signals. That is to say, a band-pass filter is the matched filter for a pulse of single frequency and an FM sonar employs a crosscorrelator. Since all of these

systems are identical with regard to processing, the only difference in detection performance one should expect, other than that due to signal energy, would be that due to the difference in the signal waveform or coding. The present paper is not concerned with any of the problems of search in bearing, range or velocity, nor is it concerned with the instrumentation problems, attendant on employing a pseudorandom or noise-like signal rather than a pulse of single frequency. The question to be discussed then is what signal waveform or coding should be used to detect a sonar echo signal against a particular background of unwanted signals, *i.e.*, noise and reverberation (clutter); that is, simply what waveform will produce the largest echo-to-noise and echo-to-reverberation ratios in various sonar situations.

SIGNAL ANALYSIS AND SIGNAL SPACE

Since we are primarily concerned with signal waveforms or coding, let us first take a look at the description of a signal in terms of sampling theory. It is the result of sampling theory, as illustrated in Fig. 1, that a signal whose significant positive spectrum lies within the finite band between 0 and W_i can be specified by a sequence of sample amplitudes s_r spaced $1/2W_i$ apart. Thus a signal of finite duration T has $2W_iT$ samples. The energy \mathcal{E} of such a finite duration signal is

$$\mathcal{E} = \frac{1}{2W_i} \sum_1^{2W_iT} s_r^2. \quad (1)$$

Different finite signals then are represented by different sequences of samples. It is philosophically interesting and theoretically expedient to represent a signal as a point (or a vector from the origin to the point) in a multidimensional signal (function) space⁵ which has as many dimensions as there are sample intervals and where the different coordinates of the point are the sample values s_r . Each arbitrary combination of sample values establishes a point in this space, and thus all possible signals are represented by all points in signal space. In this representation, a crosscorrelator resolves the one input signal (function) vector along the other input signal (function) vector, *i.e.*, forms the scalar product of the two vectors, and thus a matched filter may be considered to be a time-sequence array which isolates a direction in signal space in the same sense that a space array isolates a direction in real space.

ECHO-TO-NOISE RATIO AND THE ENERGY PRINCIPLE

First, we shall look briefly at the problem of the echo signal buried in Gaussian noise. Employing sampling

* Original manuscript received by the IRE, November 30, 1958.

† U. S. Navy Electronics Lab., San Diego, Calif.

¹ P. M. Woodward, "Probability and Information Theory, with Applications to Radar," McGraw-Hill Book Co., Inc., New York, N. Y.; 1953.

² W. W. Peterson, T. G. Birdsall, and W. C. Fox, "The theory of signal detectability," IRE TRANS. ON INFORMATION THEORY, vol. IT-4, pp. 171-212; September, 1954.

³ W. McC. Siebert, "A radar detection philosophy," IRE TRANS. ON INFORMATION THEORY, vol. IT-2, pp. 204-221; September, 1956.

⁴ S. Goldman, "Information Theory," Prentice-Hall, Inc., New York, N. Y., p. 230; 1953.

⁵ *Ibid.*, p. 166.

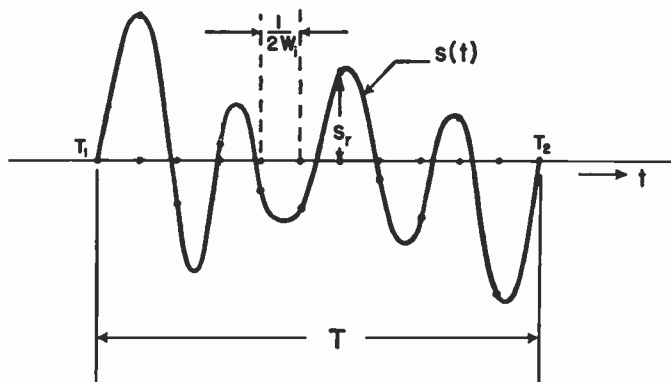


Fig. 1—A sequence of independent time samples (s_r) which represent a signal $s(t)$ of limited duration T and nominal bandwidth $W \leq W_i$.

analysis of signal waveforms, we will restate our problem as that of isolating our signal in signal sequence or signal space. Let us imagine that we have at the input of the receiver an echo $s(t)$ represented by the sequence (s_r), a time delayed version of the signal sent out, buried in a noise $n(t)$ represented by the sequence (n_r), which has a zero mean value. The mean input signal power S over the interval T is expressed as

$$S = \frac{1}{2W_i T} \sum_1^{2W_i T} s_r^2.$$

The mean input noise power is given as

$$N = \langle n_r^2 \rangle,$$

where the symbols $\langle \rangle$ signify an ensemble average of the enclosed quantity over all realizations. Since the mean value of n_r is zero, $\langle n_r^2 \rangle$ is also the variance of n_r . The signal plus noise is processed by multiplying (correlating) it with a comparison signal $x(t) = s(t)/k$ proportional in amplitude to the transmitted signal, but delayed in time to coincide with the echo in the receiver, and by integrating or summing this product over the duration of the signal. The peak output signal power is defined to be

$$\begin{aligned} \hat{S}_0 &= (\sum s_r x_r)^2 = k^2 (\sum x_r^2)^2 \\ &= (\sum s_r^2) (\sum x_r^2). \end{aligned}$$

The mean output noise power is defined as the ensemble average

$$N_0 = \langle (\sum n_r x_r)^2 \rangle.$$

Since the ensemble mean of the n_r is zero, and the n_r quantities are all independent of each other and the x_r quantities, the above ensemble average is the variance of the sums $\sum n_r x_r$, hence

$$\begin{aligned} N_0 &= \langle (\sum n_r x_r)^2 \rangle = \sum \langle n_r^2 x_r^2 \rangle \\ &= \sum \langle n_r^2 \rangle x_r^2 = N \sum x_r^2. \end{aligned}$$

Thus the output signal-to-noise power ratio is

$$\begin{aligned} \hat{S}_0/N_0 &= (\sum s_r^2)(\sum x_r^2)/N \sum x_r^2 \\ &= \sum s_r^2/N \\ &= 2W_i T S/N. \end{aligned}$$

The noise power in turn can be expressed as the noise power spectrum density \mathfrak{N} multiplied by the bandwidth W_i . With (1) this gives

$$\hat{S}_0/N_0 = 2TS/\mathfrak{N} = 2\mathcal{E}/\mathfrak{N}. \tag{2}$$

Eq. (2) is a simple derivation of the energy principle. The energy principle states the output signal-to-noise power ratio of a crosscorrelator is twice the ratio of the total energy in the signal to the noise power spectrum density per cycle at the input. This is independent of the waveform, bandwidth, power, and duration of the signal and assumes only that the noise bandwidth contains the signal and the noise has a constant spectrum density over this bandwidth. *The conclusion is then that as far as signal to Gaussian noise is concerned, the choice of signal waveform and bandwidth is entirely arbitrary. One can send out the same energy in a long pulse of low power or a short pulse of high power and obtain the same result.* To make the energy principle reasonable it should be pointed out that it is simply a restatement of the usual signal processing equation derived from power spectra as

$$\frac{S_0}{N_0} \approx \frac{S}{N} \frac{W_i}{W_0} \approx \frac{ST}{N/W_i} \approx \frac{\mathcal{E}}{\mathfrak{N}}. \tag{3}$$

Here, in addition to the symbols defined above, W_0 is the output bandwidth which is the reciprocal of the averaging time T .

For the case of a signal buried in impulsive noise, a long pulse of low power is superior to a short pulse of high power. This is obvious since a peak clipper ahead of the correlator can reduce the amplitude of each impulsive noise peak to that of the signal. Since each noise impulse has only one time sample, its energy will then be below that of the signal.

Before leaving the topic of echo-to-noise ratios it seems worthwhile to indicate the order of magnitude of the signal-to-noise ratio (SNR).

$$R = \hat{S}_0/N_0 = 2\mathcal{E}/\mathfrak{N}$$

necessary to insure a reasonable probability of detection and make negligible the chance of ambiguity due to noise. This sidesteps the problem of decision in accordance with the black and white theory of detection. The above ratio compares peak signal power to mean noise power in the output; however, it is the output noise peaks which may be confused with the signal peak to produce ambiguities. The occurrence of noise peaks

⁶ The factor of 2, difference between (2) and (3), is due to the difference in definition of \hat{S}_0 as peak output signal power in (2) and S_0 as rms output signal power in (3).

large enough to be confused with the signal peak depends on the number Ω of independent samples of the output as well as on the output SNR R . Woodward⁷ has provided a statistical treatment for the less general case of zero Doppler. The ambiguity A due to noise is given by

$$A = 1/[1 + (\Omega R)^{-1} \exp (R/2)], \quad (4)$$

which may be written

$$\frac{A}{(1 - A)\Omega} = R \exp (-R/2). \quad (5)$$

For $A \approx \frac{1}{2}$ (i.e., a 50-50 chance that a noise peak in the independent output samples will be large enough to be confused with the signal) this reduces to

$$\frac{A}{\Omega} \approx (R/2) \exp (-R/2) \quad (A \approx 1/2), \quad (6)$$

while for $A \ll \frac{1}{2}$ it becomes

$$\frac{A}{\Omega} \approx R \exp (-R/2) \quad (A \ll 1/2). \quad (7)$$

In general for $A \leq \frac{1}{2}$, A/Ω lies on or between the curves defined by (6) and (7) and as indicated in Fig. 2. Referring to this figure it is seen that the output signal needs to be at least 15 db above the output noise for almost any practical sonar situation, e.g., $A \leq \frac{1}{2}$, $\Omega = 10^6 = 10^3$ range intervals \times 20 Doppler intervals \times 50 bearing intervals. However, the ambiguity drops so rapidly above this threshold region, that a signal 18 db above the noise seems satisfactory for almost any real sonar situation. An additional 3 to 6 db may be required to compensate for instrumentation losses.

ECHO-TO-REVERBERATION RATIO AND SIGNAL AMBIGUITY

Having considered echo-to-noise ratios, attention will now turn to echo-to-reverberation ratios when cross-correlation or matched filtering is employed. Reverberation is the sum of the signals returned from unwanted targets or paths which produces a large amplitude in the output of the correlator at the instant the echo is expected. The problem of reverberation suppression is then to find a signal waveform such that a minimum of signals due to other targets, unwanted signal sources or paths, correlated with the expected signal from the desired target, a local replica of which is in the other channel of the correlator. One is reminded that in echo-to-reverberation discussions for narrow-band signals all signals wanted and unwanted are but modified replicas of the original signal and differ only in amplitude, sign, and time or frequency displacement, i.e., if one thinks of one of the targets or scatterers as being at range r and having a velocity v , then for a narrow-band signal

⁷ Woodward, *op. cit.*, p. 107.

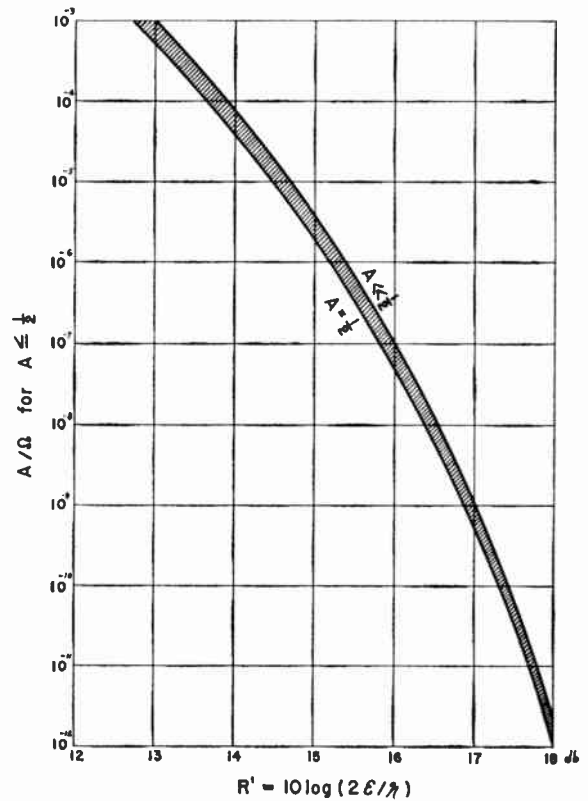


Fig. 2—A plot of the ratio of the total *a posteriori* probability A of false alarms (for the case of $A \leq 1/2$) to the number of independent τ, ϕ intervals searched as a function of the output SNR R' (R expressed in db) which is equal to twice the ratio of the input signal energy \mathcal{E} , to the input noise power spectrum density per unit bandwidth \mathcal{N} , i.e., $R = 2\mathcal{E}/\mathcal{N}$. The curve illustrates the marked threshold effect in detection that once $R' > 15$ db, the already small probability of error A for any practical Ω rapidly approaches zero.

with center frequency f_c the corresponding echo will have a frequency displacement $\phi = 2fv/c$ and a time delay $\tau = 2r/c$ where c is the velocity of sound. Thus the two-parameter autocorrelation function of the signal waveform with respect to both time displacement τ and frequency displacement ϕ (as will be discussed) is a measure of the relative response of the correlator to the echo and reverberation signals. The region of the τ, ϕ domain within which the autocorrelation function still has a significant magnitude relative to the central peak is called the region of signal ambiguity. (The signal ambiguity treated here should not be confused with the noise ambiguity treated earlier. The noise ambiguity is a function of the SNR and of the number of independent outputs observed and is relevant to the statistical decision problem. The signal ambiguity is the uncertainty of position of the target signal in the τ, ϕ domain.) Thus the aim of this approach to the problem of echo-to-reverberation will be to choose a waveform such that a minimum of reverberation signals lies within the ambiguity region when the central peak of the correlation function is in the region of the expected echo signal. For the start of our analysis it will be assumed, as is done in radar, that the bandwidth of the signal is a small fraction (1/10) of the carrier frequency such that Dop-

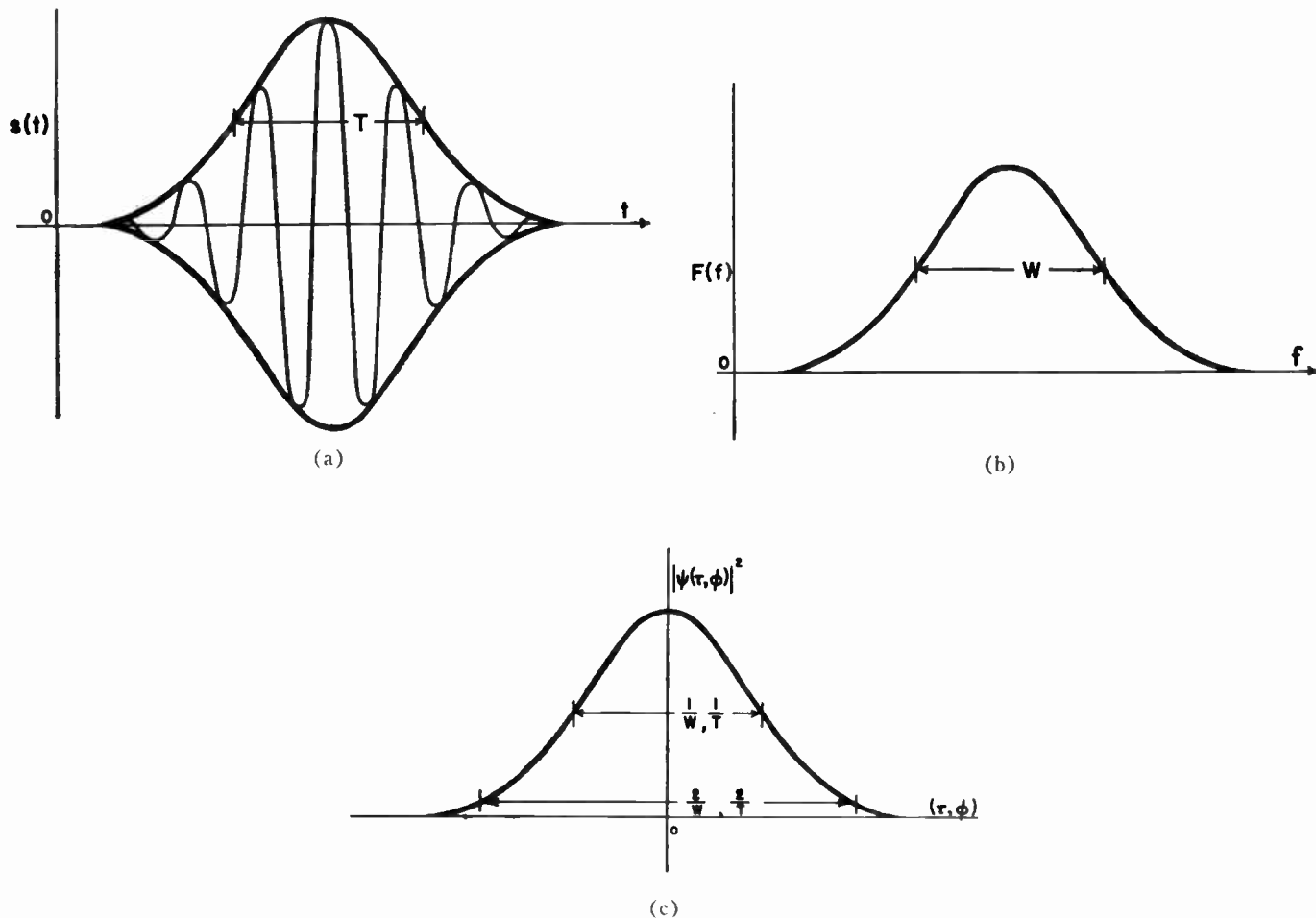


Fig. 3—A pulse with a Gaussian amplitude envelope in (a) time and (b) frequency illustrating the definitions of T and W . (c) The square of the magnitude of the correlation function for time or frequency displacement.

pler shift is approximately the same for all frequency components of the signal. More will be said about this assumption later.

Here some mathematical and physical liberties will be taken to simplify the exposition. The signals will be restricted to pulses with Gaussian-shaped amplitude envelopes in both time and frequency, as illustrated in Figs. 3(a) and 3(b). The nominal duration T of such a pulse is approximately the time interval between the half amplitude points, and the nominal bandwidth W is approximately the frequency interval between the one half value points of the amplitude spectrum. For such a pulse the square of the envelope of the correlation function as a function of time or frequency displacement is also Gaussian-shaped, as is illustrated in Fig. 3(c). This function may be viewed as the output power of the crosscorrelator to signals displaced in time or frequency with respect to the desired signal.

It is found that the nominal extent of the correlator output power between its $\frac{1}{2}$ -power values in time shift τ is approximately $1/W$ and in frequency shift ϕ is approximately $1/T$. In Fig. 3(c) it is noted that the output power of the correlator is reduced to a negligible value at the time displacement $\pm 1/W$ and at the frequency

displacements $\pm 1/T$. Properly the correlator output power $|\psi|^2$ should be pictured in a three-dimensional diagram as a function of the combined time shift τ and frequency shift ϕ .

To simplify the diagrams in Figs. 4 and 5 the correlator output power will be represented simply as the density of points on the two-dimensional τ, ϕ plane. Such diagrams are called signal ambiguity diagrams. The region within the $\frac{1}{2}$ power points of $|\psi|^2$ is indicated by a dark region called the region of complete ambiguity. The region between the $\frac{1}{2}$ -power points and the arbitrary point beyond which the power is considered negligible is indicated by grey and is called the region of partial ambiguity. Finally, the white region beyond is the region of essentially zero ambiguity.

The ambiguity diagrams for a long pulse (LP) and a short pulse (SP) of so-called single frequency are shown in Figs. 4(a) and 4(b). In the case of these simple signals, which contain only two time samples, $TW=1$. Here it is seen that the LP has low extent in ambiguity or high resolution in frequency shift, and a great extent in ambiguity or low resolution in the time shift, while the SP has high resolution in the time shift but low resolution in the frequency shift.

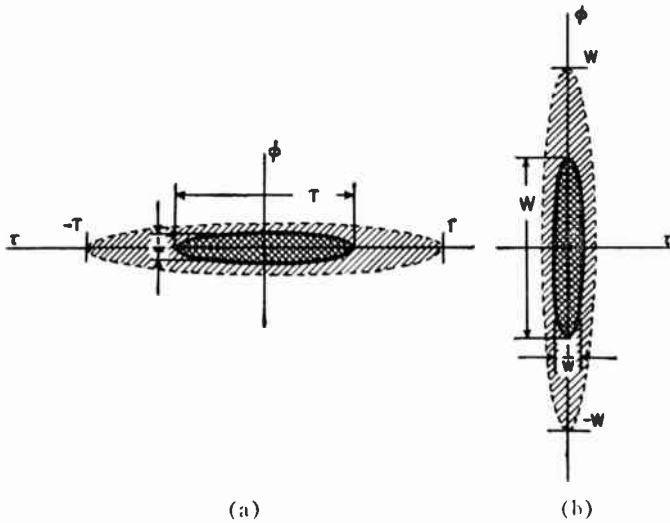


Fig. 4—The ambiguity diagrams for (a) single frequency LP of duration T and bandwidth $1/T$, and (b) single frequency SP of duration $1/W$ and bandwidth W . These are plots of the envelope of the normalized and squared correlation function $|\chi(\tau, \phi)|^2$ as a density of points in the τ, ϕ plane.

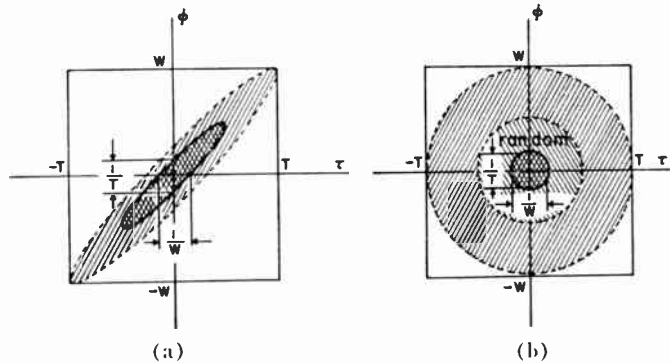


Fig. 5—The ambiguity diagrams for (a) FM and (b) PR pulses of duration T and bandwidth W . These are plots of the envelope of the squared normalized correlation function $|\chi(\tau, \phi)|^2$ as density of points in the τ, ϕ plane. Note the region of partial ambiguity for the PR signal within which the $|\chi|^2$ function has a random component of average amplitude $1/TW$.

It is noted in the above argument that ambiguity of $|\psi(\tau, \phi)|^2$ is relative to the height of the central maximum $|\psi(0, 0)|^2$. That is, whenever the height of $|\psi(\tau, \phi)|^2$ becomes equal to the height of the central peak $|\psi(0, 0)|^2$, there is complete or unit ambiguity. Thus it is natural to normalize $\psi(\tau, \phi)$ by dividing by $\psi(0, 0)$ if it is to be interpreted as a measure of the ambiguity distribution. Following Woodward, this normalized correlation function will be designated by $\chi(\tau, \phi)$. A very interesting generalization observed is that the total volume or ambiguity under the normalized $|\chi|^2$ function is the same in Fig. 4(a) and 4(b). This generalization can be demonstrated as a necessary condition for all waveforms and is known as the uncertainty principle. Since in these cases of simple signals most of the volume is in the central region of total ambiguity, the dark areas are the same. The uncertainty principle is here expressed in the form of an integral:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\chi(\tau, \phi)|^2 d\tau d\phi = 1. \quad (8)$$

The corresponding ambiguity diagrams for FM and pseudorandom (PR) waveforms are given in Figs. 5(a) and 5(b). These signals, in the special cases illustrated, occupy the total available bandwidth W and duration T , here drawn equal by an appropriate choice of scales. Here it is seen that the FM signal is not radically different from the LS or SP of single frequency. It is after all a simple signal to describe and may be regarded as a transformed single-frequency pulse. As a result its dark area of total ambiguity is equal to that for the simple ping. For such an FM pulse, if one knows the time delay of the signal, the resolution in frequency is high or, if one knows the frequency shift, the resolution in time delay is high but, if one knows neither, the total area of complete ambiguity in time and frequency or range and velocity is the same as that for the single-frequency pulses.

The case of the PR or noise-like but known signals is quite different. This is not surprising since it takes, for example, $2WT$, essentially independent, time samples to describe such a signal. Since a time compressed or expanded [that is, a Dopplered (see Appendix II)] PR signal does not correlate with a time delayed PR signal, the region of total ambiguity represented by the dark area of extent $1/TW$ is much smaller than for the other signals. However, since the uncertainty principle must hold, the total volume under this $|\chi|^2$ function must be unity also. This is achieved by a region of partial ambiguity of extent TW which is due to the self-noise or random component of the $|\chi|^2$ function of the noise-like PR signal. It may be argued as follows that once two samples of a PR signal are displaced beyond $\pm 1/W$ in time or $\pm 1/T$ in frequency they no longer are correlated and therefore may be treated as noise with respect to each other. The random products thus formed are an output noise which is reduced by the factor $1/TW$ relative to the output signal as in (3). As T approaches infinity, this factor approaches zero as noted in the usual treatment of the correlation of infinite noise samples.

Now to return to the question of echo to reverberation. If the unwanted targets (reverberation sources) are uniformly distributed in range r and velocity v over the entire region about the desired target, then the uncertainty principle states there is no choice among the various waveforms. However, for a nonuniform distribution of unwanted scatterers there is a marked difference in the reverberation return by the different waveforms. (In the discussion which follows, T and W will signify the maximum available values of these parameters, and it will be assumed that $TW \gg 1$. The SP of duration $1/W$ will employ the maximum bandwidth W . The LP of maximum duration T will employ the relatively narrow bandwidth $1/T$. The FM and PR

signals, on the other hand, will employ the maximum bandwidth W as well as the maximum pulse duration T .)

Fig. 6 illustrates the situation of a distribution of the unwanted signals or scatterers, which is partially resolvable in Doppler $\phi = 2vf_e/c$ from the target but which is unresolvable in time delay, $\tau = 2r/c$ from the target. That is, while the nominal extent in Doppler, b , of the distribution of the unwanted targets is greater than the available resolution $1/T$ in frequency shift, the nominal extent in time delay, a , of the distribution is less than the available resolution $1/W$ in time delay. Note that only the case of the target at the point A , well inside the distribution, is considered. In the case of the target at point B , well outside the distribution of the unwanted scatterers, the target is almost completely resolvable in range from the unwanted signals by a SP and hence this is the best waveform for resolving target B in this case.

Also in Fig. 6 are illustrated the regions of total ambiguity for the four different classes of signal, SP and LP of single frequency, FM, and PR signals, which give rise to the systematic component of the reverberation. The large circle labeled PR (ran) encloses the region of amplitude $1/TW$ of the $|\chi|^2$ function for the PR signal which gives rise to its random component of reverberation. It must, of course, be assumed that the product TW is large compared to unity for there to be a worthwhile distinction among these signals. The figure illustrates the special case for which $b \geq W \gg 1/T$. For this case the relative magnitudes of the reverberation for the different waveforms are approximately

$$\begin{aligned} \text{SP} &= aW & \text{PR (systematic)} &= a/T \\ \text{LP} &= a/T & \text{PR (random)} &= aW/TW = a/T \\ \text{FM} &= a/T & \text{PR total} &= 2a/T, \end{aligned}$$

from which it can be concluded that

$$\begin{aligned} \text{LP} &= \text{FM} = \text{PR (systematic)} = \text{PR (random)} \\ \text{LP} &= \text{FM} = (1/2) \text{PR (total)} \ll \text{SP}. \end{aligned}$$

If everything is assumed, as the above case, except that $W > b > 1/T$, the result is the same, except that $\text{SP} = ab$ and $\text{PR (random)} = ab/TW$, thus giving

$$\text{PR (ran)} < \text{PR (sys)} = \text{LP},$$

with the conclusion that

$$\text{FM} = \text{LP} < \text{PR (total)} < 2\text{LP} < 2\text{SP} = 2ab.$$

More generally, in the case where only Doppler resolution is possible, it may be concluded that all signals of the same duration have the same systematic reverberation and the greater the duration, the better the echo-to-reverberation ratio. The PR signal has in addition a random component which is equal to or less than the common systematic component.

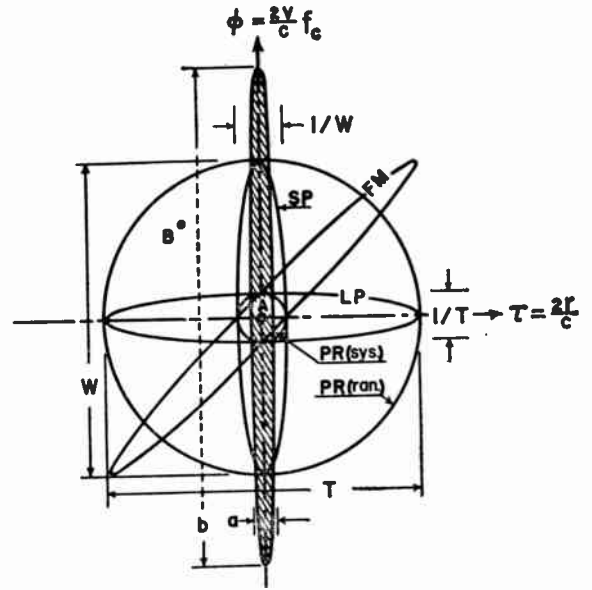


Fig. 6—The crosshatching represents a distribution of unwanted signals or scatterers (about a target at A) with a nominal extent in Doppler, b , greater than the available frequency resolution $1/T$, and a nominal extent in time delay, a , less than the available time resolution $1/W$. The curves marked SP, LP, FM and PR (sys) represent the regions of total ambiguity of the respective waveforms. The large circle marked PR (ran) encloses the region of low ambiguity for which $|\chi(\tau, \phi)|^2 \approx 1/TW$ for the PR signal.

Fig. 7 illustrates the situation of a distribution of the unwanted signals or scatterers which is partially resolvable in time delay or range from the target but unresolvable in Doppler or velocity from the target. That is, while the nominal extent in time delay, a , of the distribution is greater than the available resolution in range $1/W$, the nominal extent in Doppler, b , of the distribution is less than the available resolution in velocity $1/T$. Note that we are concerned only with resolution of unwanted relative to wanted signals and thus we are considering only the target located at point A well inside the distribution. In the case of the target at point B , well outside of the distribution of the unwanted signals, the target is almost completely resolvable in Doppler from the unwanted signals by a LP of single frequency, and is hence the best waveform for resolving target B in this case.

Again in Fig. 7 are shown the regions of total ambiguity for the four different classes of signals and partial ambiguity of amplitude $1/TW$ for the PR signal centered about the target at position A . The figure illustrates the particular case for which $a > T \gg 1/W$. For this case the relative magnitudes of the reverberation for the different waveforms are approximately

$$\begin{aligned} \text{SP} &= b/W & \text{Pr (systematic)} &= b/W \\ \text{LP} &= bT & \text{PR (random)} &= bT/TW = b/W \\ \text{FM} &= b/W & \text{PR (total)} &= 2b/W, \end{aligned}$$

from which it can be concluded that

$$\text{SP} = \text{FM} = \text{PR (sys)} = \text{PR (ran)}$$

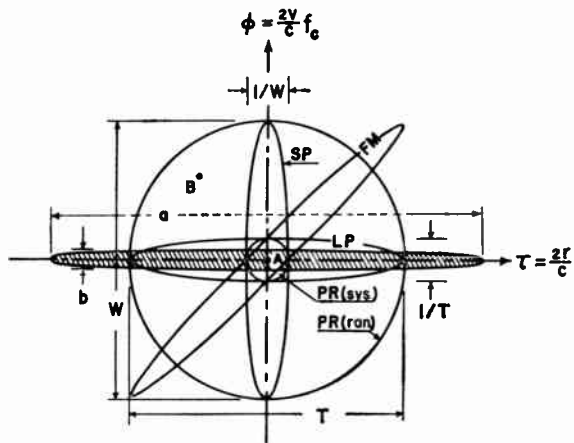


Fig. 7—The crosshatching represents a distribution of unwanted signals or scatterers (about a target at *A*) with a nominal extent in time delay, *a*, greater than the available time resolution $1/W$, and a nominal extent in Doppler, *b*, less than the available frequency resolution $1/T$. The curves marked SP, LP, FM and PR (sys) represent the regions of total ambiguity of the respective waveforms. The large circle marked PR (ran) encloses the region of low ambiguity for which $|x|^2 \approx 1/TW$ for the PR signal.

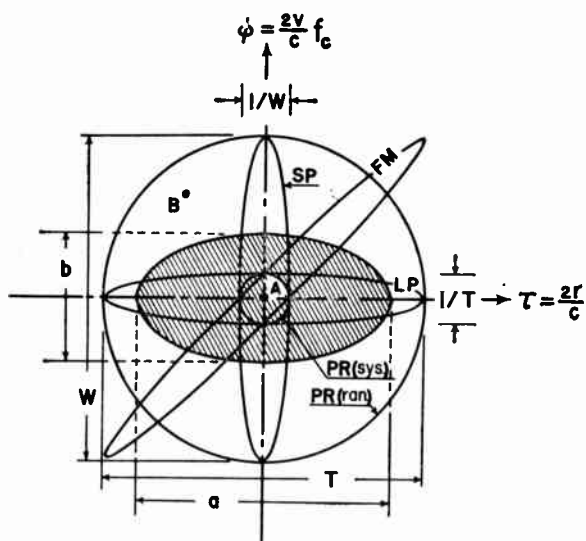


Fig. 8—The crosshatching represents a distribution of unwanted signals or scatterers (about a target at *A*) whose nominal extent in time delay, *a*, and Doppler, *b*, are each respectively greater than the resolutions in time, $1/W$ and frequency, $1/T$. The curves marked SP, LP, FM and PR (sys) represent the regions of total ambiguity of the respective waveforms. The large circle marked PR (ran) encloses the region of low ambiguity for which $|x|^2 \approx 1/TW$ for the PR signal.

and

$$SP = FM = (1/2) PR \text{ (total)} \ll LP.$$

If everything is assumed the same as in the above case except that $T > a > 1/W$, the result is the same as above except that $LP = ab$ and $PR \text{ (ran)} = ab/TW$, thus giving $PR \text{ (ran)} < PR \text{ (sys)} = SP$, with the conclusion that

$$FM = SP < PR \text{ (total)} < 2SP < 2LP.$$

More generally, in the case where only range resolution is possible, it may be concluded that all signals of the same bandwidth have the same systematic reverberation and the larger the bandwidth, the better the echo-to-reverberation ratio. The PR signal has, in addition, a random component which may be equal to or less than the common systematic component.

Fig. 8 illustrates a more general case of a distribution of unwanted signals or scatterers which is partially resolvable in both time delay and Doppler, relative to the unwanted signal or target. That is, $a \geq 1/W$ and $b \geq 1/T$ while, in addition, $T \geq a$ and $W \geq b$ are assumed. Again a target at *A* well inside the distribution is being considered and not the target at *B* well outside of the distribution of unwanted signals, where it may be almost completely resolved either by a LP or SP.

The relative magnitudes of the reverberation are approximately

$$\begin{aligned} SP &= b/W & PR \text{ (systematic)} &= 1/TW. \\ LP &= a/T & PR \text{ (random)} &= ab/TW \\ FM &= \min(b/W, a/T) & PR \text{ (total)} &= 1/TW + ab/TW. \end{aligned}$$

To evaluate this, ab must be specified within the limits imposed by the previous conditions, *i.e.*,

$$\frac{1}{TW} \leq ab \leq TW,$$

and as before $TW \gg 1$ is assumed.

1) $a \geq T$ and $b \geq W$ assumed, the reverberation magnitudes are all the same and equal to unity by the uncertainty principle.

2) $a = 1/W$ and $b = 1/T$ assumed, again the reverberation magnitudes are all the same, equal to $1/TW$.

3) Assuming $a = 1/W$ and $b = W$ and holding $ab = 1$,

$$LP = FM = 1/TW < PR = 2/TW < SP = 1.$$

4) Assuming $a = T$, $b = 1/T$ and holding $ab = 1$,

$$SP = FM = 1/TW < PR = 2/TW < LP = 1.$$

5) $a = \sqrt{T/W}$, $b = \sqrt{W/T}$ assumed, then $ab = 1$ and since $TW \gg 1$,

$$PR = 2/TW \ll SP = LP = FM = 1/\sqrt{TW}.$$

The latter case is one for which the reverberation for the PR signal may be much less than for the other waveforms. In general, the reverberation for the PR signal may be much less than that for the other waveforms when either $a/T \geq b/W \gg 1/TW + ab/TW$ or $b/W \geq a/T \gg 1/TW + ab/TW$, and in particular this will be true if $ab \gg 1$ and $W \gg b$ and $T \gg a$ or if $ab \leq 1$ and $a \gg 1/W$ and $b \gg 1/T$.

There are many other possibilities not discussed above. For example, eccentric SP, LP and FM filters such that only the end cap of the ambiguity ellipse projects into the region of the unwanted distribution to en-

compass the target, thus cutting the reverberation approximately in half. One might also ask whether or not there exist still other waveforms having lower reverberation returns than those discussed. For example, in the case of the distribution in Fig. 8, one might hope to find a waveform or code such that its ambiguity function has a large circle or ellipse of zero ambiguity in the region of the unwanted scatterers, except for a sharp peak of total ambiguity at the center in the vicinity of the target, and has the rest of its region of total or partial ambiguity occurring beyond the region occupied by the unwanted scatterers. The physical realizability of such a waveform has not been completely investigated.

In the case where the propagation losses to the wanted and unwanted scatterers are the same, the echo-to-reverberation ratio is simply the ratio of the target strength or scattering cross section of the target to that of the reverberation. This target strength or scattering cross section of the reverberation is given by

$$\mathcal{R}(\tau_0, \phi_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\chi(\tau - \tau_0, \phi - \phi_0)|^2 \mathcal{R}_{11}(\tau, \phi) d\tau d\phi$$

where now the expected target is at τ_0 and ϕ_0 and where \mathcal{R}_{11} is the target strength or scattering cross section per unit range-Doppler interval for the solid angle resolved by the source and receiving arrays. (It should be noted that the distribution in Doppler frequency shift $\phi = (2v/c)f_c$ varies with the carrier frequency f_c while the distribution in v or τ may not.) In general, the problem of the optimum waveform might be phrased as that in which ambiguity function is highest where the ratio of the distribution of wanted to unwanted signals is largest and is lowest where the ratio is lowest. (This is analogous to a likelihood ratio criterion.) However, not every ratio of distribution functions is an ambiguity function (square magnitude of the two parameter correlation function) for a real waveform. Some of the analytical properties of the ambiguity function have been investigated by Siebert.⁸

In general, it is also not possible, nor is it necessary, to have one waveform optimal for all locations of the target relative to the unwanted scatterers. It would, in fact, seem better to transmit two or more waveforms optimized for different target locations.

Finally, it should be remembered that echo to reverberation is being optimized subject to the condition that crosscorrelation or matched filtering has been employed to maximize echo-to-Gaussian noise. While it has not been demonstrated, in general, that matched filtering optimizes echo to reverberation, the reductions in the output reverberation in the cases above are the direct result of the matched filtering.

⁸ W. M. Siebert, "Studies of Woodward's Uncertainty Function," Electronics Res. Lab., Mass. Inst. Tech., Cambridge, Mass., Quart. Prog. Rep., p. 90; April 15, 1958.

While in the above discussions the differences among the SP, LP, FM and PR waveforms are emphasized, it should be pointed out that there is a continuum of possible waveforms connecting these and other extremes. This is obvious in the case of the SP and LP of single frequency. It is also seen from time-sampling analysis that PR and FM waveforms can be represented by sequences of short pulses and that by Fourier analysis of analytic signals that the PR and FM signals can be represented by a sum of long pulses of different frequencies.

Some problems still under investigation are presented in the appendixes to suggest the limitations to the above arguments.

APPENDIX I STABILITY CONDITIONS

As is well known, the targets, the platforms and the medium in sonar are not stable but have systematic and random accelerations which limit the duration over which one can say the desired target has a given range and velocity.

The stability condition 1) is that the time spent in the resolved range interval Δr must be greater than the averaging time T of the corresponding output filter; that is, $\Delta r/\bar{v} > T$ where \bar{v} is the difference in the average velocity or compressions ($d\tau/dt$) of the two signals to be compared, *i.e.*, $v = d\tau/dt = (c/2)(d\tau/dt)$. Since one can always compress or sweep the local signal to match the expected echo Doppler, \bar{v} is in reality limited to the interval between the output Doppler compression matching channels, and Δr will in general be a moving range interval which moves with the expected target. Now the resolved range interval Δr may be variously expressed as

$$\Delta r = \frac{c}{2} \Delta\tau_{\text{corr}} = \frac{c}{2W}$$

The condition 1) becomes

$$(c/2W\bar{v}) > T$$

or

$$\bar{v} < c/2TW$$

or

$$TW < c/2\bar{v}. \quad [\text{Condition 1}]$$

This condition limits the maximum processing gain which can be achieved in inverse proportion to the maximum relative average velocity between the two signals being compared. As indicated above, this condition merely amounts to an indication of the minimum separation $\Delta v = \bar{v}$ between output Doppler compression matching channels necessary to yield a given TW . Since the total range of velocities to be investigated is fixed, the total number of output Doppler channels must in-

crease as the desired TW increases as indicated in Table I below.

TABLE I
PROCESSING GAIN VS VELOCITY INTERVALS AND
TOTAL NUMBER OF OUTPUT CHANNELS

TW	Δv	Relative total number of channels
20 db	15.00 knots	1
30	1.50	10
40	0.15	100

This condition 1) will be seen again below as the condition that Doppler can be matched by simple frequency shifting of the comparison signal. For velocity differences higher than that given by $\bar{v} = c/2TW$ the comparison signal must be time compressed for proper matching.

The stability condition 2)⁹ states that the time spent in the resolved velocity interval Δv must be greater than the averaging time T of the output filter; that is, $\Delta v/\bar{a} > 2T$ where \bar{a} is the difference in the average acceleration or accelerated compression of the two signals being compared. One could, of course, match a systematically accelerated signal but since this would mean another order of magnitude in the number of output filters, it will be assumed that the local signal is unaccelerated and therefore unmatched in this parameter. The minimum resolvable Δv is expressed in terms of the usual Doppler relationship as follows

$$\phi = \frac{2v}{c} f_c$$

where ϕ is the Doppler displacement and f_c is the carrier frequency. Here the narrow-band approximation that all frequencies are shifted by approximately the same amount ϕ is assumed. Since it is assumed that matching in velocity or time compression has been achieved, we are dealing in relative displacement only between the two most nearly matched signals, thus

$$\Delta\phi = \frac{2}{c} \Delta v f_c$$

or

$$\Delta c = \frac{c}{2f_c} \Delta\phi.$$

⁹ The independent statements of conditions 1) and 2) ignore the possibility that a target might vary its range and velocity in such a manner as to remain in a single output filter of an FM system for a longer time than indicated above due to the combined time delay and frequency shift ambiguity illustrated in Fig. 4. The occurrence of the peculiar acceleration which results in such a favorable circumstance is so unlikely that its possibility has been ignored here.

Now the minimum resolvable frequency interval $\Delta\phi = 1/T$

$$\therefore \Delta v = \frac{c}{2f_c T},$$

and the condition 2) becomes

$$\frac{c}{\bar{a} 2f_c T} > 2T$$

or

$$T < \left(\frac{c}{4\bar{a}f_c} \right)^{1/2}. \quad [\text{Condition 2a)]}$$

This condition is a very severe one in terms of sonar parameters.

If one combines Condition 2a) with the practical limit that the bandwidth W must be less than or equal to the carrier frequency f_c , one has

$$TW \leq (cf_c/4\bar{a})^{1/2}. \quad [\text{Condition 2b)]}$$

This indicates a limit to processing gain if one does not match acceleration as well as range and velocity. For an active sonar operating at 100 kc on targets having a medium acceleration of 0.01 knot/second, maximum processing gain would be about 50 db. Again for an active sonar operating at 1 kc on targets having a high acceleration of 0.1 knot/second, the maximum processing gain would be about 35 db. These are possibly significant limits. For radar, on the other hand, other factors will probably limit the processing gain before one can achieve the gain of about 82 db indicated by this condition 2b) for an average acceleration of one g and a carrier frequency of 3×10^9 cps.

APPENDIX II

TIME COMPRESSION VS FREQUENCY DISPLACEMENT

In the radar studies cited and in the sonar treatment above, it has been assumed that for the narrow-band signals considered, all frequencies have approximately the same Doppler shift. This assumption depends on the condition that for sufficiently narrow-band signals, time compression or expansion is equivalent to frequency translation. In terms of correlation functions this means that for sufficiently narrow-band signals the autocorrelation function with respect to time delay τ and time compression $\alpha = dt_1/dt$,

$$\psi_1(\tau, \alpha) = \int_{-\tau}^{\tau} s(t)s([1 + \alpha]t + \tau)dt \quad (9)$$

is essentially equivalent to the autocorrelation function with respect to time delay and frequency displacement

$$\psi_2(\tau, \phi) = \int_{-\tau}^{\tau} s(t)[s(t + \tau) \cos(2\pi\phi t)]dt \quad (10)$$

where the bracket around the product in the integrand is used to denote the upper sideband for positive values of ϕ and the lower sideband for negative values of ϕ . In these expressions T , as before, denotes the nominal duration of the signal pulse, so broadening the limits of integration would make little change in the value of the integral. In the case of a single frequency pulse, $s(t)$ may be taken as $\cos 2\pi ft$ and for $\tau = 0$ the expressions in (9) and (10) may be written as

$$\psi_1(0, \alpha) = \int_{-T}^T \cos(2\pi ft) \cos(2\pi f[1 + \alpha]t) dt$$

and

$$\psi_2(0, \phi) = \int_{-T}^T \cos(2\pi ft) \cos(2\pi [f + \phi]t) dt,$$

respectively. For $\phi = \alpha f_c$ it is seen that these two expressions are identical when $f = f_c$. When $f = f_c + \Delta f$, they are approximately equal provided the quantity $2\pi \bar{\alpha} T \Delta f$ does not change by more than π within the limits of integration. This gives

$$4\bar{\alpha} T \Delta f \leq 1$$

where $\bar{\alpha}$ is the mean value of α or

$$\bar{\alpha} \leq 1/4T\Delta f.$$

There are values of Δf as great as $W/2$; thus it is seen that for a nominal bandwidth W the limitation on the $\bar{\alpha}$ is

$$\bar{\alpha} \leq 1/2TW. \tag{11}$$

For values of $\bar{\alpha}$ satisfying this condition we have

$$\psi_1(\tau, \alpha) \approx \psi_2(\tau, \phi). \tag{12}$$

For this equivalence the frequency displacement ϕ is taken as equal to αf_c where f_c is the center frequency of the signal band. From (11) this gives

$$\phi \leq f_c/2WT \tag{13}$$

as the limitations on a frequency displacement which can be taken as equivalent to a time compression in computing the correlation function.

Referring to (9), it is seen that α can also be interpreted as the time derivative of τ . Writing

$$\tau_1 = \tau + \alpha t = \tau + t \frac{d\tau_1}{dt}, \tag{14}$$

it is seen that (9) may be written as

$$\psi_1(\tau, d\tau_1/dt) = \int_{-T}^T s(t)s([1 + d\tau_1/dt]t + \tau)dt.$$

This is termed "swept correlation." For α satisfying (11) τ_1 and τ will not differ sufficiently over the period of integration to affect the value of the integral appreciably, and the subscript can be dropped giving

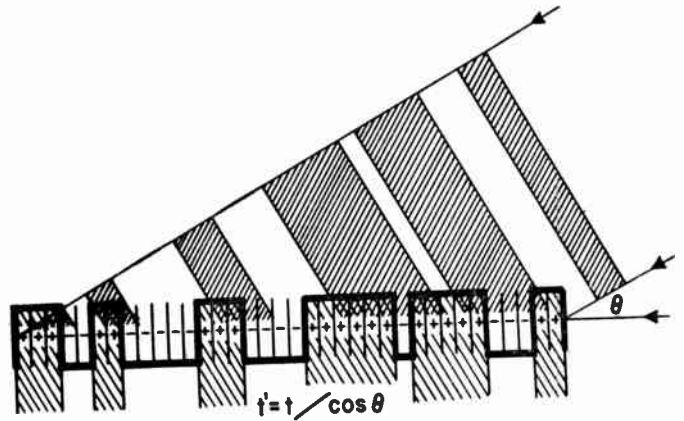


Fig. 9—A combined space-time array is represented by the horizontal sequence of + and - signs indicating the polarity of the hydrophones. It is designed to be the matched filter (time array) for the signal shown advancing along the axis of the array from the right. The same waveform advancing at an angle θ with respect to the array axis has an effective time expansion by a factor of $1/\cos \theta$ which would be ambiguous with a Dopplered signal traveling along the axis.

$$\psi_1(\tau, d\tau/dt) \int_{-T}^T s(t)s([1 + d\tau/dt]t + \tau)dt.$$

It is observed that the condition given by (11) above where $\pi = d\tau/dt = 2v/c$ is essentially the same as the stability condition 1). It is the intention of the authors to extend the treatment of Doppler to broad-band signals (the order of an octave) and treat it correctly as time compression. This may make some modifications in the applicability of the uncertainty principle and ambiguity diagrams.

APPENDIX III

OTHER SPACE AND TIME AMBIGUITIES

In the general treatment of isolation of a signal from detection in real space and in signal space, as treated at various times in this paper, a point was omitted. It is that one cannot always separate the real-space bearing localization from the signal-space resolution in Doppler. This is illustrated in Fig. 9 for the following proposal of a combined real-space and time-space array. Here, illustrated by the plus and minus signs is a real-space array with a choice of polarity outputs from each hydrophone. This particular array has been designed to be a matched filter for the particular waveform illustrated advancing from the right with wave fronts perpendicular to the axis of the array (end fire receiving array). Now it will be observed that the same wave front if approaching from an angle θ with respect to the axis will have a projected time scale $t' = t/\cos \theta$ which is the same as that for a Doppler expansion. Thus one can see that another set of ambiguity or uncertainty principles may exist between bearing and Doppler information for a combined real-space and signal-space array. (This is a well-known effect for single frequencies and has been employed in the steering or phasing of narrow-band arrays by F.M.)

The DOFL Microelectronics Program*

T. A. PRUGH†, SENIOR MEMBER, IRE, J. R. NALL†, AND N. J. DOCTOR†, MEMBER, IRE

Summary—A program on microelectronics (electronic micro-miniaturization) has been underway for the past two years at the Diamond Ordnance Fuze Laboratories (DOFL). Fourteen-component-part, transistorized binary counters were fabricated on $\frac{1}{2}$ -inch squares of steatite ceramic, 1/50 inch in thickness; and 5-part NOR's, on $\frac{1}{4}$ -inch squares of the same thickness. The techniques that have proved most useful in the program have been 1) photolithographic procedures for the accurate placement at microscopic dimensions of physical masks and electrical insulation, 2) thin film deposition using vacuum, chemical, and screening methods, 3) the use of a conductive adhesive which yields high strength and low resistivity connections, and 4) ultrasonic drilling and air abrasion enabling substrates and devices to be formed to desired sizes and shapes. Concurrent with the physical fabrication, studies were made of detailed circuit design and system applications.

Future work will include 1) refinement of the present techniques to achieve inexpensive, reliable circuit wafers capable of mass production by industry, and 2) development of advanced methods more fully utilizing thin film deposition and diffusion techniques.

I. INTRODUCTION

THE proximity fuze program¹ has required development of compact, rugged electronic assemblies in order that these ordnance devices operate under severe environmental conditions. The Diamond Ordnance Fuze Laboratories (DOFL) have been actively engaged in research on techniques for fabricating miniature electronic assemblies. This work has progressively made use of subminiature vacuum tubes, the early point contact transistors, conventional junction transistors, and more recently, ruggedized transistors designed specifically for military environments.

Much interest has developed in the past several years in the possibilities opened up by the discovery of the transistor. In particular, the small size and weight and the low power requirements of the device are attractive features of value in portable electronic equipment. Along with progress in improved techniques have gone the increasing demands in military electronics for more compact and lighter electronic systems which can be carried in missiles and satellites. Conventional techniques utilizing commercially available subminiature parts have about reached the limit in size reduction. At the 1957 IRE National Convention, Brunetti² introduced a new term—"microminiaturization"—to cover more advanced techniques made possible by rapid progress in solid state devices. At DOFL we have referred to electronic microminiaturization as "microelectronics."

Approximately two years ago (May, 1957) a team was

formed at DOFL to search for important new ways of fabricating electronic assemblies. Participating were representatives from the passive component, semiconductor devices and circuit-system areas. The objective of the program was to find ways of overcoming the lower limit in size of assemblies using individually cased component parts. As Brunetti and others have pointed out, significant improvements could be made by eliminating the individual part cases and utilizing only the essential operating core of the part. The various printed circuit techniques under development for years offered promising methods to achieve these further advances.³ It was decided to work with a thin wafer substrate onto which would be deposited or attached uncased component parts. High parts densities would be achieved using an almost 2-dimensional package, *i.e.*, one containing finite surface area yet extremely thin cross section. This technique was called the DOFL-2D approach, or more simply 2D.

Since the DOFL program was started the Signal Corps has announced a Micro-Module Program⁴ aimed at solving a broad variety of electronic applications by the use of small modular elements. DOFL's program has been more specifically directed at specialized applications where one-shot, highly reliable, ammunition concepts are important.

The following sections will mainly be devoted to a description of the detailed steps and problems in the fabricating, testing, and use of two basic circuit types. However, before describing the specific steps for a particular circuit, a brief review will be made of some recent techniques for processing and handling materials.

II. REVIEW OF TECHNIQUES

A. Photographic Processes

1) *Photoresist*: Photoresist⁵ is an organic, insulating, presensitized liquid surface coating. The use of photoresist in the production of etched wiring boards is quite well known. Other properties of this material besides its chemical resistance to copper etchants have been employed in the DOFL microelectronics program.

Coating or sensitization of a substrate is done on dry materials using conventional techniques such as spin-coating, plate coating, air brush spraying, or painting. After a substrate has been coated with photoresist

* Original manuscript received by the IRE, January 5, 1959.

† Diamond Ordnance Fuze Labs., Washington, D. C.

¹ W. S. Hinman, Jr. and C. Brunetti, "Radio proximity-fuze development," Proc. IRE, vol. 34, pp. 976-986; December, 1946.

² C. Brunetti, "A new venture into microminiaturization," 1957 IRE NATIONAL CONVENTION RECORD, pt. 6, pp. 3-10.

³ N. J. Doctor and E. M. Davies, "Microminiature components for electronic circuits," *Elec. Mfg.*, vol. 62, pp. 34-37; August, 1958.

⁴ A. W. Rogers, "The micro-module design concept in electronics," *Elec. Mfg.*, vol. 62, pp. 46-49; July, 1958.

⁵ "Industrial Uses of Kodak Photo Resist," Eastman Kodak Co., Rochester, N. Y., Pamphlet No. Q-24; August, 1956.

and allowed to dry, it is exposed through a pattern negative or positive to ultraviolet light and then developed. These steps are illustrated schematically in Fig. 1. The equipment necessary for this processing is similar to that used for contact printing and includes a light source, vacuum frames, jigs, and fixtures for proper registration. Development of an exposed pattern is done in the vapors of an organic solvent such as trichloroethylene. Unexposed resist washes away during development as shown in Fig. 1 (c). These steps result in the substrate surface's being covered only in specific areas by the film.

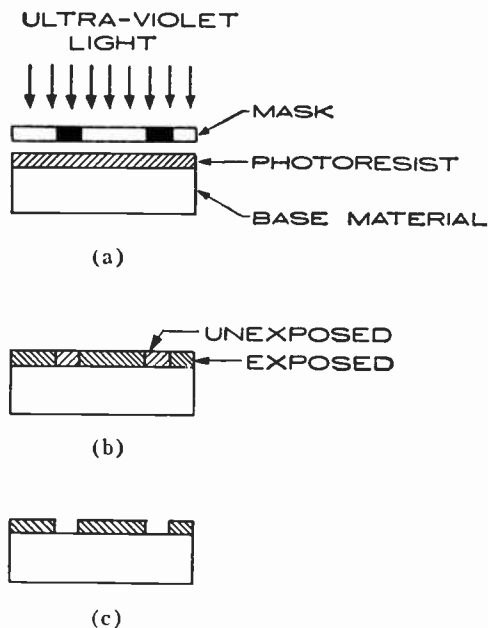


Fig. 1—Steps in the use of photoresist: (a) during exposure, (b) after exposure, (c) after developing.

Photoresist can be used as the basis of a number of processes. The photoresist can be formed into an accurate, detailed mask on the surface of a base material. Then thin layers can be either added to or subtracted from the exposed surface regions of the base material. Evaporation, printing, electro-deposition or chemical deposition can be used to add material. Electro-etching or chemical etching can be used to remove material. Following any of these processes the protecting photoresist can be stripped off the surface of the base material. Methylene chloride is an excellent stripping solvent.

2) *Photographic Plates*: Present-day research in spectroscopy and astronomy demands the use of spectrographic plates with high resolving power, high contrast, maximum sharpness, and minimum granularity. A typical spectrographic plate can resolve 1000 lines per millimeter. Photographic plates with these properties are being used as master negatives for achieving the desired geometry in certain components of the micro-electronic assemblies.

B. Deposition Processes

1) *Vacuum Deposition*:⁶ Vacuum deposition requires the heating of a metallic or dielectric material in a vacuum until vaporization takes place. The vaporized material travels in all directions from the source and condenses on arrival at the first cool surface. A substrate on which it is desired to deposit a thin film is covered with a mask of the proper pattern. The mask can be a mechanical mask or a photoresist surface mask which remains in intimate contact with the substrate during the deposition process. Thus, selected areas of the substrate can be coated with a thin metallic or dielectric film.

2) *Electro-Deposition*: Electro-deposition is a process by which metal is electrically plated onto a conducting substrate. The substrate is made the cathode; the metal to be plated, the anode; and both are immersed in an electrolytic bath. Electro-deposition is a wet process and this requires the use of electrolyte-proof masks in achieving selective deposition of various geometric metallic patterns.

Electro-deposited metals and alloys have been used for ohmic and rectifying contacts on semiconductor materials. Alloys of the III-V elements of the periodic table, such as indium and antimony, as well as the noble metals have been electro-deposited to produce the active electrode areas of transistors and diodes.

3) *Chemi-Deposition*: Using strictly chemical means a metal may be deposited from solution onto either a conducting or an insulating substrate without the use of an electric current. One process consists of sensitizing the substrate by dipping it into several metal salt solutions and then depositing copper onto the sensitized substrate from another solution which is believed to be a modified Fehling solution with formaldehyde as the reducing agent.⁷ Gold may be chemi-deposited in a bath by allowing it to displace another metal.

Gold and nickel chemi-deposited films have been used as both protective and ohmic areas on metallic and semiconducting substrates.

4) *Screen-Printing*: Screen-printing is a commonly used process for the application of inks so as to form particular geometric patterns on substrates.⁸ The process of screen-printing an ink onto a substrate consists essentially of the following steps: 1) preparing a screen so that areas through which ink is to pass are clear while other are masked, 2) accurately positioning the substrate below the screen, and 3) pushing the ink down into and across the mesh of the screen with a squeegee so that the portion of the screen beneath the edge of the squeegee momentarily touches the substrate and transfers ink to it as the squeegee passes over.

⁶ L. Holland, "Vacuum Deposition of Thin Films," John Wiley and Sons, Inc., New York, N. Y.; 1956.

⁷ E. D. Olson, "Fine line etched wiring," *Electronic Design*, vol. 7, pp. 38-41; February 4, 1959.

⁸ C. Brunetti and R. W. Curtis, "Printed Circuit Techniques," NBS Circ. 468; November 15, 1947.

There are both cutting and photographic methods by which the stencil-like patterns may be applied to the screens, but the photographic methods are required in microminiature circuit work because patterns having great detail can be obtained from them.

The photographic methods employed primarily in the DOFL work involved 1) making a contact print on Kodak Ektagraph film from a positive of the desired pattern, 2) developing the film in water to wash out the pattern area, 3) transferring the film and its plastic backing to a stainless steel screen, and 4) stripping away the plastic backing.

Screen printing can be used for the formation of conductive wiring, resistors, protective coatings, and coils.

C. Mechanical Forming to Size and Shape

1) *Scribe and Break*: Ruling of thin ceramic or germanium material bonded to cellophane with paraffin is done with a tungsten carbide or diamond tipped scribe. Pressure is exerted by drawing the cellophane over a sharp edge thus breaking the material at the scribed lines. The paraffin is removed by dissolving in an organic solvent which separates the isolated scribed areas.

2) *Molding*: Ceramic materials may be compression molded into various shapes from powdered constituents and subsequently sintered at high temperature into finished pieces. This would be a logical method by which to fabricate large numbers of substrates in a desired geometry.

3) *Sawing*: Hard materials such as ceramics, glasses, and semiconductors can be shaped to a desired form by sawing with diamond-impregnated bronze, silicon carbide, or alumina saw blades. In the production of micro-electronic assemblies all the above methods have been used; however, diamond-impregnated blades are preferred because of their resistance to wear. Circular as well as rectangular shapes can be fabricated with ease. An extension of the above mentioned technique is wire sawing. In this case a wire is impregnated with an appropriate abrasive to perform the desired sawing operation.

4) *Air Abrasive Technique (Sand Blasting)*: In air abrasion a hard abrasive material of small particle size is propelled by air under pressure against the substrate and abrades the substrate away. Air abrasion methods are usually used to form holes and irregular shapes in hard, brittle materials such as those used as substrates in DOFL microelectronic assemblies. Some of the abrasive materials used are sharp sand, alumina, and silicon carbide. Selective abrading can be achieved by forming masks of resilient materials such as pressure sensitive plastic tape. Since the abrasive will rebound from the tape, only the area of the substrate under a hole in the tape will be abraded. Preformed masks of hard materials such as glass, ceramic, and metals can also be used if these masks abrade at a slower rate than the substrate.

5) *Ultrasonic Machining*: Ceramic bodies can be drilled or formed to desired shapes by ultrasonic ma-

chining. A magnetostrictive transducer is used to drive at ultrasonic frequencies a soft metal tool in an abrasive slurry. The action of the tool and abrasive on the material being shaped causes micro-chipping of the hard ceramic. The major advantage of ultrasonic machining is that inexpensive tools can be made of soft materials such as brass which can be easily machined into complex shapes.

D. Conductive Adhesive

In the fabrication of printed circuit assemblies, it is often necessary to provide for the electrical and mechanical connection of various materials. The use of solder has sometimes proved inadequate in making connection to the thin printed wiring patterns on microminiature circuit wafers. Not only does the connection usually have low mechanical strength but the thermal shock caused by the soldering iron may even crack the wafers.

The recent development of a conductive plastic adhesive⁹ has allowed the fabrication of mechanically strong connections (tensile strength: ca. 5000 psi) with low resistivity (resistivity: <1 ohm-cm). The adhesive is predominantly a liquid epoxy resin heavily loaded with silver flake. Tiny quantities of the adhesive can be applied to specific positions on a circuit plate from the tip of a glass needle. A cure of one hour at 100°C sets the resin and yields a finished connection.

III. STEPS IN FABRICATION OF CIRCUIT WAFERS

The four different types of circuits that have been fabricated in a 2D form so far in the program are amplifier stages, binary counters, NOR circuits, and free running multivibrators. Most of the effort has been put on the binary counter and NOR circuits because of their usefulness in counting and timing operations of particular interest to DOFL.

In order to illustrate how the various techniques just reviewed are utilized, the fabrication steps for a binary counter stage will be described. The particular binary counter stage chosen contains fourteen component parts and is shown schematically in Fig. 2. In certain of the fabrication steps alternate approaches are available and will be referred to if they seem to have any advantages.

A. Substrate

The choice of substrate was based on obtaining a dimensionally stable, easily formed, rugged material which could stand up under the high temperatures encountered in the fabrication processes. The types of substrate used in the DOFL microelectronics program have been steatite, forsterite, alumina, quartz, and barium titanate. Barium titanate was used in an early unit where the substrate was required to be the dielectric for a capacitor. Steatite, a material that was readily available, was used in most cases.

⁹ T. J. Kilduff and A. A. Benderly, "Conductive adhesive for electronic applications," *Elec. Mfg.*, vol. 61, pp. 148-152; June, 1958.

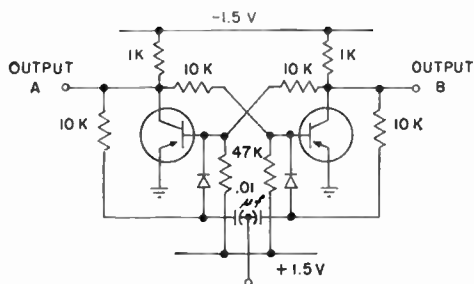


Fig. 2—Schematic diagram of binary counter circuit.

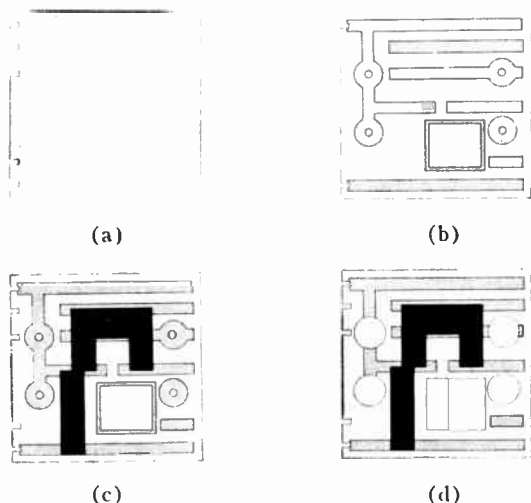


Fig. 3—Assembly steps for DOFL-2D binary counter wafer: (a) blank substrate, (b) after application of conductive wiring, (c) after application of resistors, and (d) after sandblasting holes and recesses.

The steatite was purchased as large wafers 0.020 inch in thickness. The wafers were stacked and cut in quantity with a diamond saw to 0.500 by 0.500 inch. While the wafers were still stacked, five grooves were cut into one side of the stack and hence into each wafer. The wafers, shown in Fig. 3(a), were separated and from this point on were treated individually.

B. Conductive Wiring

The conductive wiring was screened onto the steatite wafers as a silver ink [Fig. 3(b)]. This silver ink, available commercially, contained, in addition to silver and volatile solvents, glass frit which provided the permanent bond to the ceramic wafer when the wafer was subsequently fired in an oven for 30 minutes at 700°C. The conductivity of this ink is described as "excellent" in the manufacturer's literature. This means that using a simple ohmmeter one would measure zero resistance between any two connected points on the wafer. More precise measurements have not been made in these laboratories.

Oxidation of the silver wiring will not have a detrimental effect on the circuit performance since silver oxide is also a good conductor of electric current. Continuous operation of a circuit employing silver wiring in a high humidity environment could lead to circuit mal-

function due to silver migration.¹⁰ Since it is contemplated to seal hermetically completed circuit assemblies, the high humidity environment will not be encountered.

Conductive wiring could also have been applied using other metals and other techniques such as vacuum deposition and chemical deposition.

C. Resistors

Resistors were screened at their required positions on the wafer and overlapped appropriate areas of the conductive wiring as shown in Fig. 3(c). For the binary counter circuit a resistance ink which yielded a resistivity of approximately 10,000 ohms/square was used. It was composed of 100 parts by weight (pbw) liquid epoxy resin, 7 pbw boron trifluoride piperidine complex (BF₃P), and 15 pbw carbon black. It should be noted that no solvents were used in the ink formulation. It was believed that the elimination of volatile solvents and hence elimination of the need to control the amount of solvent in an ink would result in more reproducible resistors. Since BF₃P was a latent curing agent, it had a very slow reaction rate with the epoxy resin until a critical temperature was reached. For BF₃P this critical temperature was 150°C. For this reason the ink, once prepared, had a usable life at room temperature of several months. Resistors were cured for one-half hour at 225°C.

This installation has conducted research on resistance inks for many years. This research has been directed toward inks for use with an injection molding apparatus for printing resistors.¹¹ Therefore all the data available at the present time are based on injection molded printed resistors. The injection molding apparatus was not employed in the present work because in early stages of the program the greater versatility offered by the screening approach was desirable. The ink used, however, was the same, and, although some difference in the properties of the resistors resulting from different methods of application are expected, the data for the injection molded resistors are listed in Table I.

TABLE I
PROPERTIES OF PRINTED RESISTORS*

Resistance-temperature characteristic	±5% average change in resistance from value at +25°C between -55°C and +85°C
Voltage coefficient	0.0035%/volt
Temperature cycling, 5 cycles between -55°C and +85°C	-0.3% average change from initial value

* Coated resistors in the range from 10 kilohms-100 kilohms.

¹⁰ J. C. Williams and D. B. Herrmann, "Surface resistivity of non-porous ceramic and organic insulating materials at high humidity with observations of associated silver migration," IRE TRANS. ON RELIABILITY AND QUALITY CONTROL, vol. RQC-6, pp. 11-20; February, 1956.

¹¹ R. S. Marty, E. M. Davies, and P. J. Franklin, "A new injection-molding process for printed resistors," Elec. Mfg., vol. 55, pp. 56-63; January, 1955.

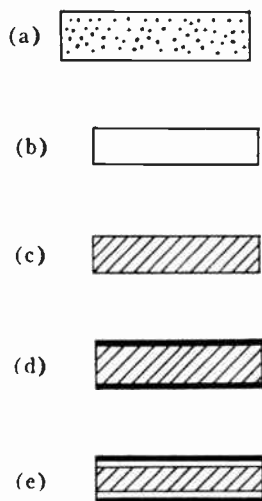


Fig. 4—Steps in fabrication of layered capacitor; cross-section schematics: (a) original film, (b) sintered, (c) reduced, (d) silver paint, (e) fired.

In some cases resistors as screened were wide of nominal value by as much as 60–80 per cent. In these cases low values were abraded with a scribe until they were within 20 per cent of nominal value. Resistors which had resistances higher than nominal were subjected to a hot air torch until their resistance also fell within the ± 20 per cent tolerance. It should be noted that resistance spreads on printing proved much higher when the resistors were screened than when they were injection molded.

Resistors could also have been applied using vacuum deposition and perhaps even chemical deposition in conjunction with appropriate pattern masks.

D. Capacitors

Extremely high capacitance per unit volume was required of the capacitors used in the binary counter stage. The particular units used were made by special processing of a ceramic material predominantly barium titanate in composition.¹²

The fabrication was started by preparing a mixture of powdered titanates, trace additives, resin binder, and solvents. This mixture was formed into a wet film approximately 0.024 inch in thickness and then dried by a heat lamp to produce a flexible sheet 0.012 inch in thickness. The sheet was cut into 7/16-inch squares for later case of processing [Fig. 4(a)].

The squares were then fired in three successive steps. The first firing sintered the material into a hard ceramic [Fig. 4(b)]. The next firing in a dry hydrogen atmosphere chemically reduced the barium titanate to render it conductive [Fig. 4(c)]. The third firing was done briefly in air and produced a thin surface film of unreduced barium titanate. During this third firing

¹² L. N. Maxwell, D. M. Freifelder, and P. J. Franklin, "Layerized high-dielectric-constant capacitors," *Electronic Design*, vol. 7, pp. 70-74; March 4, 1959.

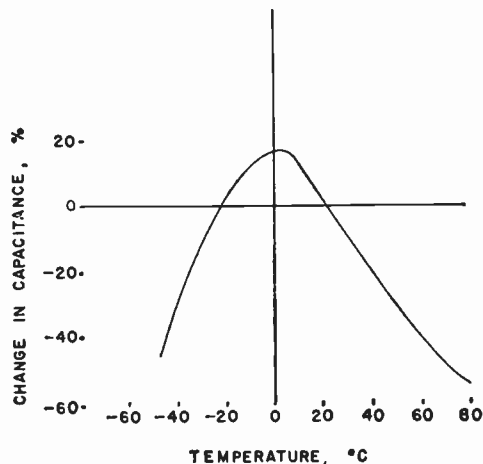


Fig. 5—Capacitance-temperature characteristic for a layered capacitor.

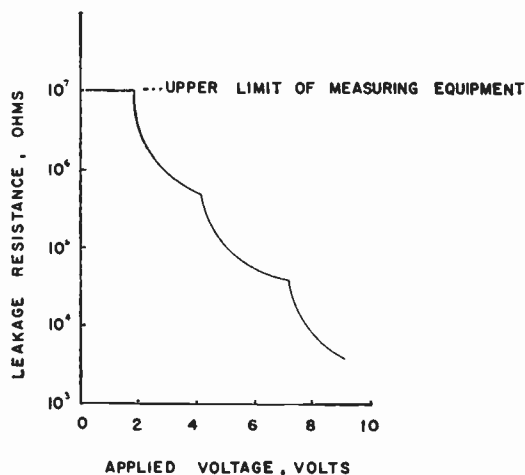


Fig. 6—Leakage resistance as a function of applied voltage for a layered capacitor.

step, silver electrodes, which had previously been printed or screened on the surface [Fig. 4(d)] were also processed. The squares were finally cut to 0.1 by 0.1 inch using scribe-and-break techniques.

The resulting capacitors had a cross section as shown in Fig. 4(e). The major thickness of the body was contributed by the conductive, reduced barium titanate; on each side of this was a thin region of unreduced barium titanate which was the dielectric of the capacitor. The silver electrodes formed one plate of each of the two dielectric regions at the surfaces. The complete capacitor was made up of two capacitors in series. The dimensions of the 0.01 microfarad capacitors used in the binary counter were 0.1 by 0.1 by 0.008 inch in thickness.

Electrical characteristics of a typical unit are shown in Figs. 5 and 6. Fig. 5 shows that the capacitance changes quite significantly with temperature. This amount of change, however, is tolerable in the circuit under consideration. Modifications in the capacitance-temperature characteristic can be accomplished by variations in the ceramic constituents in the formulation. Fig. 6 shows that the leakage resistance decreases

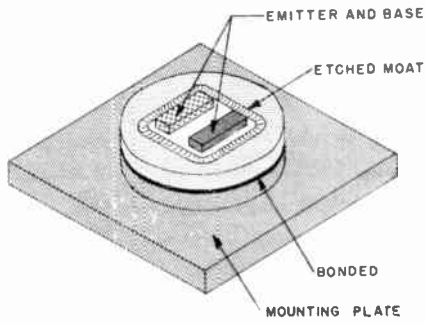


Fig. 7—Diffused base transistor bonded to a mounting plate.

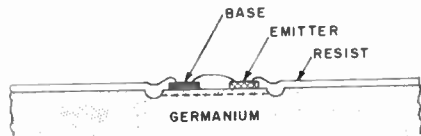


Fig. 8—Cross-sectional view of diffused base transistor ready for insertion into wafer.

sharply above 2 volts and therefore these capacitors are very low voltage types having principle use in transistor circuitry.

The capacitors were incorporated into the circuit wafer in the following manner. Following printing and curing of the resistors the wafer was sandblasted so that it contained four round holes which went through the entire plate and two overlapping square recesses, one on each side of the plate, which went half way through the plate [Fig. 3(d)]. Conductive adhesive was employed to attach two capacitors back to back but slightly displaced so that one capacitor would seat into each recess on the wafer. A ribbon of conductive adhesive was used to make connection from the silver wiring on the surface of the wafer to the bottom of the recess. The capacitors were then placed into the recess on top of the adhesive and the adhesive was cured. Epoxy resin was next used to fill in the space around the capacitors and to cement them ruggedly in place. Vacuum deposited aluminum or painted conductive adhesive was used to make connections from the top of the capacitors to the printed wiring.

E. Transistors

Both diffused base¹³ (mesa) and alloy types of germanium transistors have been used in the DOFL microelectronic assemblies. Germanium diffused units were chosen initially because of their ease of fabrication using photolithographic techniques¹⁴ and the natural extension of these techniques to permit connections to be made from active areas on the transistor to the printed wiring on a wafer by vacuum deposition. Fig. 7 shows

¹³ C. A. Lee, "A high-frequency diffused base germanium transistor," *Bell. Sys. Tech. J.*, vol. 35, pp. 23-34; January, 1956.

¹⁴ J. W. Lathrop, J. R. Nall, and R. J. Anstead, "Two dimensional transistor packaging," *Electronic Design*, vol. 7, pp. 50-52; February 18, 1959.

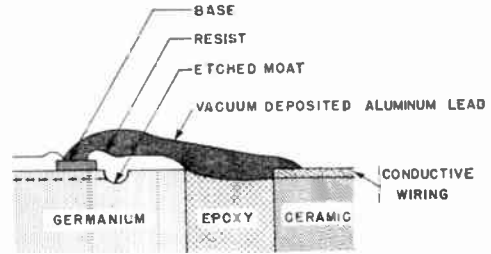


Fig. 9—Cross-sectional view of a diffused base transistor inserted in a ceramic wafer showing a deposited aluminum lead.



Fig. 10—Photomicrograph showing diffused base transistor and deposited aluminum leads.

the electrode configuration and the mounting plate of a diffused base transistor. The germanium die pictured was circular in shape, measured 0.040 inch in diameter and 0.005 inch in thickness and was soldered or bonded to the mounting plate which became the collector contact. Earlier dies were square and measured 0.045 by 0.010 inch. The areas of the base and the emitter contacts were each 0.004 by 0.012 inch and were located inside an etched moat 0.0005 inch across and enclosing an area 0.014 by 0.017 inch. The top of the die was coated with photoresist and exposed and developed through the proper pattern to give unprotected regions over the electrode areas as shown in Fig. 8. The remaining photoresist acted as a protective coating over the active areas of the device. The transistor was next inserted into a hole in the ceramic wafer and the area between the transistor and ceramic wafer was filled with an epoxy resin which provided mechanical bonding of the transistor to the wafer and an insulating bridge over which leads would be evaporated. The leads from transistor electrodes to silver wiring on the wafer were formed by vacuum deposition of aluminum through a fine-line, engraved, mechanical mask. Fig. 9 shows in a cross-sectional view such a vacuum deposited aluminum lead. Fig. 10 is a photomicrograph of a top view of the connections.

Typical small signal characteristics of the DOFL diffused base transistor are given in Table II.

TABLE II
TYPICAL CHARACTERISTICS OF DOFL
DIFFUSED-BASE TRANSISTORS

Common base		Common emitter	
h_{ib}	38.9 ohms	h_{ic}	660 ohms
h_{rb}	170×10^{-6}	$h_{fc}(\beta)$	15.9
h_{ob}	2.1×10^{-6} mhos	h_{oe}	29.5×10^{-6} mhos
h_{fb}	-0.935		
$(1-\alpha)$	0.065		
	I_{CBO} 1.2 microamperes		
	C_c 13.9 μf		
	f_{ab} approx. 48 mc		
	operating point $I_E = 1$ ma, $V_C = -5$ volts		

In low frequency circuits such as for the binary counter being described, the high frequency cutoff of the transistor was hardly a limiting parameter. It should be noted, however, that diffused base transistors have been made at DOFL with an f_{ab} of 250 mc. This was accomplished by simply reducing the areas of the base and emitter electrodes to 0.002 by 0.006 inch.

Commercial germanium alloy junction transistors of the $p-n-p$ type, normally intended for audio frequency amplifier service, were removed from their enclosures and integrated in microelectronic wafer assemblies in both nonhermetic and hermetically-sealed forms. The alloy junction device has emitter and base electrodes on one side of the die and the collector electrode on the opposite side. Fig. 11 shows an alloy junction unit incorporated in a ceramic wafer. A recess the proper size to receive the transistor was machined in the ceramic wafer already containing the other component parts. The transistor was coated with photosensitive lacquer except over its electrodes, placed into the recess, and potted with epoxy resin in a manner similar to that used in the incorporation of the diffused base transistors. The epoxy acted both as a mechanical bond and as an insulating bridge over which wiring from the transistor electrodes to the silver wiring on the ceramic wafer was made using conductive adhesive.

The hermetically-sealed transistor which was incorporated in a wafer is illustrated in Fig. 12. The method of incorporation was basically the same as for the non-hermetically sealed unit except the emitter and base side of the transistor was placed in the machined recess. Also, the ceramic was metallized in the following areas: 1) near the emitter and base electrodes, and 2) around the perimeter of the rectangular recess. Soldering to the metallized ceramic hermetically sealed the transistor. Units fabricated in this manner were exposed to ammonia vapor for several hours and to an atmosphere of 95% humidity at 71°C for a period of 32 hours without degradation of transistor characteristics. Ammonia vapor is one of a group of gaseous ambients frequently used to test semiconductor surfaces. Ammonia causes a rapid change in the characteristics of unprotected transistors.

F. Diodes

Diffused germanium diodes were the only type of

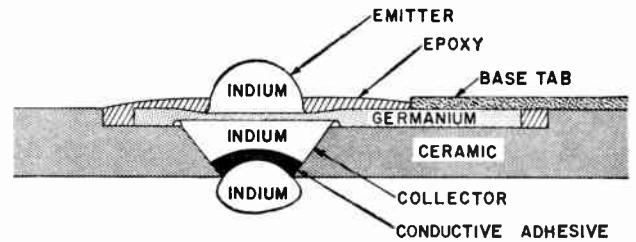


Fig. 11—Cross-sectional view of an alloy transistor inserted in ceramic wafer; connections from emitter and base electrodes not shown.

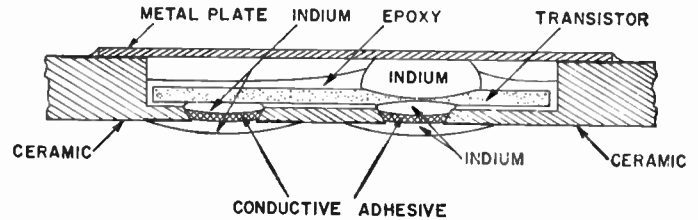


Fig. 12—Cross-sectional view of an alloy transistor hermetically sealed in a ceramic wafer.

diodes used in the microelectronic wafer assemblies and were constructed in a manner similar to that used for the construction of diffused base transistors. The diode was a simple device to construct because it contained only one junction. The junction was formed by solid state diffusion on p -type material forming an n -type surface and creating a $p-n$ junction. The device was resist-masked and gold was electro-deposited in a circular pattern and alloyed on the n -type surface to produce an ohmic contact. An etched moat was formed by photoengraving techniques and this operation completed the device. The diode was photolithographically processed so as to leave uncoated an ohmic area over the gold. The diode was next inserted into the wafer in the same manner as the diffused base transistor, using epoxy for a mechanical bond and to form an insulating bridge between the wafer and the die. A mechanical mask formed by photoengraving techniques was used to define the area for the vacuum deposition of aluminum which made connection between the printed wiring on the ceramic wafer and an electrode on the diode. Fig. 13 shows a photomicrograph of a diode with the evaporated leads.

Reverse current of a typical diode at -3 volts was 2.0 μa ; at -10 volts was 2.2 μa . The forward current at 1 volt was 150 ma.

G. Lead Wires

The lead wires from each 2D microelectronic wafer were nickel metal approximately 0.015 inch in diameter by 0.5 inch in length. Jigs were made for mounting all five lead wires in one operation. This operation involved coating of the ends of the wires with conductive adhesive, inserting the coated leads into the slots provided in the wafer, and curing the adhesive at 100°C for one hour.

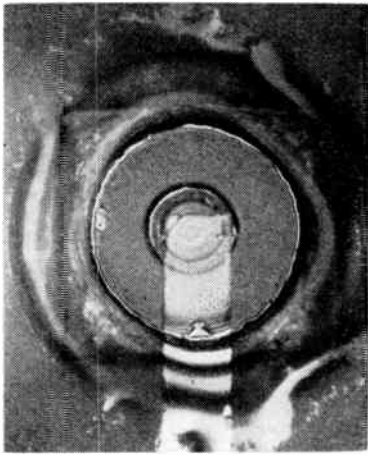


Fig. 13—Photomicrograph of a diffused diode showing vacuum deposited aluminum lead.

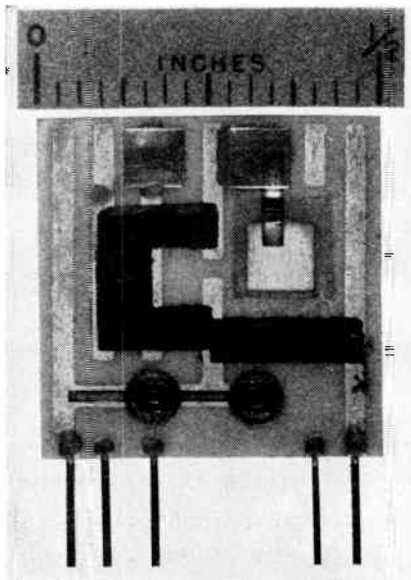


Fig. 14—DOFL-2D binary counter wafer.

II. Completed Wafers

Fig. 14 shows a completed 14-component part binary counter wafer. Vacuum deposited aluminum was used extensively in this unit for connection of the DOFL diffused-base transistors, diodes and the layerized high-dielectric constant capacitors to the silver wiring. It should be noted that the layout of the binary counter wafer was symmetrical and component parts were attached to both sides of the wafer.

Fig. 15 shows a completed 5-component part NOR wafer. This unit employs a commercial alloy transistor die; all connections were accomplished by conductive adhesive. The steatite substrate used for the NOR circuit measured 0.250 inch on a side and 0.020 inch in thickness.

IV. PACKAGING

A. Interconnection of Wafers

The interconnection problem was severe because of

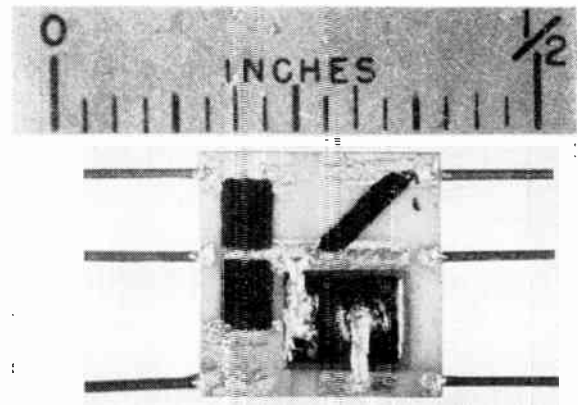


Fig. 15—DOFL-2D NOR circuit wafer.

the close spacing between leads that existed when the wafers were stacked to obtain maximum parts density. One approach¹⁵ was to pot the stacked assembly in plastic with all the leads extending from a common face. The leads and face were then milled off exposing the interconnection points as cross-sections of the lead wires flush with the face of the plastic. A metallic film pattern which interconnected the ends of the leads was then achieved by the chemical deposition of copper over the whole face followed by selective removal to leave the pattern. Fig. 16 shows how the deposited copper film was used to interconnect five NOR circuits into a half-adder. Fig. 17 shows the block diagram for such an assembly. Interconnection has also been achieved by screening silver ink on the face of the potted stack.

Another approach to wafer interconnection is to use welding techniques such as those used to make subminiature vacuum tubes.

B. Sealing

Probably the most questionable part of the 2D process is the ability to protect adequately the transistors and diodes from contamination and subsequent deterioration. Although it was never intended to operate any of the assemblies exposed to an external environment, there still remain the important questions: 1) at what level of fabrication should the protection be applied; and 2) what degree of sealing is necessary?

The most satisfactory and least expensive type of sealing would be to enclose a group of wafers in a common hermetically tight can. The problem still remains of contaminants being trapped in the can or generated from the other parts. The opposite extreme is to seal individually each transistor and diode in a hole in the ceramic plate as described previously. This reduces the advantage of the 2D approach by complicating the wafer fabrication.

An intermediate level of sealing is also possible. Each wafer could be enclosed in a separate can of metal,

¹⁵ N. J. Doctor and E. L. Hebl, "Interconnecting microminiature modules," *Electronic Design*, vol. 7, pp. 34-37; February 4, 1959.

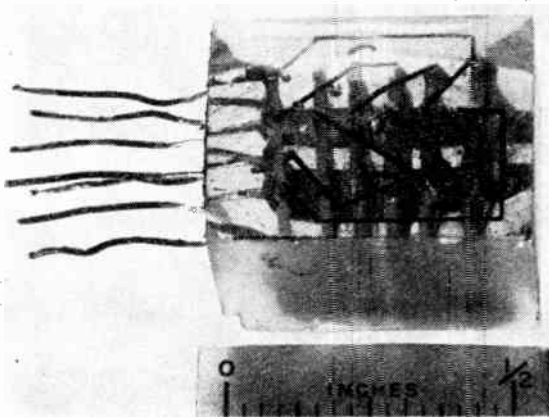


Fig. 16—5-NOR half-adder interconnected with deposited copper leads.

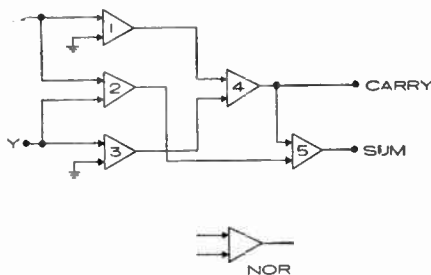


Fig. 17—Block diagram of 5-NOR half-adder.

glass, or ceramic. However, the semiconductors used in the 2D microelectronic assemblies have had some superficial protection against the ambients that will be encountered before final hermetic sealing. Considerable study of photoresist indicates that this material does not degrade the device and offers interim protection. Field effect measurements, as well as surface recombination measurements, have given considerable insight to the solution of surface stability.¹⁶

Up to the present time three wafers containing a total of 36 component parts have been sealed in a can 0.625 by 0.550 by 0.150 inch. Glass to metal seals were used in the header and it was fastened to the can by epoxy resin. A check on a vacuum leak detector showed no leakage. This is certainly not a long-term evaluation, and since plastics are known to have finite vapor transmission future enclosures are to be soldered or welded.

C. Incandescent Lamp

One desirable feature in counting and timing circuits is to have a means of visually indicating the binary state of a flip-flop or binary counter stage. Many approaches have been used including neon lamps, special thyratron tubes, and meters. A successful approach in our program has been to use incandescent lamps. Ordinary 2-volt, 60-

¹⁶ J. M. Stinchfield and O. L. Meyer, "The stability of semiconductors in microminiature assemblies," *Electronic Design*, vol. 7, pp. 42-45; February 18, 1959.

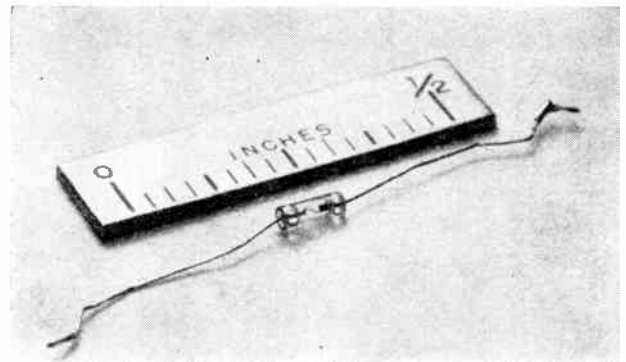


Fig. 18—Microminiature incandescent indicator lamp.

ma, radio pilot lamps can be operated from low power transistors used as switches. The size of these lamps, however, leaves much to be desired from a miniaturization standpoint.

DOFL therefore built an incandescent lamp (Fig. 18) 0.100 inch in length and 0.035 inch in diameter.¹⁷ The tungsten filament, 0.00022 inch in cross-section drew 35 ma at 1 volt and produced a light easily visible from any point in a normally lighted room.

The same filament was also packaged in a glass bulb 0.200 inch in length and 0.135 inch in diameter. In this latter version both leads projected from the base of the bulb.

One of the smaller lamps has been flashed over 70,000,000 times at approximately 10 times per second and is still operating.

V. CIRCUIT AND SYSTEM ASPECTS

The circuit and system problems are not directly concerned with the physical technologies used to fabricate operating circuits. The circuit-system designer is concerned with the interplay between circuits, fabrication techniques, and the eventual use of assemblies to make operating systems. The mutual interdependence of these various phases puts restrictions on the circuit designer and later on the systems designer. Important problems in the circuit and system area include 1) the search for and the design of an optimum number of basic electronic circuits that can be produced cheaply in a microminiature form and be easily interconnected; 2) the search for the best arrangement of the basic circuits to perform various specific functions, *e.g.*, addition, multiplication, counting, and timing; and 3) the search for new electronic systems made possible by the advances in microelectronics.

Work at DOFL on the first two problems is described in the immediately following paragraphs. The third problem area has hardly been scratched and will be briefly touched on in Section VI.

¹⁷ D. J. Belknap and L. R. Crump, "Miniature incandescent indicator lamps," *Electronic Design*, vol. 7, pp. 78-81; March 4, 1959.

A. Circuit Design

The search for an optimum number of basic circuits is a continuing struggle between standardization and optimization. At one extreme is the use of a single type circuit to perform all functions in a system. Such an approach generally is somewhat wasteful in terms of total electronic circuitry. At the other end of the scale is the use of many specialized circuits, each optimized to the specific problem at hand. This approach puts a greater burden on designers, manufacturers, and reliability engineers to prove out many types of assemblies.

Studies at DOFL have resulted in a family of circuits¹⁸ useful in low frequency counting and timing applications. Of the ten types resulting from these studies, two circuits, the NOR and the binary counter, were chosen for extensive fabrication in the 2D form.

The circuit design area followed several guiding principles:

- a) Circuits should be thoroughly understood and well behaved.
- b) Circuits should be as simple as possible to reduce number of parts.
- c) Since high-frequency performance was not required, upper operating frequency should be traded for broader margins.
- d) Circuits should be designed for a low operating power level.
- e) Designs should permit broad tolerances for individual component parts.
- f) The number of leads and interconnections should be minimized.

The binary counter and NOR type circuits have been thoroughly studied both theoretically and experimentally in order to arrive at circuit designs whose performances are well understood. Complete design equations are available^{19,20} for both of these circuits and can be used for rapid design to meet a given set of specifications. For the circuits described the following conditions were assumed:

- a) Two power supplies, each consisting of a single cell, so as to give both positive and negative 1.5 volts with respect to a common ground would be used.
- b) No speed-up capacitors would be used in either circuit. This would mean greater simplicity although at a sacrifice in maximum operating frequency.
- c) The circuits would be designed for audio type low frequency transistors. This would permit a wide choice of transistor types and a circuit design with wide margins.

¹⁸ T. A. Prugh, "Standard transistor switching circuits," *Electronic Design*, vol. 7, pp. 46-51; March 18, 1959.

¹⁹ E. L. Cox, "Designing a transistor NOR circuit for minimum power dissipation," *Electronic Design*, vol. 7, pp. 42-45; March 18, 1959.

²⁰ P. Emile, Jr., "Design of a two-transistor binary counter," *Electronic Design*, vol. 7, pp. 38-42; March 18, 1959.

- d) Germanium types would be chosen at a sacrifice in maximum operating temperature because of their wide availability, low cost, and excellent performance as switches.
- e) Transistor-resistor logic (TRL)²¹ would be used where possible in preference to direct-coupled transistor logic (DCTL) to reduce the number of transistors required.

Schematic diagrams of the two circuits are shown in Figs. 2 and 19. The NOR circuit²² has the advantage of extreme simplicity; in the two-input NOR used so far at DOFL only four resistors and one transistor are required. The NOR is a basic building block which can be used to perform all the Boolean logical operations of AND, OR, and NOT. The two-input NOR has similar properties to the two-grid pentode gate. The NOR can be designed for more than two inputs for greater utility, but the circuit design becomes more critical as the number of inputs increases.

For the present state of development the smallest number of inputs (two) was therefore chosen. The NOR can also be used to perform a binary counter function. Fig. 20 shows such a binary counter utilizing eight NOR circuits. This type counter²³ requires no short-time memory as provided by the capacitors in the ordinary binary counter of Fig. 2.

The specific binary counter circuit²⁴ shown in Fig. 2 was chosen because of its good triggering characteristics which permitted a design with known safety factors. Both dc and ac performance were readily calculated. This circuit configuration does not give good high frequency performance because of the streamlining operation used to reduce the number of parts. Additional diodes and speed-up capacitors can be added if necessary in future models to improve this characteristic.

B. Logical System Design

The design of a system involves the choice of the interconnection network between the elemental building blocks. Specific functions, such as addition or multiplication, can generally be accomplished in a number of ways. One problem then is to determine the approach which uses the fewest elemental building blocks. Two methods have been used for the solution of this problem. The first was to work with Boolean expressions and to simplify mathematically. The second method, which has proven very useful for complex functions, was to use an experimental machine having a number of basic circuits mounted in a rack with their input and output

²¹ M. W. Marcovitz and E. Seif, "Analytical design of resistor-coupled transistor logical circuits," *IRE TRANS. ON ELECTRONIC COMPUTERS*, vol. EC-7, pp. 109-119; June, 1958.

²² W. D. Rowe, "The transistor NOR circuit," 1957 WESCON CONVENTION RECORD, pt. 4, pp. 231-245.

²³ W. H. Ware, "The logical principle of a new kind of binary counter," *PROC. IRE*, vol. 41, pp. 1429-1437; October, 1953.

²⁴ G. D. Bruce and J. C. Logue, "An experimental transistorized calculator," *Elec. Eng.*, vol. 74, pp. 1044-1048; December, 1955.

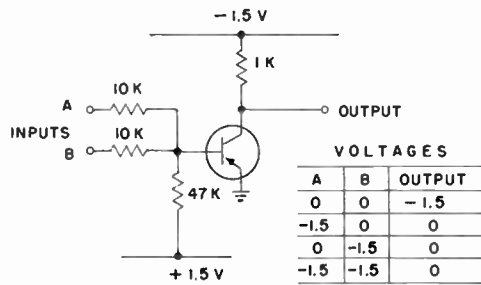


Fig. 19—Schematic diagram of NOR circuit.

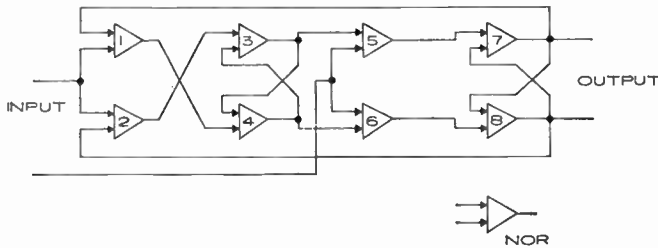


Fig. 20—Binary counter circuit built using 8 NOR circuits; push-pull input and output.

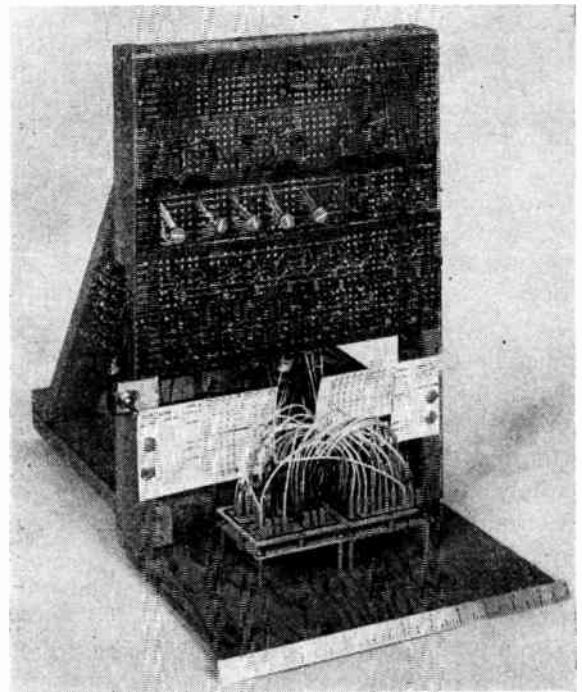


Fig. 21—Boolean Operation Logical Device, BOLD.

terminals appearing at a plug board. This setup permitted rapid trial of different schemes.

The first model of such a flexible system was built incorporating 10 binary counter stages, 20 NOR circuits, and a pulse-repetition-frequency generator operating at approximately 1 kc. A view of the equipment is shown in Fig. 21. This equipment is titled BOLD for *Boolean Operation Logical Device*. A typical problem studied on BOLD was the search for the smallest number of NOR and counter circuits required to give a 5-stage ring counter function. Another example of the use of BOLD was to confirm the logic of the NOR-type binary counter shown in Fig. 20.

C. Test and Evaluation

The design and testing of circuits and systems become more difficult when equipment is operated in a completely inaccessible location or the size of the parts is so small as to rule out conventional voltage measurement and waveform checks at key points in the system. In these electronic systems, performance must be built in and confirmed at the time of manufacture. The problem is similar to that faced in the design of underwater telephone repeaters.²⁵ Adequate tests had to be developed which could not only check that a unit was good at the time of test but also confirm that the unit was stable and thus would not drift out of tolerances during its intended operating life.

The individual circuit wafers offered similar problems. For example, the 2D binary counter circuit had only five external connections, three of which were non-sig-

nal. Thus, only one input lead and one output lead were available for test and evaluation. The possibility still existed of using a small test probe on the faces of the $\frac{1}{2}$ -inch square wafers. However, as the wafers get smaller this type of probing will become more difficult even with a micromanipulator.

One approach to the problem was to make careful studies of the output voltage waveform of the binary counter. Details of the waveform were carefully noted both for a normal circuit and for a circuit containing various out-of-tolerance component parts. The method was similar to the use of an electrocardiograph in diagnosing the condition of a human heart.²⁶ So far only an elementary attempt has been made to check out the completed wafers by such a waveform study. Further effort will be necessary to devise an automatic means of comparing the waveform of a newly manufactured unit with an acceptable waveform.

A typical collector waveform is shown in Fig. 22 for two of the binary counter stages. The upper waveform was obtained from a properly fabricated unit whose parts were within tolerances. The lower waveform was from a unit which divided properly but had a faulty part. In this case, one of the steering diodes had a bridging resistance which caused a low impedance in the reverse direction of the diode.

The NOR circuits were tested by checking each component part separately. This approach was possible with the NOR since no bridging paths exist for an in-

²⁵ H. A. Lewis, R. S. Tucker, G. H. Lovell, and J. M. Fraser, "System design for the North Atlantic link," *Bell Sys. Tech. J.*, vol. 36, pp. 29-68; January, 1957.

²⁶ O. H. Schmitt, "Where is medical electronics going? Part IV," 1956 IRE CONVENTION RECORD, pt. 9, pp. 107-112.

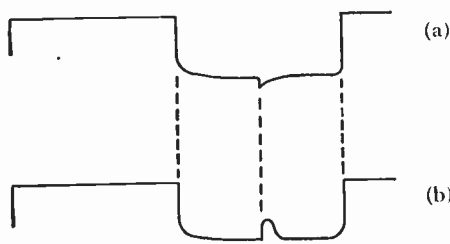


Fig. 22—Output wave forms from binary counters: (a) normal, (b) diode shunted with resistance.

dividual NOR not connected to power supplies. In these models the base connection of the transistor was accessible by probing during tests. The collector characteristic curves were taken on the transistors and the resistances of the resistors were measured with a digital ohmmeter.

The binary counter stages were difficult to check on an individual part basis because of the bridging paths that existed. As a result a complete operating circuit was checked. The circuits were set up under the standard conditions of drive, load, and power supply voltages shown in Table III. Operation was at room temperature for these tests. Each parameter was then changed in the direction shown until failure occurred. The value of the parameter at the point of failure was designated as the end point. Such a procedure gave quantitative data on each assembly for comparison with a properly operating circuit tested under the same conditions. Additional information on the circuit was obtained by determining the maximum operating temperature under the above standard conditions.

TABLE III
TEST PARAMETERS FOR BINARY COUNTER CIRCUITS

Parameter	Direction Changed	Standard Condition	Typical End Point
Collector supply voltage (V_{CC})	Reduced	1.5 volts (neg.)	0.8 volt
Base supply voltage (V_{BB})	Reduced	1.5 volts (pos.)	0.3 volt
Input driving frequency	Increased	1000 cps	12 kc
Load resistance	Reduced	Infinite (open)	700 ohms
Load capacitance	Increased	0	0.06 μ f
Input signal voltage	Reduced	1.5 volts peak to peak*	0.8 volt
Generator resistance	Increased	1000 ohms	3.5 K ohms

* The most positive part of the waveform was always held at ground potential.

The early binary counter units had three weak points: 1) some of the lead-wafer connections were mechanically weak; 2) certain of the alloy transistors had high I_{CBO} which seriously lowered the high temperature limit of the associated counter; and 3) several of the diodes had low resistance bridging paths which impaired their electrical operation. The first two problems were solved in more recently fabricated units while the third problem is still under study.

VI. CONCLUDING DISCUSSION

Two directions of work are being pursued at DOFL. One is the refinement of the present 2D approach to give low cost, high reliability wafers. The second is to intensify the search for techniques that integrate the component part functions into a "solid" circuit.

To further the objectives in the first direction, 250 2D NOR circuits made substantially like those produced at DOFL are being fabricated by a contractor. The larger quantities will permit a number of important studies such as

- Quality control of wafer performance.
- Evaluation of life under storage in various environments.
- More comprehensive consideration of interconnection and packaging problems.
- Fabrication of more complex systems to demonstrate feasibility of the whole program.

Some specific items that need consideration to help extend the usefulness of the 2D wafer design include the following:

- Incorporation of silicon devices.
- Resolution of the semiconductor surface protection problem.
- Consideration of injection molded or deposited film resistors.
- Design of an alloy germanium transistor whose physical shape is tailored to the 2D assembly.
- Design of an alloy germanium transistor with significantly (10:1) lower I_{CBO} than existing "low" power types.

The present techniques are not well suited for the fabrication of inductors and transformers or for circuitry which must operate at high power levels. New ideas are needed to permit extension of the 2D approach to these areas.

With regard to techniques more advanced than the 2D, one suggestion is the more complete use of thin films to produce operating circuits. Also under consideration at DOFL and elsewhere is the idea of starting with a block or slab of semiconductor such as silicon and converting various portions of the block to proper resistivity and type so as to achieve circuit function without resorting to separate resistors and capacitors. The NOR circuit, because of its simplicity and limited number of parts, is particularly attractive for the latter approach.

The 2D and other microelectronic approaches will affect significantly two groups of people in the electronics industry. The first group consists of the parts manufacturers; the second, the users and potential users of electronic systems. Complete operating circuits will of necessity have to be fabricated as integral sub-assemblies in a closely controlled sequence of steps. The assembler of electronic systems will then use these wafer

circuits as his basic building blocks instead of a multitude of independent component parts.

These very compact sub-assemblies have opened up new system possibilities for the user. More sophisticated systems can be substituted for existing less intelligent versions. More important still, electronic systems can be used in places and ways never before considered feasible. Missile and satellite applications are widely publicized examples of the former. The field of medical electronics would seem to be one of the most fruitful areas for microelectronic systems.

There are at least three categories of use in which microelectronics can benefit the field of medicine. First is the use of microelectronic systems inside the human body to aid in the diagnosis of ailments.²⁷ The second is for the replacement or repair of faulty human organs with electronic equivalents. External systems are already being used for such functions as control of breathing.²⁸ Internal systems may soon be possible. The third category is the study of complex electronic systems with a view toward learning more about the human nervous

²⁷ S. MacKay and B. Jacobson, "Pill telemeters from digestive tract," *Electronics*, vol. 31, pp. 51-53; January 3, 1958.

²⁸ G. H. Meyers and G. A. Saxton, Jr., "A servomechanism for automatic regulation of breathing," 1958 WESCON CONVENTION RECORD, pt. 5, pp. 224-230.

system.²⁹ Here the study may prove reciprocal and even more important knowledge may be gained from the study of the human system and permit better design of electronic systems.

The first phase of the DOFL microelectronics program described in the previous sections has demonstrated the feasibility of incorporating semiconductor devices as integral parts of operating wafers or assemblies. The next phase is to refine the various techniques so that commercial production becomes possible while retaining the advantages of small size, ruggedness, and low cost so necessary for success in the program.

VII. ACKNOWLEDGMENT

The work reported is the result of the efforts of many people at DOFL. In addition to those DOFL personnel whose papers have already been referred to in the text the authors wish to thank T. M. Liimatainen for many helpful discussions, A. Goodman for conducting evaluation tests, and G. B. Wetzel, M. L. Jones, and B. Mermelstein for their help in fabrication of experimental devices and wafers.

²⁹ J. von Neumann, "The Computer and the Brain," Yale University Press, New Haven, Conn.; 1958.

The Micro-Module: A Logical Approach to Microminiaturization*

S. F. DANKO†, MEMBER, IRE, W. L. DOXEY†, MEMBER, IRE, AND J. P. McNAUL†, MEMBER, IRE

Summary—Among the diverse activities of research and development of electronic parts at the U. S. Army Signal Research and Development Laboratory, particular emphasis has been placed on the evolution of a coherent and practical microminiaturization "system." An objective analysis of the several techniques suitable for a system base indicates that the advantages of each of the techniques could be combined logically into an approach called the "micro-module." The flexibility of this system to assimilate advancements in the state of the art, its compatibility to mechanized assembly, and its short- and long-range applicability as a universal construction system for all low and medium power circuits are reviewed. The weakness of "miniaturization for miniaturization's sake" is discussed frankly, and perspective is drawn for microminiature, circuits, for digital circuits and for general electronic usage. Latest experimental micro-modules are illustrated and their features are discussed. Several complex solid-state circuits are shown and quantitatively assessed to show the ultimate capabilities of this approach and its adaptability to the micro-module construction philosophy.

* Original manuscript received by the IRE, December 5, 1958; revised manuscript received, March 9, 1959.

† U. S. Army Signal Res. and Dev. Lab., Fort Monmouth, N. J.

INTRODUCTION

A SIGNIFICANT change has taken place during the past ten years in the military's attitude towards electronic research and development; there is a growing emphasis on development of new passive and active electronic parts to provide stimuli and new capabilities in system and equipment design. This metamorphosis reflects the growing recognition by research and development management that progress in basic parts over the past decade has significantly shaped the practicing electronic art as we know it today. The classical example of yesterday's audion as the "herald of a new electronics" has been complemented in the last decade by the transistor, the solar converter, printed circuits, ferrite memories, the thermal battery, and dozens of other component and material accomplishments, each opening intriguing new capabilities in the electronic art. Today, still newer component concepts

and techniques absorb our research and development interests—parametric amplifiers, cryogenic logic elements, semiconductors and pure crystals, atomic and molecular resonance devices; new materials born under superpressures and supertemperatures, new techniques for thermoelectric conversion. These, like the historical example, are the melodramatic “heralds for the next generation of electronics.”

Within the U. S. Army Signal Research and Development Laboratory at Fort Monmouth, N. J., these major component programs are carefully balanced against the systems and equipment activities whose work product reflects the prime mission of the Laboratory: to provide the Army with the best communications and surveillance capability that science and technology can provide. This intimate association of components, systems, and equipment development under one roof has produced a particular and mutually sympathetic sensitivity to the perennially acute problems of size and weight of the Army's growing new electronics, whether borne on the back of the foot soldier, crammed into the fuselages of surveillance drones, or packed into multi-ton trailers of the new communications centers.

An obvious brute force solution to these problems consists of making the parts smaller and increasing the density of parts in assembly. This approach was reflected in the production of the Army's Handy-Talkie (Fig. 1), an excellent example of practical and effective miniaturization in the pretransistor period. Space for finger manipulation during fabrication was minimal, and the parts were essentially layered one on top of the other; maintenance was difficult and further miniaturization was only marginal. Accordingly, it was necessity that spurred the Signal Corps' development of the dip-soldered printed wiring concept in 1949 which, coupled with the timely invention of the transistor by Bell Telephone Laboratories, made possible a new dimension in electronic construction (Fig. 1).

In the ten years since the fortuitous mating of these two concepts, industry has contributed a dozen process innovations and layout disciplines, and has even stimulated its own development of varieties of new parts which are mechanically and electrically compatible with this construction technique. But once again, the miniaturization capability has become asymptotic, approaching an over-all limit in practical parts density of about 50,000 parts per cubic foot. Even the most compact layouts (based on an industry standardized 25-mil grid system) still involved considerable intercomponent waste space; only about 30 per cent of the over-all chassis volume, at best, was occupied by the parts themselves. As for the future of this technique, only small incremental improvements in miniaturization could be foreseen if the individual conventional components, at the fringes of their own arts, continued their slow and sporadic shrinkage in size. For these reasons, interest has been directed to development of a more effective

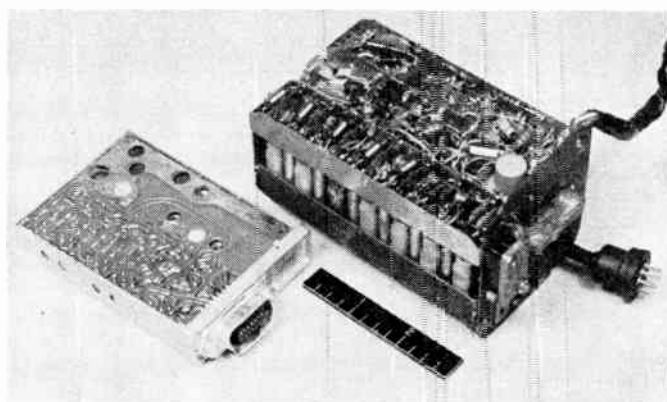


Fig. 1—Conventional construction techniques. (Left) TransistORIZED printed wiring receiver-transmitter with parts density of about 50,000/cubic foot. (Right) Hand-wired receiver-transmitter with parts density of about 8000/cubic foot.

system of practical miniaturization that could reduce the size of Army equipments at least ten times further without compromise in performance, and would be adaptable to machine production, and reducible to usage within 5 years.

BASIC TECHNIQUES AND CONCEPTS

A series of studies was undertaken in 1957 to provide perspective and guidance for the selection of such a practical, coherent system of microminiaturization applicable to all electronic circuitry operating at low and medium power levels. First, an estimate was made of the potential capabilities of the several techniques of depositing (integrating) circuits on inorganic substrates, and then of the intriguing possibilities of the new techniques of solid-state physics. This was complemented by an investigation of the history of the industry's experience with construction systems, the very high degrees of specialization inherent in quality parts production, and the growing interest of industry in the pre-packaged circuit or module as a vendor item. The results of this study (abstracted in Table I) provided a clear basis for decision: the system that could best satisfy the criteria established for size reduction, quality of performance, and machine producibility in the desired time frame was the “micro-module system” (Figs. 2 through 6).

Perhaps the most significant feature of the micro-module (MM) is the fact that the elemental part, as such, has lost its identity to the ultimate user; the smallest aggregate that the consumer will see and be able to service will be a standard shaped monolithic nugget having a specific electronic circuit function, an IF stage, a computer gate, discriminator, audio section, etc.

Prior to discussion of the technical aspects of this MM system, it may be helpful for fuller understanding to cite some of the practical considerations that guided its selection as a model system for the next generation of Army electronics.

TABLE 1
LOW AND MEDIUM POWER LEVEL MODULE CONSTRUCTION SYSTEMS

System	Approximate Practical Parts Densities Attainable (parts/cubic foot)	Relative Sizes of Modules	Availability of Parts for System	Remarks
1) Integration of conventional parts by printed wiring	50,000	200	Almost all conventional quality parts adapted to system; advances in state of art readily accommodated.	Microminiaturization potential limited: modules are repairable. Problem: Machine assembly complicated by varied sizes of parts.
2) Deposited parts on inorganic substrates (integrated circuits)	10 ⁶	10	Essentially limited to <i>R</i> and moderate <i>C</i> 's; small inductances possible for VHF; active parts and others that cannot be deposited may be attached. Research needed for complete system of parts (vapor deposition, sputtering, sintering, chemical deposition, layering, etc. are applicable techniques) yielding two-dimensional discrete (or distributed) parts.	Microminiaturization potential good: good producibility with special manufacturing facilities; parts quality is fair to good; non-repairable; relatively large area substrates require extra thickness to permit handling in processing and in usage, cutting down volumetric efficiency. Elements fabricated and assembled by same manufacturer. Problem: Parts system not complete. The two-dimensional integrating means which would give a rugged assembly.
3) Wafer assembly (Navy "Tinker-Toy" system)	10,000	1000	Basically, passive and active parts are separately fabricated and integrated with "printed" conductors on standard ceramic substrates, $\frac{7}{8}$ inch square. Deposited <i>R</i> and <i>C</i> of 2) above could be used. Any quality part can be adapted to system by using wafer as carrier. Extensive engineer program required to complete family of parts for wafer compatibility. Deposited parts require research, as in 2) above.	Microminiaturization potential—limited: excellent wafer assembly features compatible with mechanization; quality of special parts for system are fair to good; special manufacturing facilities required; module is non-repairable. Elements fabricated and assembled by same manufacturer. Problem: Substrate and spacing between wafers consume considerable volume. Parts system not complete.
4) Micro-modules (Army composite system)	500,000 at present; can expand into $\gg 5 \times 10^6$ as techniques 2) and 6) provide new capabilities	20	Passive and active parts are separately fabricated as in 3) but in form of standard shape (0.3 inch square) elements. Low wafer volume and spacing permit high-volume efficiency. Most quality parts can be adapted to two-dimensional shape of standard wafer. Techniques of 2) applicable to fabrication of wafer. Extensive engineering program requires to complete family of parts in micro-element size.	Microminiaturization potential good; excellent wafer assembly feature with module integration possible with simpler machinery. Nonrepairable. Elements will be procured and terminations affixed by assembler prior to module assembly. Can absorb benefits of future research on techniques of 3)–5), providing a "system" utilization for these techniques. Problem: Vendor supply sources needed for completion.
5) Deposited parts on single crystal substrates	up to 5×10^6	2	Limited experimental models make use of deposited techniques for some passive parts, using doping, etching, alloying, etc., of the crystal to form active parts, <i>R</i> 's and <i>C</i> 's. In research stages as yet. Full range of passive parts cannot be realized on the substrate as in the deposited system, 2). Considerable research needed.	Microminiaturization potential is excellent; adaptable to MM 4); nonrepairable modules. This is a logical bridge between deposited circuitry 2) and the ideal solid-state circuit 6). Problem: A system for integration of the modules is needed; parts system is not complete.
6) Solid-state circuitry	$\gg 10 \times 10^6$	1	Limited experimental models of complex digital circuitry have been made. Comments on substrate same as for 5). Full range of parts functions cannot be duplicated as yet. Vapor deposition of single crystals will provide powerful stimulus for this approach. Extensive research needed to develop full capabilities.	Microminiaturization potential is excellent; adaptable to MM 4); nonrepairable modules. This is the ultimate technique for microminiaturization in the long run. Problem: The solid-state circuit techniques are in the early stages; usage will be evolutionary with substantial capabilities developing 10–15 years from today.

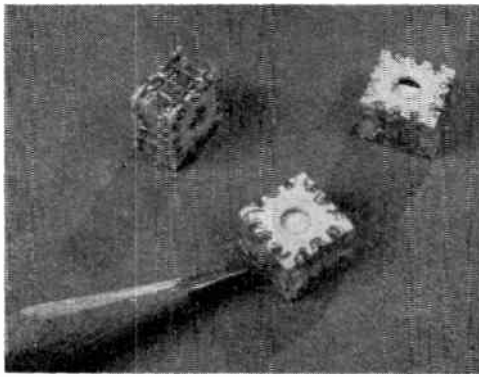


Fig. 2—Micro-module (MM) circuit functions. Typical encapsulated receiver sections in the standard 0.3-inch square cross section. Length of the MM is variable and dependent on the number of parts in the circuit. Parts density for receiver circuitry averages 500,000 to 600,000/cubic foot.

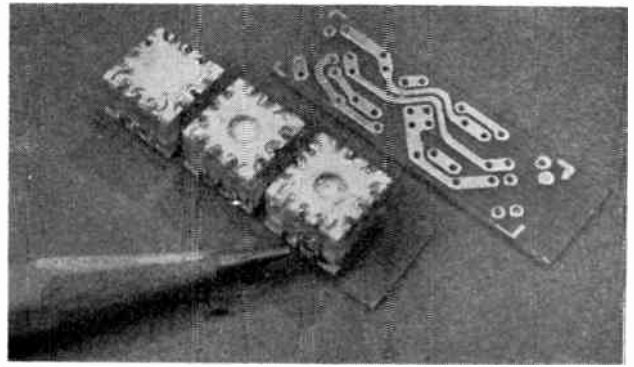
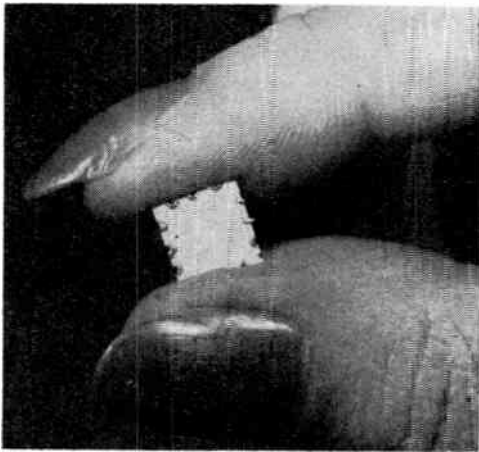
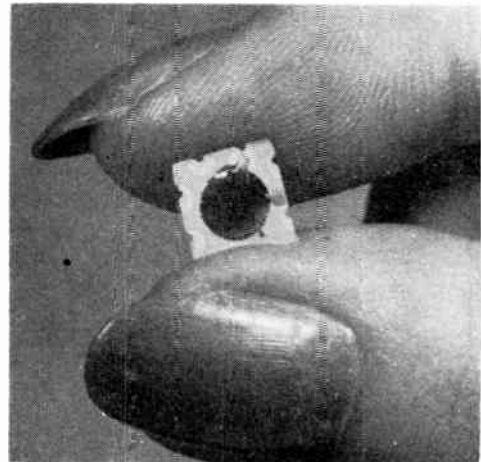


Fig. 3—Micro-module integration. Modules can be axially combined or efficiently integrated with printed wiring boards as shown. Uniform shape of MM means simplified machine handling in mechanized systems.



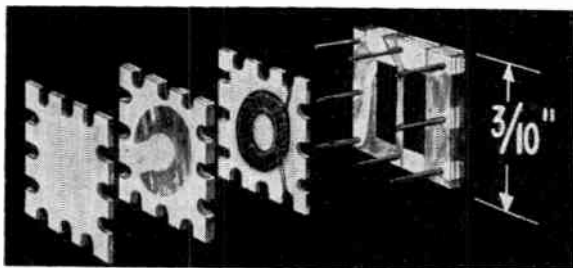
(a)

(a) This basic wafer is the foundation of all micro-elements. Presently made 0.3 inch square and 0.01 inch thick, it will lend itself to automatic manufacture. Two types of micro-wafers are being considered: the notched design for insertion of connecting wires, and the notchless design for connection with flat ribbon conductors. Twelve-notch design.



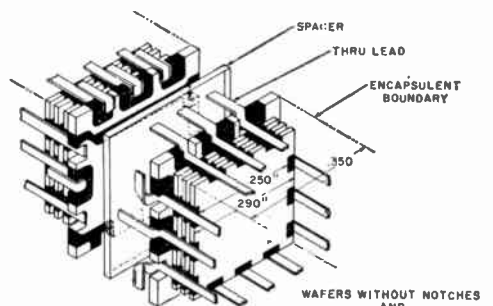
(b)

(b) The micro-element is processed using the best applicable process and materials, controlled to give the desired component characteristics. The micro-element shown here is a resistor with organic protection. Single- or double-sided elements can be made; these are now vendor-procedure. The eight-notch wafer is an early design.



(c)

(c) Micro-module, exploded view, before stacking and interconnecting. As many micro-elements as desired may be assembled to form the module. The over-all length is the flexible dimension of the system. A 0.01-inch space is allowed between micro-elements to provide electrical decoupling and allowance for joints and tolerances. Experimental machine assembly facilities are in use.



(d)

(d) Micro-module assembly, in this case using notchless micro-wafers and flat ribbon conductors. After assembly, the module is sealed by molding or encapsulation to form a solid body.

Fig. 4—General features of micro-module assembly.

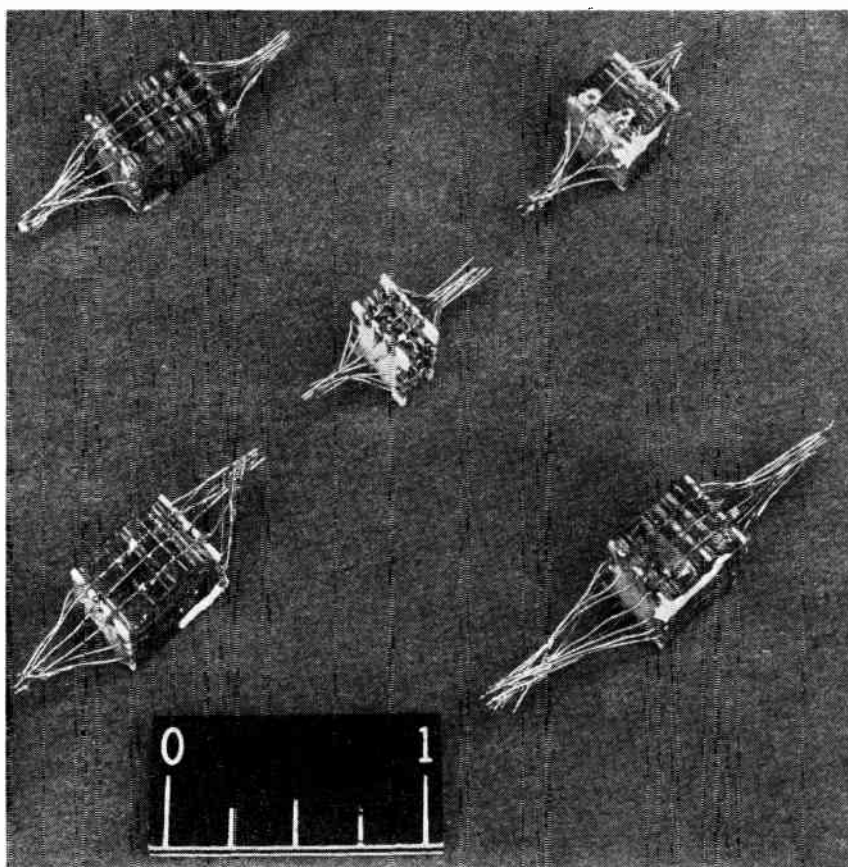


Fig. 5—Micro-module functions in experimental use. (Top left) Pulse generator. (Top right) 192-kc timer. (Center) Unencapsulated gate. (Lower left) 4.3-mc discriminator. (Lower right) Pulse shaper.

MM: THE RATIONAL SYSTEM

The micro-module concept is predicated on the continued use of the know-how, skills, and facilities (possibly modified) of the individual component and "printed circuit" manufacturers. The talents and production techniques extant in the industrial body today is the cumulative product of diverse development extending as far back as 30 or more years, and is perhaps the best index of our electronic progress. Not to use this background in any new proposed system of construction would be akin to stepping back to the infant days of radio and repeating the cycle of slow, progressive development of a new parts technology. It would seem more logical to use this existing talent to the maximum degree possible as a steppingstone to the new micro-miniaturization capabilities while still retaining the flexibility for absorption of any future advances in the state of the art. When future research provides new capabilities in the vapor deposition of functioning parts (or even complete circuits) these will be used to replace the micro-element or the micro-module, as the case may be. The same can be said of future accomplishment in solid-state circuits. There is only one basic requirement within the MM concept in this connection—the final shape of the new micro-element, or the new complete function,

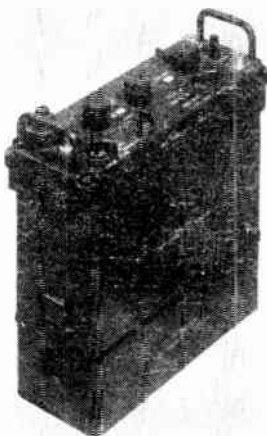
must conform to the geometric disciplines of the MM system.

The MM system can be recognized as a composite of the best features of the several systems within the experimental and practical production experiences of our electronic industry. As a logical next step, it permits use of any of the existing and effective micro-miniaturization techniques of ceramic printed circuits wherever the product of this construction meets the physical and electrical requirements of the MM system. It uses the very flexible assembly technique suggested in the Navy Tinkertoy program, and possesses the disciplined geometry so essential for machine assembly and testing. It offers applicability to all electronic circuitry. It satisfies the military micro-miniaturization need by a very substantial and adequate 10 times size reduction with no risky dependence on anticipated "breakthroughs" for success; when a breakthrough does occur—for example, if an amplifier can be made by solid-state circuit techniques—such an amplifier, in the requisite geometry, can be readily integrated into the system as another micro-module.

The MM miniaturization objective of ten times reduction of *over-all equipment sizes* is sufficient to satisfy the Army's presently projected long-term needs. Micro-

1960 ARMY DEVELOPMENT

AN/PRC-25



Weight 15 lbs.
Size 0.27 cu. ft.

AN/PRC-25

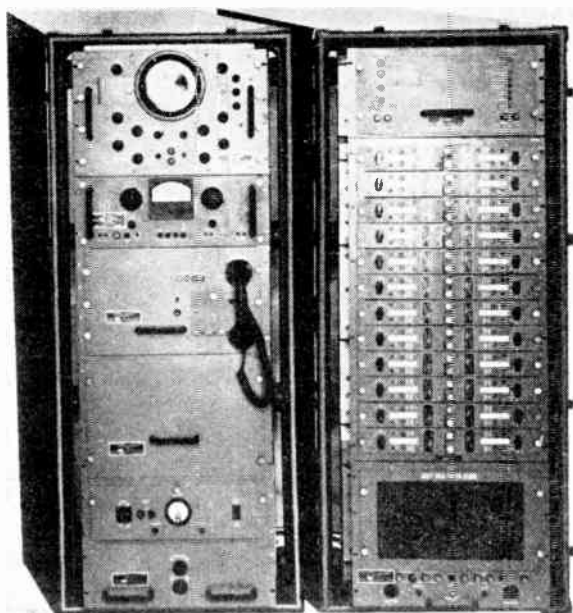
MICRO-MODULARIZED



Weight 5 lbs.
Size 0.014 cu. ft.

(a) Printed wiring-transistorized combat pack set now in advanced development is contrasted with the planned micro-module equivalent.

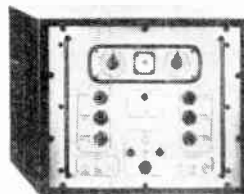
AN/TCC-13



PRESENT TUBE TYPE NEW

Weight 1200 lbs.
Size 40 cu. ft.

AN/TCC-26



NEW
TRANSISTORIZED TYPE

Weight 80 lbs.
Size 2-1/2 cu. ft.

AN/TCC-26



PROJECTED
MICRO-MODULARIZED
TYPE

Weight 3 lbs.
Size 0.07 cu. ft.

(b) Printed-wiring-transistorized telephone multiplexer equipment in advanced development is contrasted with a prior handwired tube version and the planned micro-module equivalent. The microminiature multiplexer is programmed for field tests by 1962.

Fig. 6—Planned micro-modularized equipments.

miniaturization, as an objective by itself, can easily become an obsession, compounding its problems into abstractions if taken far enough. Quality of performance, manufacturability, life, environmental resistance, reproducibility from different sources of supply, and economy are military considerations that *cannot* be dissociated from a realistic microminiaturization goal. This is not to deny research the license to unfettered pursuit of any other approach that could perhaps lead to entirely new concepts, but it is important to categorize such a pursuit as *research* and recognize the timetable that will be involved in the subsequent development and production engineering. Such research, as a matter of fact, is essential today in the broad areas of integrated circuitry and solid-state circuitry for the further growth and maturity of the MM construction system.

GEOMETRY AND COMPONENTS RANGES

The selection of the 0.3 by 0.3-inch dimensions for the micro-elements was an engineering decision based on a review of parts techniques, manufacturing capabilities, and the voltage and power level maxima established for micro-modules. A 1- to 2-watt power dissipation per micro-module, working frequencies up to 100 mc, 75-volt maximum levels, and applicability to general circuitry (IF, RF audio, filter, oscillator and computer logic circuits) were influencing parameters in the selection of this geometry. The 0.3-inch square was the smallest size in which the desired ranges of many components could be accommodated, including the electrolytics, high-K capacitors, quartz crystals of 7 mc and higher, coils up to 10 millihenries, transistors, diodes, ceramic resonators, glass capacitors, metal film resistors, and even certain electromechanical parts such as potentiometers and trimmer capacitors. The metal film resistors, in values up to about one megohm (adequate for almost all uses in the transistor circuitry), could be easily accommodated in this space. Actually, as many as four resistors could be placed on one wafer substrate; this component presented no serious problems.

The selected dimensions also permitted the use of any ceramic substrate (alumina, ferrite, glass, steatite) of only 10 mils thickness, as contrasted with the 20 mils or more that would have had to be used with a $\frac{1}{2}$ -inch square substrate; these requirements of thickness stem from the very critical demands for flatness and adequate ruggedness for machine handling. Since most of the micro-elements will be made by two-dimensional techniques (film resistors, deposited capacitors, deposited conductors), and other parts will have one dimension de-emphasized (crystals, ceramic filters, solid tantalum capacitors), it was essential that the substrate not consume any substantial percentage of the volume. The thinner 10-mil substrate also permits higher values of capacitance in the low-K dielectric system for precision and temperature compensating capacitors, and in the high-K system for the coupling and bypass capacitors.

Fig. 7 illustrates in part the adaptability of existing parts techniques to fabrication of micro-elements.

It is expected that up to 70 or 80 per cent of all electronic circuitry can be ultimately accommodated directly in the micro-module form. No compromise of effectiveness of the system is foreseen in the use of special oversize components, if needed, if the dimensions of such special parts are multiples of the basic 0.3-inch square geometry. For example, a battery constructed with a 0.9 by 0.6-inch cross section could be readily nested with the standard 0.3-inch square micro-modules to make a compact electronic package. The system basically strives for such geometric compatibility of all its elements.

Fig. 8 shows an experimental micromodule receiver operating at 49 mc.

DIGITAL CIRCUITS

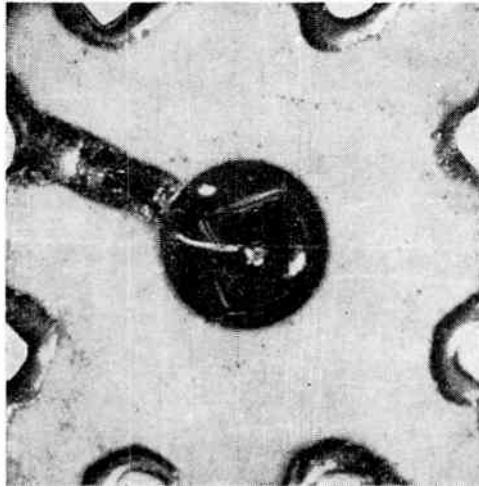
In speaking objectively of microminiaturization systems and capabilities, it is essential that a distinction be made between "general circuitry" (in receivers, amplifiers, filters, computers, etc.) and "digital circuitry," which is most frequently associated with the computer arts. Most of these digital functions can be realized through use of just resistors, capacitors, diodes, and transistors, parts which are adaptable to two-dimensional construction (Fig. 9). Precise frequency determining elements such as capacitor-inductor combinations, and quartz crystals in general are not representative of the digital circuit parts populations. Accordingly, computer circuitry will be the first beneficiary of microminiaturization. In such applications, module parts densities of above 10^5 parts per cubic foot are presently attainable.

Solid-state circuits for digital circuit usage have been made experimentally (Figs. 10-12) with parts densities of the equivalent circuits as high as 30×10^6 parts per cubic foot. The circuits are physically adaptable to the micro-module wafer as a substrate, with essentially the same parts density.

These parts densities attainable for digital circuitry are contrasted with the 500,000 parts per cubic foot average achievable for *general* circuitry using the MM (or any other) construction system. Further research and development in parts, materials, and related techniques are needed to advance capabilities beyond this point.

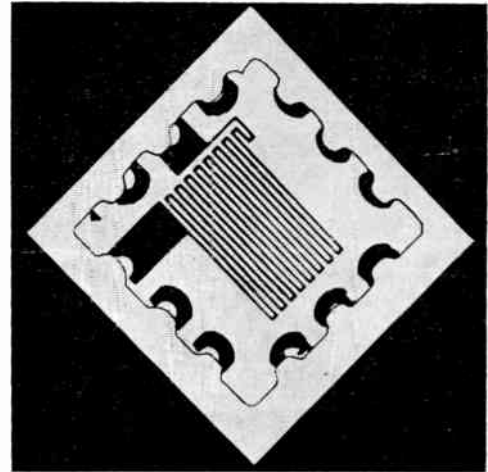
CONCLUSION

The first production uses of printed wiring barely eight years ago were attended by rigorous efforts of industry to apply some measures of uniformity in practices of this new art, to establish standards for good layout systems such as hole sizes to characterize the base materials—all this and more to provide common denominators among manufacturers wishing to exploit this new construction tool. This same type of discipline,



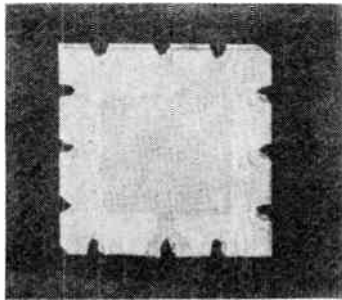
(a)

(a) Transistors and diodes. Semiconductor devices are the active devices for all micro-modules, and must be repackaged as micro-elements. In the illustration, the basic semiconductor device is contained between two recessed ceramic cases of the micro-element. The semiconductor elements are connected to the terminal notches by printed conductors on the cover surfaces. (Early 8-notch design.)



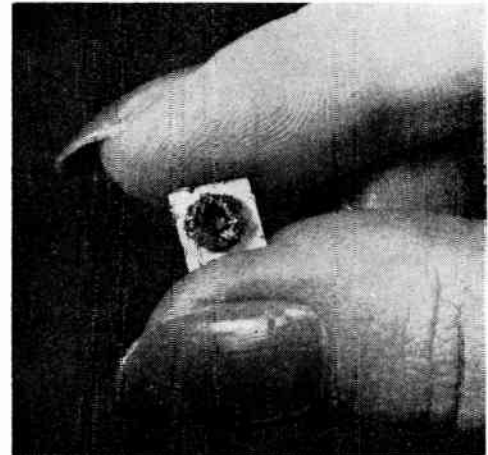
(b)

(b) Resistors. Micro-element resistors are now metal alloy or oxide precision types. Values up to 1 megohm can be accommodated. Unit shown here is a metal film deposit on a glass substrate. Up to 4 resistors can be placed on one wafer.



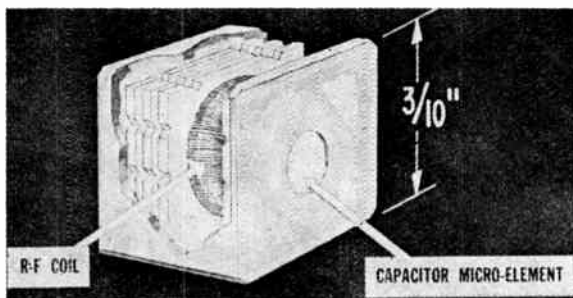
(c)

(c) Capacitors. General purpose capacitors with values up to 0.01 mfd and higher can be formed using conventional high-K dielectric ceramics. Multilayer structures extend this range to 0.3 mfd. larger capacitor; values from 0.3 mfd to about 10 mfd are obtained using solid electrolyte structures. Precision capacitors up to 100 mmf are made by using conventional precision, temperature-compensating types of ceramic bodies.



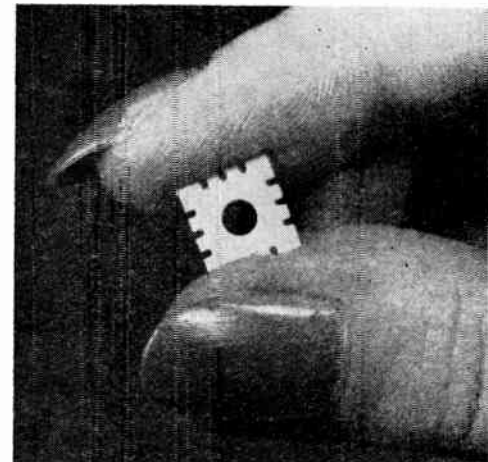
(d)

(d) Interim inductors in the form of toroidal windings on ferromagnetic cores are accommodated in values ranging up to 10 mh, adequate for RF coil coverage from 100 kc to about 100 mc. Shaped ferrite wafers with controllable air gaps are being considered.



(e)

(e) Tuned circuits and IF transformers. These units use a toroidal coil at present with an associated fixed capacitor so placed within the MM that the low potential electrode surface of the capacitor is exposed through a terminal wafer, even after encapsulation (see Figs. 2 and 3). Adjustment is made by removal of a portion of this electrode area. Thus, tuned circuits are pre-adjusted to value after assembly and can be considered fixed tuned. The voltage sensitive reverse biased diodes show excellent possibilities for a micro-module tuning system. Micro-element shaped cases can also accommodate ceramic filters and quartz crystals.



(f)

(f) Terminal wafer with aperture for circuit adjustment.

Fig. 7—Some micro-element characteristics.

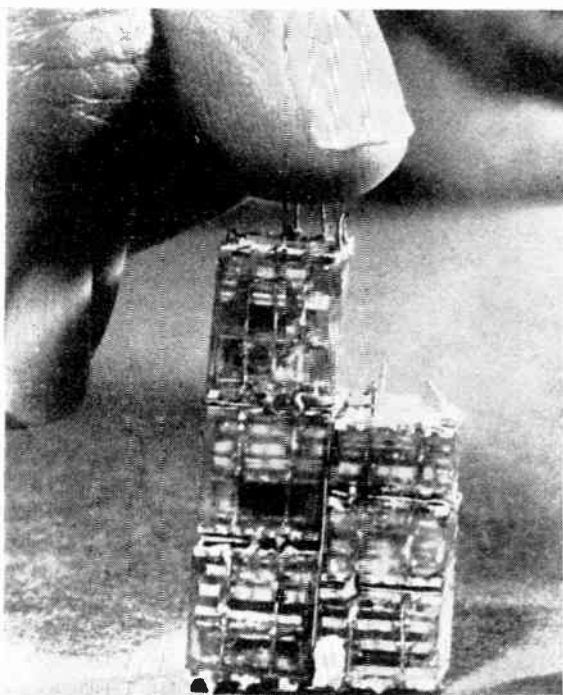


Fig. 8—The total electronics of an experimental micro-modularized FM receiver, which is fixed tuned to operate at 49 mc. The receiver uses a 4.3 mc. IF and includes squelch circuitry.

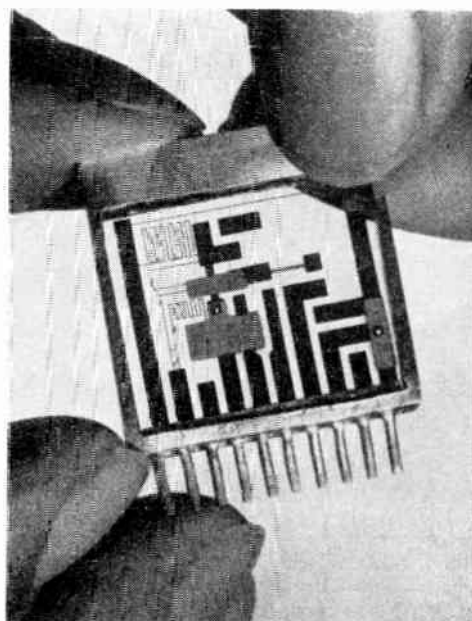
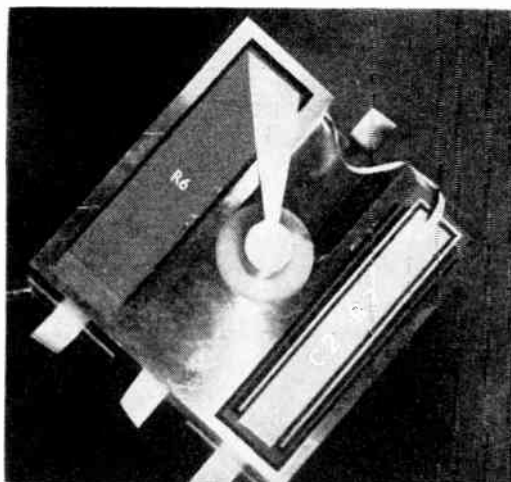


Fig. 9—An advanced type of "integrated circuit function" for digital circuit applications. (Photograph courtesy Hughes Aircraft Company.) Individual circuit elements are deposited on an inorganic substrate to yield a parts density of over 10⁸/cubic foot. This unit contains 12 resistors, 4 capacitors, 4 diodes, and uses 4 transistors which plug into the end of the circuit module.



British Crown Copyright

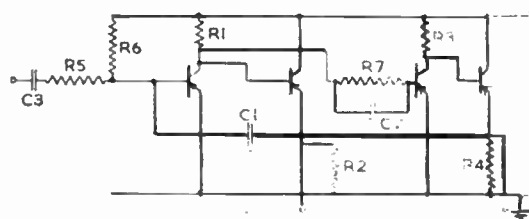


Fig. 10—A combination of the integrated circuit and solid-state circuit techniques is illustrated in this multivibrator. (Developed by Plessey, Ltd. for the Royal Radar Establishment, Malvern, Eng.) A silicon crystal substrate is doped and shaped to form 4 normal transistors; 4 of the resistors are formed by silicon bridges in the substrate. Other resistors and capacitors are deposited in film form directly on the silicon block with intervening insulating films. Unit is about 0.24 inch square and 0.08 inch thick. Equivalent parts density of this circuit function is about 5×10^6 /cubic foot.

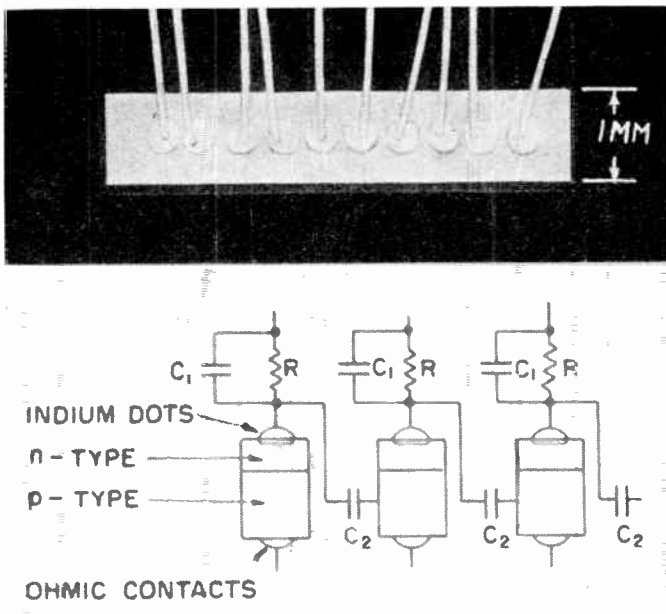


Fig. 11—Solid circuit (shift register, developed by RCA). A single germanium crystal is treated to form the active and passive elements. Equivalent parts density for an encapsulated unit will be of the order of 10×10^6 /cubic foot.

perhaps to a different degree, is needed in any venture in which the talents and resources of many people are required.

Certainly some such discipline, at least in shapes and sizes, is desirable for the increasing number of diverse microminiature circuits being fabricated today. The micro-module system is being forwarded as just *one* such discipline based on a logical and feasible extension of existing techniques and capabilities, an excellent automatic manufacturing capability, and a flexibility which will accommodate any product of future parts research and development.

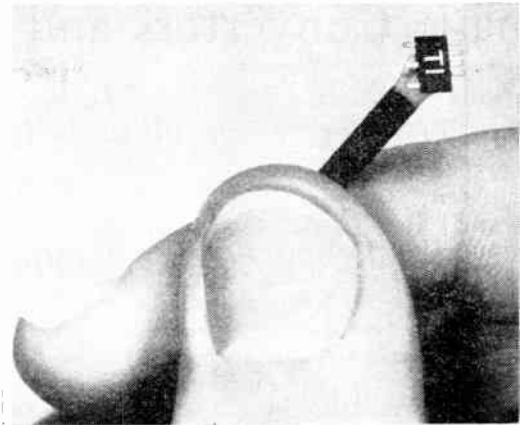


Fig. 12—Solid circuit. (Multivibrator developed by Texas Instruments, Inc., Dallas, Tex.) Controlled diffusion, photographic masking and etching are performed on a single crystal wafer to form the active and passive elements properly connected to form the desired circuit. All components are within the single crystal wafer and no additional substrate is required. The passive and active elements formed are resistors, capacitors, diodes, and transistors. The unit is adaptable to MM system use.

A hermetically sealed multivibrator unit is approximately $\frac{1}{4}$ by $\frac{1}{8}$ by $\frac{1}{32}$ inch. The equivalent parts density is about 30×10^6 parts per cubic foot.

BIBLIOGRAPHY

- [1] G. W. A. Dummer, "Problems affecting the design of service components and a survey of present and future UK developments," *Proc. Internatl. Symp. on Electronic Components*, Malvern, Eng.; September 24-26, 1957.
- [2] "Micro-Module Production Program," Radio Corp. of America, Camden, N.J. First Quart. Rep., MM-206, Contract DA36-039sc-75968; April 1-June 30, 1958.
- [3] V. J. Kublin and S. F. Danko, "Micro-modules: components and materials requirements," *Proc. Second Natl. Conf. on Production Techniques*, New York, N. Y.; June 4-6, 1958.
- [4] "The Micro-Module Program" and "The Micro-Module Concept," Radio Corp. of America, Purchasing, Bldg. 8-9, Camden, N. J.
- [5] "Micro-Module Production Program," Radio Corp. of America, Camden, N.J., Second Quart. Rep., Contract DA 36-039sc-75968; July 1-October 1, 1958.
- [6] "Micro-Module Production Program," Radio Corp. of America, Camden, N.J., Third Quart. Rep. Contract DA 36-039sc-75968; October 1, 1958-January 1, 1959.

Image Converters and Image Intensifiers for Military and Scientific Use*

MYRON W. KLEIN†

Summary—Since World War II, the infrared image converter tube used by the Armed Forces for night observation has been greatly improved. Recent developments in electronics have led to a number of improved image converter tubes sensitive to infrared or ultraviolet radiation, as well as image intensifier devices which will permit observation under starlight illumination with no projected radiation whatsoever. The devices described are of interest to astronomers, radiologists, and nuclear physicists, as well as to the military. Some of the devices have already been constructed while others face serious technological problems. The Corps of Engineers is supporting a program of research in this field to provide image intensifier systems for the Armed Forces.

INTRODUCTION

Background

THE process of photoemission, whereby electrons are emitted from certain materials under the influence of light, is still not thoroughly understood. It is, however, the basis of many very interesting, valuable, and highly developed industrial, military, and scientific devices ranging from radioactive measuring instruments to television camera tubes.

Of interest to the military is the image converter or infrared image tube, which is shown schematically in Fig. 1. It consists essentially of an evacuated glass and

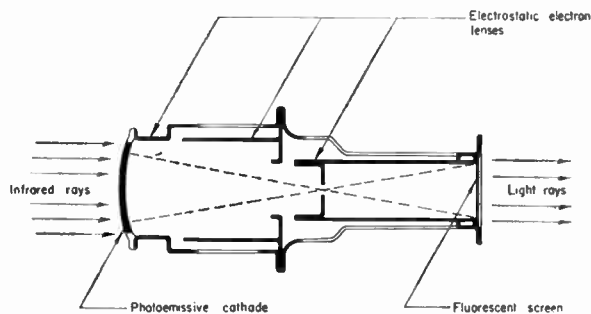


Fig. 1—Image tube schematic.

metal tube, one end of which bears an infrared sensitive photoemissive coating on its inside surface while the other end is coated with a fluorescent powder which emits visible light when bombarded with high velocity electrons. Within the tube are one or more coaxial metal rings to which high voltage may be applied. If an invisible infrared image is focussed on the photoemissive coating, electrons will be emitted within the image tube. The electrons emitted from this photocathode and having a distribution corresponding to the original in-

frared image are accelerated by the electrostatic fields surrounding the metal rings and focussed on the fluorescent screen at the other end of the tube. The resulting fluorescence produces a visible reproduction of the original invisible infrared image, which was focussed on the photoemissive cathode.

Infrared Image Converter Tube

Since the invention of the photoemissive infrared image converter tube in 1934^{1,2} scientists have attempted to utilize it both as a component of military-weapons³ and as a research tool.⁴ While the armed forces have used the tube primarily in infrared viewers for night battlefield surveillance and weaponsights, there have also been important scientific applications for the tube in the fields of vision, ophthalmology, dermatology, entomology, infrared microscopy, and photomicrography. Unfortunately, attempts to use the early tubes for such scientific research were greatly limited because of the military restrictions on the use of the tube which were necessarily imposed in the national emergency. A limited quantity of the World War II image tube, Type 1P25 (see Fig. 2), was released to the commercial market, and a supply of a British image tube found its way into this country. These two types were the only tubes available to the public up to 1953 when the Government removed the restricted classification from several new designs. The one shown in Fig. 3 can now be purchased in either infrared or visible sensitive models from commercial sources. These new infrared tubes have already been used for inspection of photographic film, for observation of animal behavior, and, in one case, to view through a microscope the dissected retina of an eye, using infrared radiation instead of visible light, which would destroy the specimen. Recently, an ultraviolet microscope incorporating an ultraviolet sensitive image converter tube was placed on the market and should be of great use in scientific research. Also available are special image converter tubes^{5,6} incorporat-

¹ G. Holst, J. H. Deboer, M. C. Teves, and C. F. Veenemans, "Apparatus for transformation of light of long wavelength into light of short wavelength," *Physica*, vol. 1, p. 297; 1934.

² V. K. Zworykin and G. A. Morton, "Applied electron optics," *J. Opt. Soc. Amer.*, vol. 26, p. 181; 1936.

³ G. A. Morton and L. E. Flory, "Infrared image tube and its military applications," *RCA Rev.*, vol. 7, p. 385; 1946.

⁴ A. B. Dember, "Image tubes," *Med. Phys.*, vol. 2, pp. 414-424; 1950.

⁵ R. G. Stoudenheimer and J. C. Moor, "An image converter tube for high speed photographic shutter service," *RCA Rev.*, vol. 18, pp. 322-331; September, 1957.

⁶ B. R. Linden and P. A. Snell, "Shutter image converter tubes," *Proc. IRE*, vol. 45, pp. 513-523; April, 1957.

* Original manuscript received by the IRE, November 26, 1958; revised manuscript received, February 16, 1959.

† U. S. Army, Fort Belvoir, Va.

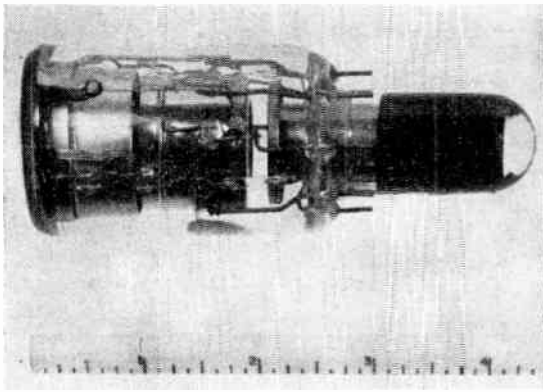


Fig. 2—1P25 image tube.

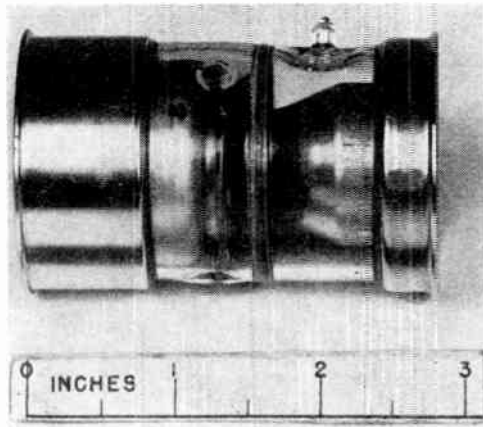


Fig. 3—IC-16-2 image tube.

ing a fine mesh grid which permits gating the tube on and off for 10^{-8} seconds. These tubes find application as shutters for high speed cameras and have been used successfully in explosion propagation studies.⁷ In France, at the Paris Observatory, a demountable image tube⁸ sensitive to visible light for astronomical photography has also been constructed. The film is placed inside the evacuated image-tube housing and is exposed directly by electrons providing an appreciable gain over conventional photographic methods. As a result of this work, a group of American astronomers have undertaken the development of an image-converter tube^{9,10} which permits the electron image to emerge through a thin membrane or foil to expose an external photographic plate. The tremendous improvement both in sensitivity and quality which has been achieved in these newer tubes makes them much more useful to science than has been heretofore believed.

⁷ J. S. Courtney-Pratt, "A new method for photographic study of fast transient phenomena," *Research*, vol. 2, p. 287; June, 1949.

⁸ A. Lallemand, M. Duchesne, and G. Wlerick, "Electronic photography," *Proc. Image Intensifier Symp.*, Ft. Belvoir, Va.; October 6-7, 1958. (To be published.)

⁹ R. H. Miller, W. A. Hiltner, and J. Burns III, "Image converters with thin protecting foils," *Astrophys. J.*, vol. 123, pp. 368-370; March, 1956.

¹⁰ W. A. Baum, W. K. Ford, and J. S. Hall, "Recent astronomical tests of thin film image converters," *Astrophys. J.*, vol. 63, p. 47; 1958.

Recent Developments

A number of new developments over the past five or six years have pointed strongly toward the possibility of a starlight viewing device which would be capable of forming bright visible images under extremely low levels of natural night illumination. Such a device would completely eliminate the need for any projected artificial radiation. Many of these developments were spurred by the astronomers who are seeking a method of extending the range of their large telescopes. The radiologists, who would like methods of reducing X-ray dosage,^{11,12} also entered the field, along with the nuclear scientists,¹³⁻¹⁶ who wish to track high energy particles by their faint scintillations in crystal chambers. In addition to these, the Army, Navy, and Air Force saw a need for improved methods of night observation, and sponsored a great deal of work in this field. Much research had already been performed on photoemissive surfaces, phosphors, photo-etched glass screens and other special target materials related to intensifier tubes and closed circuit television systems for low light-level operation.^{17,18} Fortunately, the goals of all the interested parties were quite similar, although each had its specific requirements and limitations imposed by its special applications, which dictated the approaches to be followed. Many of the nations' scientists began to combine these developments into image intensification processes, and numerous schemes for image intensifier tubes and systems were suggested.

Multi-Alkali Photosurface

One of the important steps forward was the development of a more sensitive photoemissive surface, the multi-alkali surface, whose spectral range (see Fig. 4)¹⁹ extended from 300 to 800 millimicrons in wavelength. This surface is formed as an evaporated semitransparent coating of cesium, potassium, sodium, and antimony, and it is three to four times as sensitive to visible light with only one tenth the dark emission of the conventional cesium-antimony surfaces. The cathode sensitivity of the multi-alkali surface has been measured at over 200 μa per lumen, while that of the cesium-antimony surface is seldom higher than 50 in the semi-transparent form.

¹¹ J. W. Coltman, "The scintillation limit in fluoroscopy," *Radiology*, vol. 63, pp. 867-875; 1954.

¹² R. H. Morgan, "Screen intensification," *Amer. J. Roentgenology*, vol. 75, pp. 69-76; 1956.

¹³ L. J. Mandel, *Rev. Sci. Instr.*, vol. 32, p. 405; 1955.

¹⁴ E. K. Zavoisky, M. M. Butslav, A. G. Plakhov, and G. E. Smolkin, *J. Nuclear Energy*, vol. 4, p. 340; 1957.

¹⁵ G. T. Reynolds and P. E. Condon, *Rev. Sci. Instr.*, vol. 28, p. 1098; 1957.

¹⁶ R. Kalibjian, University of Calif., Berkeley, Rad. Lab. Rep. No. 4732; 1956.

¹⁷ G. A. Morton and J. E. Reudy, "The low light level performance of the intensifier orthicon," *Symp. on Photoelectronic Image Devices*, Academic Press, London, Eng.; 1959.

¹⁸ R. K. H. Gebel, "Daytime detection of celestial bodies using the intensifier image orthicon," *Proc. Image Intensifier Symp.*, Ft. Belvoir, Va.; October 6-7, 1958. (To be published.)

¹⁹ A. H. Sommer, "New photoemissive cathodes of high sensitivity," *Rev. Sci. Instr.*, vol. 26; July, 1955.

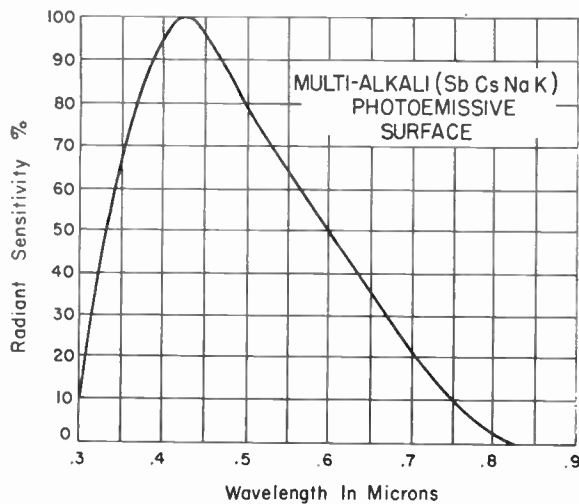


Fig. 4—Multi-alkali photosurface.

PHOTOELECTRIC IMAGE CONVERTER INTENSIFIERS

Low Magnification Image Intensifier Tube

The simplest form of image intensifier which has been built is the image converter tube wherein the image is demagnified electron-optically (see Fig. 5). Since the electron density in the electron image increases inversely as the square of the magnification, a reduction of the image diameter to one sixth gives an electron density and a resultant brightness gain of thirty-six times. Using the multi-alkali surface described above along with accelerating potentials as high as 25 kv, tubes of this type have been able to form images five hundred to a thousand times brighter than the original scene brightness. By introducing an X-ray fluorescent screen into the tube directly in front of the photosurface, this tube has been used as an X-ray image intensifier²⁰ for fluoroscopy with greatly reduced X-ray dosages while still providing increased image brightness.

Cascade Image Tube

A second type of image intensifier tube is the cascade image tube (see Fig. 6),²¹ in which two or more image converter tubes are constructed in tandem in one envelope such that the phosphor screen of the first is separated from the photocathode of the second only by a transparent glass membrane. In this way the first tube serves as a preamplifier for the second, and additional sensitivity may be obtained. In actual practice the phosphor and photocathode are coated on opposite sides of a very thin glass membrane which separates the tube sections. These tubes are difficult to construct, but are capable of brightness gains of several thousand times. Single image converter tubes may also be cascaded using fast lenses between phosphor screen and photocathode, but the brightness gain obtained is greatly reduced.

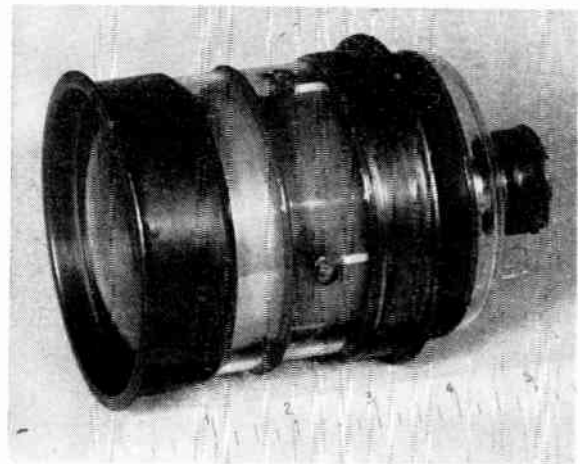


Fig. 5—Low magnification image intensifier tube.

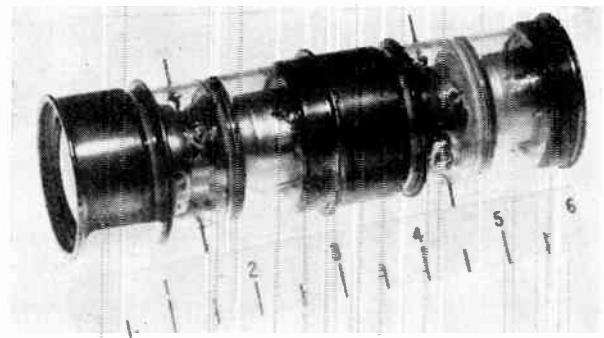


Fig. 6—Cascade image tube.

Secondary Emission Mesh Intensifier Tube

Another type of image intensifier tube which has been proposed is the secondary emission mesh intensifier shown in Fig. 7. This tube is constructed much like a conventional image converter tube but contains a series of very fine mesh screens coated with a secondary emitting material such as magnesium oxide. When a faint visible image is focussed on its photoemissive cathode, the emitted electrons are imaged on the first mesh by the magnetic focussing coils surrounding the tube. Secondary electrons are emitted in the ratio of three or four to one and are reimaged on successive mesh where the secondary emission amplification process is repeated. Upon emerging from the final mesh, the amplified electron image is accelerated by an electrostatic field of the order of 16 kv and focused upon the fluorescent screen on the rear faceplate of the tube. Experimental tubes of this type have been attempted, but much work will be needed before this tube will be usable in practical image intensifier equipment.

Transmission Secondary Emission Intensifier Tube

A similar tube, the transmission secondary emission intensifier (see Fig. 8)^{22,23} has been devised using thin

²² E. J. Sternglass, *Rev. Sci. Instr.*, vol. 26, p. 1202; 1955.

²³ M. M. Wachtel, D. D. Doughty, and A. E. Anderson, "The transmission secondary emission image intensifier," *Proc. Image Intensifier Symp.*, Ft. Belvoir, Va.; October 6-7, 1958. (To be published.)

²⁰ F. H. Marshall, "Fluoroscope image amplifying tube," *Electronics*, vol. 26, pp. 172-173; October, 1953.

²¹ R. G. Stoudenheimer and J. C. Moor, *Symp. on Photoelectronic Image Devices*, Academic Press, London, Eng.; 1959.

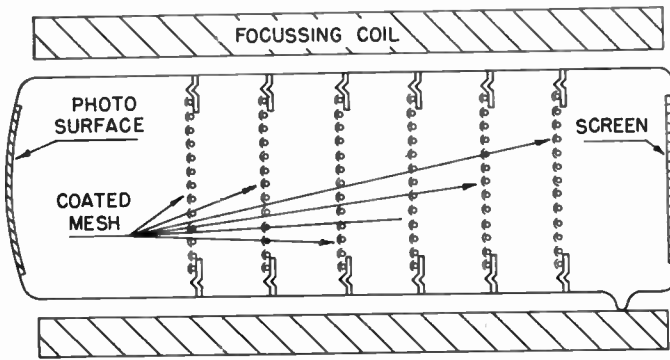


Fig. 7—Secondary emission mesh intensifier tube.

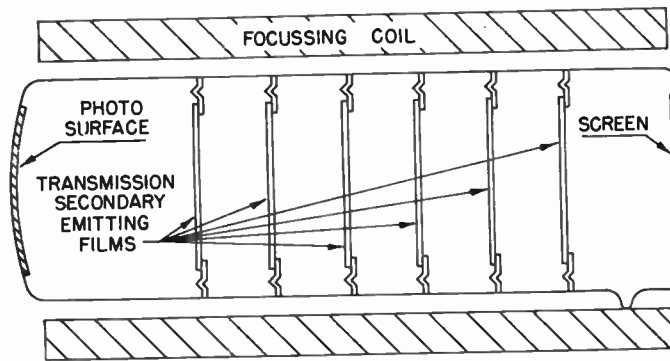


Fig. 8—Transmission secondary emission intensifier tube.

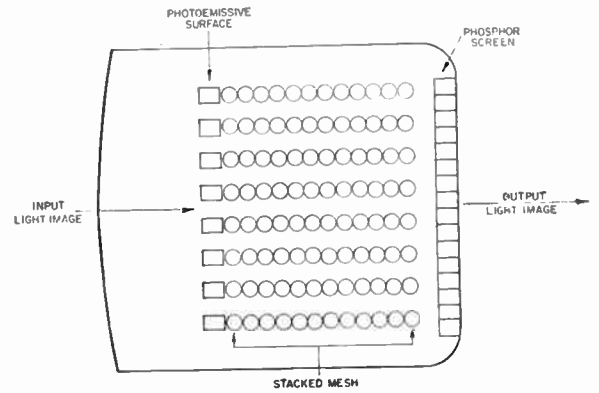


Fig. 9—Channeled secondary emission intensifier tube.

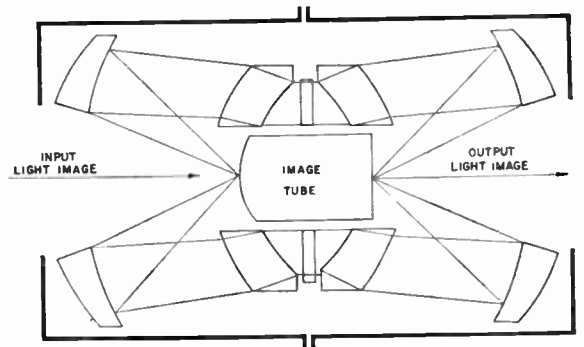


Fig. 10—Regenerative image intensifier system.

foils or membranes on which secondary emitting coatings are deposited. These membranes have been made from aluminum, aluminum oxide, or silicon monoxide coated with a secondary emitting layer of potassium chloride. In this case the secondary electrons are emitted from the rearside of the membranes and, since no mesh is used, higher resolving power is obtained. Electron amplification of 1500 has been achieved through six foils with resolutions of six or eight line pairs per millimeter. Serious problems which must be overcome in this device are instability of the secondary emitting coating and film penetration of primary electrons.

Channeled Secondary Emission Intensifier Tube

Still another image intensifying device is the channeled secondary emission intensifier (see Fig. 9).^{24,25} This technique makes use of a bundle of very fine metal tubes or perforated screens which may be stacked up in register to form long tubes. If the walls of the tubes thus formed are coated with a secondary emitting material, an electron image focussed on the surface of the bundle or stack will be channeled along the tubes and amplified by secondary emission. The emergent image is then accelerated to a phosphor screen where a greatly intensi-

fied image is formed. Some comparatively coarse forms of this device have been built using metal tubes as well as perforated glass and plastic screens. The resolution obtained has been much lower than with the previously mentioned systems.

Regenerative Image Intensifier System

A regenerative image intensifier system (see Fig. 10)²⁶ has also been built. This is an optical feedback technique whereby a portion of the output image from the screen of an image converter tube is fed back to the corresponding point in the image on the photocathode by means of lenses and mirrors. Although the resolution obtainable is obviously much too low for military use, it is still of interest to the nuclear physicist as a simple and inexpensive intensifier for scintillation photography.

IMAGE ORTHICON INTENSIFIER DEVICES

Image Orthicon Tube

The tube which has received the greatest amount of study,²⁷ both by commercial and military research groups, is the well known image orthicon television camera tube, shown schematically in Fig. 11. As in the other tubes described, the visible image is focused on

²⁴ A. Roberts, "Regeneration in conventional and channeled image intensifiers," *Symp. on Photoelectronic Image Devices*, Academic Press, London, Eng.; 1959.

²⁵ J. Burns III and M. J. Nueman, "The channelled, shaped-screen image intensifier," *Proc. Image Intensifier Symp.*, Ft. Belvoir, Va.; October 6-7, 1958. (To be published.)

²⁶ L. W. Jones and M. L. Perl, "Studies of regenerative image amplification and its application to particle track imaging," *Proc. Image Intensifier Symp.*, Ft. Belvoir, Va.; October 6-7, 1958. (To be published.)

²⁷ A. Rose, "Television pickup tubes and the problem of vision," *Adv. in Electronics*, vol. 1, pp. 131-166; 1948.

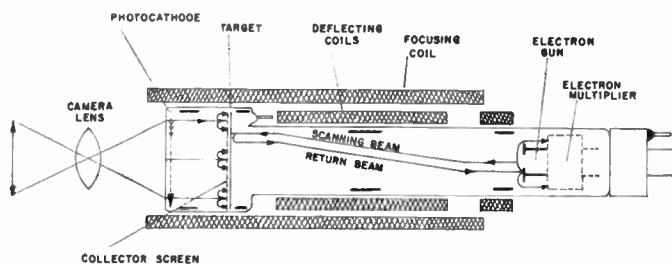


Fig. 11—Image orthicon tube schematic.

the photocathode from which electrons are emitted and focused on a thin glass target. As a result of secondary emission this target is left with a positive charge image which is scanned from the rear by an electron beam. Electrons in the electron scanning beam are decelerated when they near the target and discharge the positive portions of it. The electron beam thus modulated by the loss of electrons to the target returns to an electron multiplier section and forms the video signal, which is displayed on a conventional monitor in a closed TV system. This tube holds great potentialities as an image intensifier by virtue of the signal integration which may be accomplished on its target before introduction of the now-limiting noise in the electron scanning beam. This is considered especially important for operation at extremely low light levels, where the number of light photons received from the scene per second may be too low to instantaneously convey all the image information. It is possible to store the signal on the target for several seconds before reading it out and to present the picture on a storage display tube. A second reason for the importance of the image orthicon tube is found in the scanning technique itself which permits introduction of special electronic circuits for increasing contrast and beam modulation, special display methods and, of course, remote viewing. Unfortunately, in the image orthicon tube, the beam noise is still very large when the signal is small, so that there is an inherent limit to its detectability.

Image Intensifier Orthicon Tube

By incorporating the image converter tube as a pre-amplifier to the image orthicon tube, the gain of this system can be increased to a value where beam noise is no longer limiting, resulting in usable pictures under starlight conditions. This type of tube, the intensifier image orthicon, shown in Fig. 12, has been built^{17,28} with both one and two stages of image converter section and has been studied intensively by the Armed Services.

High Gain Image Orthicon

A new image orthicon tube developed by one U. S.

²⁸ A. Rotow and D. Marschka, "The image intensifier orthicon," also C. T. Shelton, B. F. Walker, and D. E. Townsend, "A night television system," *Proc. Image Intensifier Symp.*, Ft. Belvoir, Va.; October 6-7, 1958. (To be published.)

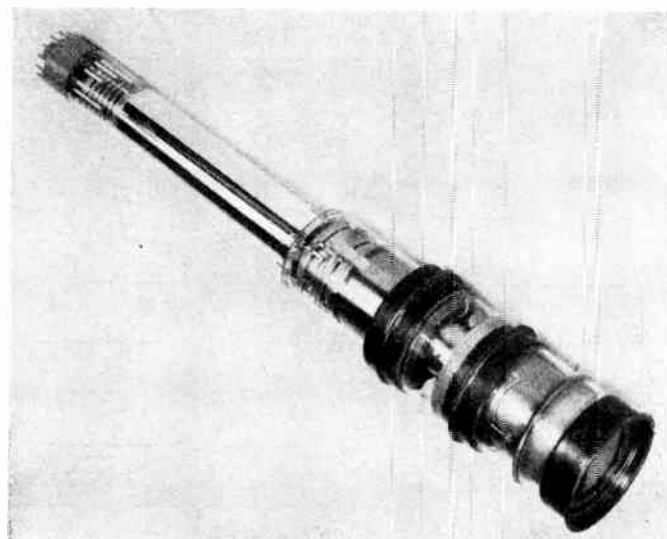


Fig. 12—Intensifier image orthicon tube.

firm utilizes a semitransparent target²⁹ enabling the tube to operate with less than one thirtieth of the illumination required by conventional image orthicon tubes.

Light-Scanned Image Orthicon Tube

It has also been proposed to build a tube in which a beam of light may be substituted for the scanning beam of electrons. A sketch of this light-scanned image orthicon tube^{30,31} is shown in Fig. 13. In this tube, the rear side of the target is coated with a photoemissive mosaic made of tiny independent islands of a photoemitter. The potential of the glass target in this case is adjusted so that the secondary emission ratio is less than one, *i.e.*, fewer electrons are emitted than land on the target, thereby forming a negative charge image on the target. This means that each element of the photoemissive mosaic is brought to a potential which corresponds to the brightness of its corresponding point in the original scene. The rear mesh which is provided in this tube is established at a potential which will just stop the electrons when emitted with the velocities normally encountered in photoemission. If, however, any element of the mosaic is brought to an increased potential, it will emit its electrons with an increased velocity, enabling them to pass through the rear mesh and arrive at the electron multiplier to become a part of the video signal. This tube, if perfected, might be capable of operating at illumination levels much lower than the conventional image orthicon; and, when provided with an image tube intensifier section, it could surpass the present image intensifier orthicon tubes.

²⁹ H. Hamm and P. Wargo, "An image orthicon with a new target," *Proc. Image Intensifier Symp.*, Ft. Belvoir, Va.; October 6-7, 1958. (To be published.)

³⁰ B. Meltzer and P. L. Holmes, "Beam temperature, discharge lag and target biasing in some television pickup tubes," *Brit. J. Appl. Phys.*, vol. 9, pp. 139-433; April, 1958.

³¹ S. A. Ward, "Light scan camera tube," *Proc. Image Intensifier Symp.*, Ft. Belvoir, Va.; October 6-7, 1958. (To be published.)

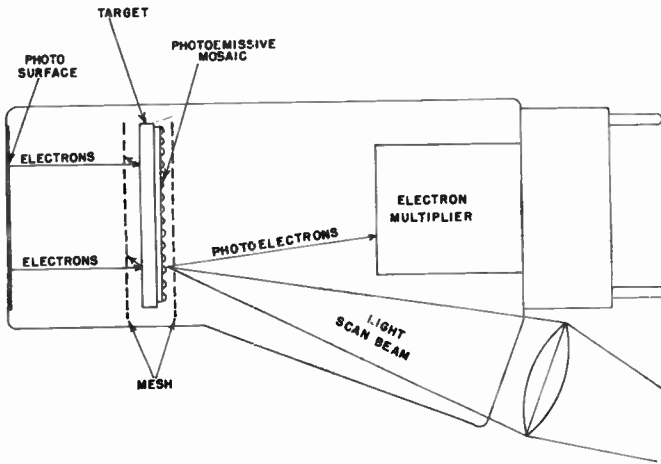


Fig. 13—Light scanned image orthicon tube.

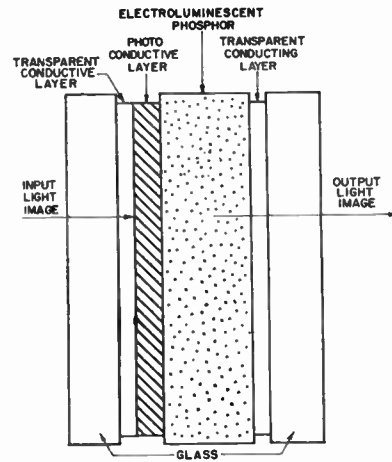


Fig. 14—Solid-state image intensifier.

SOLID-STATE INTENSIFIERS

Multi-Layer Type

An entirely different principle of image amplification is used in the solid-state image intensifier (see Fig. 14).³² In the simplest form of this device, a phosphor layer is placed between two transparent conducting plates to which an electric field of about 100 v is applied. If an ultraviolet image is focused on the screen thus formed, a marked increase in light emission is produced. In another form of the solid-state intensifier, an electroluminescent phosphor layer and a photoconductive coating are sandwiched between transparent conducting coatings on two plates of glass. The photoconducting coating is an insulator such as cadmium sulphide which, when illuminated with light, becomes conducting. In operation, a potential source is attached to the transparent conducting coatings across the photoconductive and phosphor layers. If light is focused on the photoconductive layer, it becomes conductive, permitting

the electric field to appear across the phosphor with a resulting emission of light. The conductivity developed under the influence of light corresponds to the brightness of the scene, and, although the phosphor reacts non-linearly to the electric field, halftone images are produced. Unfortunately, this type of image intensifier depends upon the characteristics of the photoconductor as well as the phosphor, and its brightness gain is presently limited by them to a range of illumination which is well above that of starlight or moonlight.

CONCLUSION

Many of the devices described have already been built and have demonstrated their capabilities for image intensification. Some are limited because of serious technological problems, others by the characteristics of some material involved. The research program of the Warfare Vision Branch of USAERDL is based on theoretical and practical consideration of all these factors and is aimed toward providing prototype components for study and evaluation on which to establish design of future military image intensifier systems.

³² D. A. Cusano and F. E. Williams, "Unique phosphors that amplify light," *G. E. Rev.*: September, 1956.

Naval Ordnance and Electronics Research*

W. W. SCANLON† AND G. LIEBERMAN†, ASSOCIATE MEMBER, IRE

Summary—Two programs of electronics research at the Naval Ordnance Laboratory are reviewed. These are photoconductor research and polarity coincidence correlation. In Part I an account is given of investigations of the lead salts, PbS, PbSe, and PbTe, which have led to a consistent theory of photoconductivity and infrared photoconductive detectors. In Part II the theory of polarity coincidence correlation is reviewed, and some of the considerations involved in developing equipment for particular applications are discussed.

INTRODUCTION

THE primary responsibility of the Naval Ordnance Laboratory is the development of naval weapons systems. Although such systems vary widely in design and employment, there are certain basic problems common to all. These are associated with detection, localization and classification of the target and delivery and fuzing of the weapon. Practical air and underwater ordnance systems which have been developed to date employ electromagnetic and acoustic means for solving such problems. Various portions of both spectra prove to be especially suitable for such well-known applications as radar, VT-fuzes, missile guidance, sonar, homing torpedoes, and mine firing devices. It is not surprising, therefore, that a large part of the research program at the NOL is directed toward an increased understanding of the generation, propagation, and processing of electromagnetic and acoustic signals. Likewise, the research effort in electronics consists mainly of investigations on semiconductors for photoconductor and transistor applications; on magnetic materials, both alloys and ferrites, for magnetic amplifier, microwave, computer, and acoustic transducer applications; and on signal processing, particularly for use in underwater acoustic systems. Two of these programs have been selected for description in this paper: photoconductor research and applications of polarity coincidence correlation. The objective in the former has been to increase our fundamental knowledge of the mechanism of photoconductivity in order that optimum infrared detectors and circuits can be designed. In the latter, the emphasis has been on the development of circuits to implement some of the advances in signal processing theory.

PART I. PHOTOCONDUCTOR RESEARCH

A. Introduction

Research on semiconductors at the NOL is concentrated principally on materials having a small energy gap. This includes some intermetallic semiconductors such as InAs and InSb and the lead salt semiconductors

PbS, PbSe, and PbTe. Other materials studied, principally for surface effects, are germanium and silicon. Much attention is devoted to infrared detection and problems of optical filters in the infrared. This discussion will be confined to the infrared detection problem and will describe some of our contributions to this subject.

One of the major goals of the work has been an understanding of the mechanism of photoconductivity in PbS, PbSe, and PbTe films. This problem has been attacked by an integrated program involving studies of the properties of the bulk crystals, properties of thin films, surface effects, and the application of information theory to the performance of infrared detectors. As a result of this work it is now possible to understand the major aspects of the photoconductive mechanism in thin films of these materials and to understand their most effective use and operation as radiation detectors.

We shall first review some of the highlights of the program starting with investigations of the properties of single crystals of the pure compounds. Following this we shall describe our studies of the properties of the materials in the form of thin films; for it is in this form that these compounds display their greatest sensitivity to infrared radiation. Finally, we shall summarize some of the results of the application of information theory to the problem of radiation detectors.

B. Properties of Single Crystals of PbS, PbSe, and PbTe

Of fundamental importance in the understanding of photoconductivity is an understanding of related electrical and optical properties of the semiconductor. When the photoconductivity research program started at the NOL there was little information available about the basic properties of the lead salt semiconductors. Even such important quantities as the energy gaps were not established.

Much of the desired information could best be obtained from measurements made on pure single crystals of the compounds. The preparation of such crystals of compounds of polar or partially polar semiconductors presents problems not generally met in dealing with elementary semiconductors such as germanium and silicon. The composition and hence the electrical and optical properties of polar crystals are strongly dependent upon the deviation from stoichiometric composition as well as upon the presence of certain foreign atoms in the lattice.

In polar crystals the solid exists as a single phase over the small range of composition near the stoichiometric proportions. This implies some thermodynamical consequences which are important in the preparation of crystals having desired electrical properties. First, the

* Original manuscript received, December 1, 1958; revised manuscript received, February 13, 1959.

† U. S. Naval Ordnance Laboratory, Silver Spring, Md.

two-phase, two-component system such as a lead sulfide crystal in equilibrium with its vapor has two degrees of freedom. Thus, the composition of the crystal is fixed only when, for example, the temperature and vapor pressure of one of the components are fixed. Second, the two components will in general have different proportions in the crystal and in the vapor. Therefore, at equilibrium a polar crystal can be maintained at a fixed composition over a range of temperature only by appropriately varying the vapor pressure of one of its components.

The first thermodynamic consequence was employed in preparing a group of PbS crystals having a wide range of conductivities, both *n*-type and *p*-type.¹ The second was demonstrated in connection with the problems of interpreting electrical measurements made at high temperatures on polar crystals.²

Both of these concepts are illustrated in the phase diagram for the PbS crystal-vapor system shown in Fig. 1.

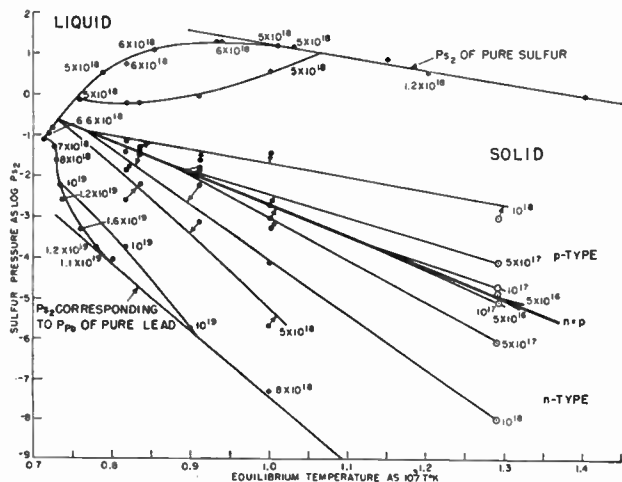


Fig. 1—Solid-vapor-phase diagram for the PbS system.

This diagram is similar to the one published by Bloem and Kroger³ except that it incorporates our data at the lower temperatures. Here a vertical section shows how the equilibrium composition of a crystal at a given fixed temperature varies with the sulfur pressure. A horizontal section shows the variation with temperature of crystal composition at a fixed sulfur pressure.

With a supply of crystals having a range of carrier concentrations from 10^{16} – 10^{19} /cm³ it was possible to study the behavior of various electrical properties as a function of composition.

Hall effect and resistivity curves were evaluated as a function of temperature. The curves show the custom-

ary intrinsic and extrinsic conductivity ranges.¹ From the intrinsic Hall data the quantity np/T^3 was evaluated and is plotted as a function of temperature in Fig. 2. The slope of this curve gives the value for the thermal activation energy at 0°K for PbS of 0.34 ev. This is in good agreement with optical data and removed one of the serious problems existing in the literature concerning the discrepancy between optical and thermal values that had previously been obtained for the energy gaps in these compounds. This work established that photoconductivity in the 1–3 micron range in PbS corresponds to a main-band transition and not to impurity photoconductivity.

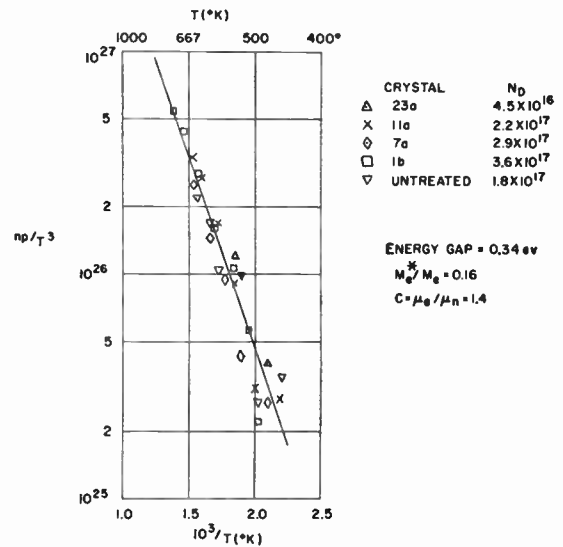


Fig. 2—The intrinsic energy gap in PbS obtained from Hall data.

The spectral distribution of sensitivity for the photo-voltaic effect in a *p-n* junction for each of the three compounds gives further evidence of the fact that photo-response in these compounds is a main-band transition phenomena.⁴ These results are shown in Fig. 3.

We have found that for PbS in the temperature range from 500°K to 80°K the mobility seems to be an intrinsic characteristic of the crystal, independent of carrier concentration from about 10^{16} – 10^{19} /cc.⁵ Attempts have been made to account for this by a combination of scattering by optical and acoustical vibrational modes of the crystal.⁶

Hall effect and resistivity measurements have been extended down to liquid helium temperature (4.2°K).⁷

⁴ W. W. Scanlon, unpublished material.

⁵ W. W. Scanlon, R. F. Brebrick, and R. L. Petritz, "The chemical and physical properties of PbS single crystals" in "Photoconductivity Conference," ed. by R. G. Breckenridge, B. R. Russell, and E. E. Hahn, John Wiley and Sons, Inc., New York, N. Y., pp. 619–635; 1956.

⁶ R. L. Petritz and W. W. Scanlon, "Mobility of electrons and holes in the polar crystal, PbS," *Phys. Rev.*, vol. 97, pp. 1620–1626; March 15, 1955.

⁷ R. Allgaier and W. W. Scanlon, "Mobility of electrons and holes in PbS, PbSe and PbTe between room temperature and 4.2°K," *Phys. Rev.*, vol. 111, pp. 1029–1037; August 15, 1958.

¹ R. F. Brebrick and W. W. Scanlon, "Electrical properties and the solid-vapor equilibrium of lead sulfide," *Phys. Rev.*, vol. 96, pp. 598–602; November 1, 1954.

² W. W. Scanlon, "Interpretation of Hall effect and resistivity data in PbS and similar binary compound semiconductors," *Phys. Rev.*, vol. 92, pp. 1573–1575; December 15, 1953.

³ J. Bloem and F. A. Kroger, "The *p-T-x* phase diagram of the lead sulfur system," *Zeit. Phys. Chem. Neue Folge*, vol. 7, pp. 1–14; April, 1956.

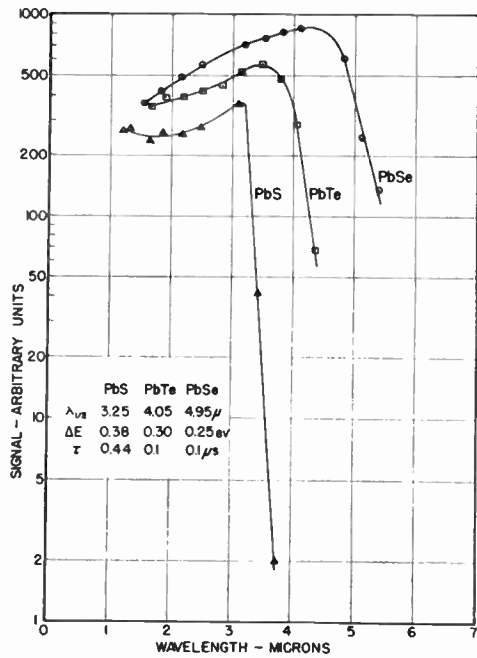


Fig. 3—Photovoltaic spectrum for *p-n* junctions in PbS, PbSe, and PbTe at room temperature.

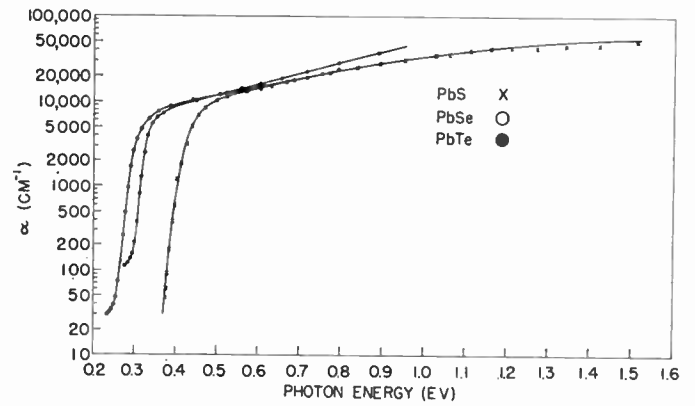


Fig. 4—Optical absorption edge in PbS, PbSe, and PbTe crystals.

The Hall coefficient is essentially constant down to the lowest temperature. The resistivity decreases with decreasing temperature and shows an absence of appreciable impurity scattering usually found in valence semiconductors. Electron mobilities approaching a million are found in PbTe crystals at 4.2°K. Most crystals of the three compounds show mobility curves which approach a constant value in 4.2°–50°K temperature range. The value reached is unrelated to carrier concentration.

We find a striking difference in the value of mobility of electrons and holes in thin photoconductive films. Woods⁸ and others⁹ have made Hall effect and resistivity measurements on films of PbS and find room temperature mobilities of the order of a hundred times smaller than in the crystals.

The spectral distribution of photoconductivity and optical absorption are related; hence it is important to obtain the detailed behavior of the optical absorption edge in these compounds. The complete optical absorption edge for the crystals from absorption coefficients of about 10 cm⁻¹–10⁵ cm⁻¹ was obtained using a special infrared microscope technique.¹⁰ For some measurements crystals had to be ground and polished down to less than one micron in thickness in order to get measurements of the absorption coefficient. The curves for

the three compounds are shown in Fig. 4.¹¹ There are three important applications of these results to photoconductivity. First the role of film thickness on the spectral distribution of photoconductivity can be calculated and explanations given for thick films appearing to have longer wavelength thresholds for photoconductivity than thin films.¹² The second application of the absorption data is the establishment of the energies for direct and indirect transitions across the forbidden energy gap in these semiconductors.¹¹ These values are listed in Table I. Finally one may calculate the time

TABLE I
ROOM TEMPERATURE VALUES FOR DIRECT AND INDIRECT TRANSITION ENERGIES, RADIATIVE RECOMBINATION RATE, AND RADIATIVE LIFETIME IN PbS, PbSe AND PbTe CRYSTALS

Material	Energy Gap (ev)		Radiative Recombination (Rate/cm ³)	Radiative Lifetime (μ s)
	Direct	Indirect		
PbS	0.41	0.37	22 × 10 ¹⁸	63
PbSe	0.29	0.26	14.6 × 10 ²⁰	8.0*
PbTe	0.32	0.29	8.9 × 10 ²⁰	6.8*

* Using m_e^*/m_e same as in PbS.

constant for radiative recombinations of electrons and holes in the three crystals from the shape of the absorption edge.¹¹ This was done using the van Roosbroeck and Shockley¹³ analysis and the results are given in the last column of Table I.

Studies of the carrier lifetime in crystals of these materials have been made using the PEM-photoconductivity method.¹⁴ The results show that in all three crys-

⁸ J. Woods, "Investigation of the photoconductive effect in lead sulfide films using Hall and resistivity measurements," *Phys. Rev.*, vol. 106, pp. 235–240; April 15, 1957.

⁹ J. L. Levy, "Sensitive Hall measurements on NaCl and photoconductive PbTe," *Phys. Rev.*, vol. 92, pp. 215–218; October 15, 1953. J. Lathrop, Ph.D. dissertation, Northwestern University, Evanston, Ill.; 1949.

¹⁰ W. W. Scanlon, "Intrinsic optical absorption and the radiative recombination lifetime in PbS," *Phys. Rev.*, vol. 109, pp. 47–50; January 1, 1958.

¹¹ W. W. Scanlon, "Recent Advances in the Electrical and Optical Properties of PbS, PbSe, and PbTe," Internatl. Conf. on Semiconductors, Rochester, N. Y.; August 18–22, 1958. To be published in *J. Chem. and Phys. Solids*.

¹² J. N. Humphrey and R. L. Petritz, "Photoconductivity in lead selenide: theory of the dependence of sensitivity on film thickness and absorption coefficient," *Phys. Rev.*, vol. 105, pp. 1192–1197; February 15, 1957.

¹³ W. Van Roosbroeck and W. Shockley, "Photon-Radiative recombination of electrons and holes in germanium," *Phys. Rev.*, vol. 94, pp. 1558–1560; June 15, 1954.

¹⁴ W. W. Scanlon, "Lifetime of carriers in lead sulfide crystals," *Phys. Rev.*, vol. 106, pp. 718–720; May 15, 1957.

tals the maximum lifetime is much less than the radiative recombination lifetime. This is to be compared to the experiments on photoconductive films of the three materials where lifetimes are many times greater than the radiative lifetime. The contrast between lifetimes observed in crystals and in photoconductive films strongly suggests a trapping mechanism in the films. More will be said about this later.

Information on the shape of energy band edges for the three compounds has been obtained from magnetoresistance studies.¹⁵ In PbTe the magnetoresistance behavior suggests that the surfaces of constant energy for both valence and conduction bands are ellipsoids of revolution oriented along $\langle 111 \rangle$ directions in K space with an eccentricity factor of from 4–6. Magnetoresistance effects are much smaller in PbS and PbSe and the eccentricity is smaller also. Since the differences in the various models decrease for decreasing anisotropy as far as magnetoresistance studies are concerned, it is not possible at this time to comment on the shape of the constant energy surfaces for PbSe and PbS.

C. Films of PbS, PbSe and PbTe

Theory of Photoconductivity: The work on single crystals has established important basic facts about photoconductivity in the lead salt semiconductors. The primary carriers generated by the radiation are electron-hole pairs resulting from photo excitation of electrons from the valence band to the conduction band. In crystals, recombination occurs very rapidly with the result that the photo current is extremely small. In films, on the other hand, the sensitization process appears to reduce the recombination rate and thereby increase carrier lifetimes by orders of magnitude. Photoconductive lifetimes in films exceed the lifetime for direct electron-hole recombination in crystals through a radiative process and therefore suggest a trapping mechanism in films.

The most important problem in understanding photoconductivity in films is to understand the mechanism by which the recombination rate of electrons and holes is reduced. A number of theories have been proposed which may be classified under three categories according to the carrier lifetime and recombination processes assumed in the theory. The first is an intrinsic carrier model in which electrons and holes recombine directly by a radiative process or at recombination centers such as dislocations in the crystal lattice.¹⁶ An obvious defect of this model when applied to the lead salts is the fact that these recombination processes are much too fast to give the observed lifetimes in films. The second category is the minority carrier model in which photoconductivity depends upon a long minority carrier lifetime. The film

is visualized as a composite of microcrystals of alternating n - and p -type material. In one theory of this type barrier modulation plays an important role.¹⁷ Petritz has proposed the third category, which is the majority carrier model in which photoconductivity results from the predominance of the majority carrier.¹⁸ In this model the free minority carrier produced by the photon is trapped in a deep-lying energy level, leaving the corresponding majority carrier free to conduct. We have performed a number of experiments which tend to confirm the majority carrier model.

Our studies of evaporated PbSe films¹⁹ show that although other oxidizing elements formed acceptor sites in the film changing the carrier sign from n - to p -type, only oxygen produced photoconductivity at room temperature. If either the intrinsic or minority carrier models were applicable, any acceptor element should produce photoconductivity. A plausible interpretation is that oxygen is the only acceptor element which acts as a minority carrier trap required by the majority carrier model. A probable trapping site in the p -type material could be a $(\text{PbO})^{++}$ ion which could attract and trap an electron, becoming $(\text{PbO})^+$. The $(\text{PbO})^+$ would repel holes, thus producing a trapping rather than a recombination center.²⁰ Other acceptor centers do not provide these conditions.

It is reasonable to expect some of the traps to be in the surface. Electrons trapped in the surface may result in a modulation of the barriers which exist between crystallites. Experiments were carried out therefore to observe barrier modulation effects. The Hall effect gives the carrier density within the crystallites when inter-crystalline barriers are present which are of high resistance compared to the crystal resistance.²¹ The film resistance on the other hand is strongly affected by barriers. Information on barrier amplification effects can therefore be obtained by studying changes in the Hall coefficient and resistivity of a photoconductive film under illumination. The experimental results obtained with chemically deposited PbS films show that over a wide range of illumination and temperature the fractional change in the Hall coefficient is equal to the fractional change in resistivity. Thus it follows that since no observable change in mobility can be found in a film under illumination, barrier modulation is negligible. The photoconductive effect is due entirely to an increase in the density of majority carriers in the PbS crystallites.⁸

¹⁷ J. C. Slater, "Barrier theory of the photoconductivity of lead sulfide," *Phys. Rev.*, vol. 103, pp. 1631–1644; September 15, 1956.

¹⁸ R. L. Petritz, "Theory of photoconductivity in semiconductor films," *Phys. Rev.*, vol. 104, pp. 1508–1516; December 15, 1956.

¹⁹ J. N. Humphrey and W. W. Scanlon, "Photoconductivity in lead selenide: experimental," *Phys. Rev.*, vol. 105, pp. 469–476; January 15, 1957.

²⁰ J. N. Humphrey and R. L. Petritz, "Photoconductivity in lead selenide: theory of the mechanism of sensitization," *Phys. Rev.*, vol. 105, pp. 1736–1740, March 15, 1957.

²¹ G. R. Wait, "Hall effect and specific resistance of silver films," *Phys. Rev.*, vol. 19, pp. 615–622; June, 1922.

J. Volger, "Note on the Hall potential across an inhomogeneous conductor," *Phys. Rev.*, vol. 79, pp. 1023–1024; September 15, 1950.

¹⁵ R. S. Allgaier, "Magneto-resistance in PbS, PbSe and PbTe at 295°, 77.4° and 4.2°K," *Phys. Rev.*, vol. 112, pp. 828–836; November 1, 1958.

¹⁶ A. Von Hippel and E. S. Rittner, "Thallous sulfide photoconductive cells. II. Theoretical discussion," *J. Chem. Phys.*, vol. 14, pp. 370–378; June, 1946.

While the preceding experiments strongly suggest a majority carrier photoconductor with traps for the minority carrier, additional experiments were desired to confirm the model and give details on the nature of the traps. Field effect studies provide this information. The field effect can be used to inject one type of carrier only and permits observation of its decay rate. Radiation, on the other hand, injects both carriers simultaneously and therefore is a somewhat more complicated injection experiment. Field effect experiments can yield information on the location of traps, their concentration and depth, so that these experiments offer critical evidence concerning the model of photoconductivity.

Field effect experiments on chemically deposited PbS films were made²² and the results may be summarized as follows. The time constant for field effect injected majority carriers is identical to the photoconductive time constant. The rate limiting process in a sensitized film is the rate of transfer of charges into the traps. Majority trap density is high while the capture cross section is very small, conditions required for the long majority time constants measured. An estimated density of surface traps required to provide the time constants measured is of the order expected in crystallites with dimensions of the order of 0.1 micron.

The field effect experiments support all the major features of the majority carrier model of photoconductivity in sensitized PbS films. Sensitization of a film means the introduction of the right density of surface states of the proper energy with the Fermi energy level in the semiconductor adjusted to the right value.

Responsivity of Photoconductive Detectors: Having established the model of photoconductivity in films it becomes possible to give an equation for the process. This is given in terms of a specific responsivity, R_s ,²³

$$R_s = \frac{(1 + B)\eta_s\tau}{4pdh\nu_s[1 + (\omega\tau)^2]^{1/2}} \quad (1)$$

where B = barrier modulation factor (zero in PbS films), η_s = quantum efficiency, τ = carrier lifetime, p = majority carrier mean density, d = film thickness, $h\nu_s$ = incident photon energy, and ω = angular frequency at which radiation is chopped. This definition normalizes the response to unit biasing voltage and unit incident radiation flux. R_s is a measure of the effectiveness with which a detector converts radiation flux into signal voltage. Responsivity calculated from the above equation is in excellent agreement with experiment and provides confirmatory evidence of the majority carrier model.

Chemical Deposition of Films: We have studied the chemical deposition process for films of PbS, PbSe, and PbTe. In particular we are interested in the relationship

of the response of films to radiation as a function of the method of preparation. This latter relationship has been studied intensively for PbS. A correlation between the photoconductive properties of the films and the chemical conditions prevailing during the deposition has been investigated.²⁴ The most sensitive films were produced only over a rather narrow range of the parameters defining the composition of the deposition solution. Within this range an impurity is incorporated within the film which acts as a minority carrier trap. Studies are also made of the kinetics of the reactions in the solution.

D. The Use and Operation of Photoconductive Infrared Detectors

In many applications for infrared detectors one is interested in knowing the minimum flux of infrared energy which can be detected. This limit is established by the electronic noise level in the system and is generally defined as the radiation flux which produces a response equal to the noise. The reduction of noise to the theoretical minimum is therefore one of the important problems in cell technology. It became clear early in our program that the prevailing concepts of noise behavior in infrared cells were incapable of accounting for the experimentally observed noise. An understanding of all the fundamental mechanisms for noise production in thin films of a semiconductor was needed. With this knowledge it is possible to predict such things as the ultimate sensitivity of a detector and its optimum characteristics for specific applications, and to provide an insight into the problem of improving cell characteristics. Our work on this subject is described in a number of publications^{25,26} and will only be summarized here.

The principal contribution is a description of an intrinsic noise mechanism occurring in semiconductors called generation-recombination or G - R noise, which remains as the dominant source of noise in a photoconductive cell when all other noise mechanisms are minimized. This noise, therefore, establishes the ultimate sensitivity of a cell.

This work has led to a resolution of the noise in a photoconductive cell into four components. First is the classical resistor noise described by Nyquist and often called "Johnson" noise which is due to fluctuations in

²⁴ F. L. Lummis, D. E. Brayton, R. F. Brebrick, E. Gubner, and R. L. Petritz, "Investigation of the sensitization of lead sulfide photoconductive films," *Bull. Amer. Phys. Soc.*, ser. 11, vol. 3, no. 2, p. 114; 1958.

²⁵ R. L. Petritz, "On the theory of noise in photoconductors," *Phys. Rev.*, vol. 86, p. 614A-615A; May 15, 1952. R. L. Petritz, "On the theory of noise in p - n junctions and related devices," *Proc. IRE*, vol. 40, pp. 1440-1456; November, 1952. F. L. Lummis and R. L. Petritz, "Noise, time constant and Hall studies on lead sulfide photoconductive films," *Phys. Rev.*, vol. 105, pp. 502-508; January 15, 1957. R. L. Petritz, "Information Theory of the Performance of Radiation Detectors," *IRIS Proc. Infrared Information*, vol. 2, pp. 18-34; June, 1957.

²⁶ R. L. Petritz, "The relation between lifetime, limit of sensitivity, and information rate in photoconductors," in "Photoconductivity Conference"; ed. by R. G. Breckenridge, B. R. Russell, and E. E. Hahn, John Wiley and Sons, Inc., New York, N. Y., pp. 49-77; 1956.

²² H. E. Sorrows, "Fast Field Effect and Photo Response Studies on Lead Sulfide Photoconductors," NAVORD Rep. No. 6164 (unpublished).

²³ R. L. Petritz and J. N. Humphrey, "Research on Photoconductive Films of the Lead Salts," NAVORD Rep. No. 6042 (unpublished).

carrier velocities. Second is shot noise at intercrystalline barriers. Third is noise which varies inversely with frequency and is called $1/f$ noise. Finally, there is the $G-R$ noise. Fluctuations in a number of processes concerning the interactions of electrons and holes with the radiation field and the crystal lattice contribute to the $G-R$ noise. Among these are fluctuations in the rate of incident photons on the detector, fluctuations in the quantum efficiency, fluctuations in the lifetime of conducting electrons and holes, and fluctuations in the interactions with the lattice. Taking all these into consideration a solution for the generation-recombination noise spectrum has been obtained.²⁶ One important consequence of this solution is that the spectrum has the form $1/[1+(\omega\tau)^2]$ where τ is the photoconductive lifetime. This component of the noise spectrum has the same form as the curve for the frequency dependence of photoconductive response to chopped radiation. See (1).

The general noise spectrum for a photoconductor containing several noise mechanisms is given as the sum of the above components.

$$G \left[I^2 \left(G - R, \frac{1}{f}, \text{shot}, R \right) \right] = \frac{4\tau I_{dc}^2(1+B)^2}{A\rho d[1+(\omega\tau)^2]} + \frac{CI_{dc}}{fAd} + \frac{2qI_{dc}}{N} + \frac{4kT}{R}, \quad (2)$$

where I_{dc} is the direct current through the cell, A is the cell area, C is a constant, N is the number of barriers, and R is the cell resistance. Generally the third term in the noise equation is negligible because of the large number of barriers. The general form of a noise curve is given in Fig. 5. The noise spectrum described by this theory is in excellent agreement with the experiments on PbS cells.

The shape of the noise curve for a detector tells which fluctuation mechanism is limiting the sensitivity. If the noise spectrum is of the $1/f$ type, one can conclude that the detector is not very sensitive. Improvements in the technology of manufacturing cells have led to a large reduction in $1/f$ noise. The most sensitive detector is one where the sensitivity is limited only by $G-R$ noise. The hump in the noise curve of Fig. 5 occurs at a frequency $f_r = 1/(2\pi\tau)$.

One other result of the noise studies concerns the application of infrared cells for detecting signals. There are two general kinds of applications. In one the objective is to determine the presence of a signal, while in the other the objective is to reproduce faithfully the detailed shape of the signal. The first is a weak signal application such as in spectroscopy where line identification is important. The second is a strong signal case such as the infrared signature of a military device.

Information theory shows that in the case of weak signal detection the information rate depends linearly on the noise level.²⁶ In the strong signal case improvements in noise level increase the information rate loga-

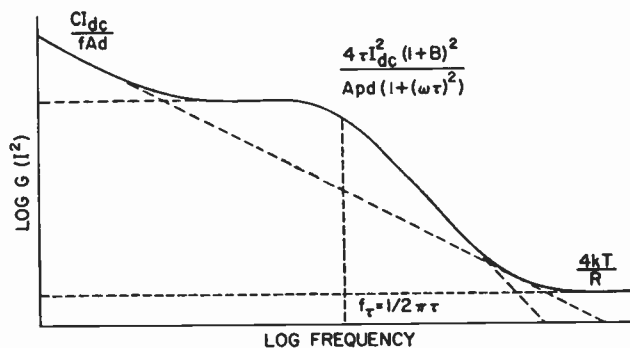


Fig. 5—Plot of current-noise spectrum vs frequency for a photoconductive film.

arithmically. The result is that the information rate in the large signal case is essentially dependent only on the bandwidth of the detector, with a short time constant detector yielding the greatest information rate.

The theory shows that the steps taken to produce shorter time constant cells for certain applications also reduce their over-all sensitivity. Although it is not commonly practiced it is possible with a simple addition to the amplifier system to extend greatly the frequency range of a long time constant detector, thereby effectively shortening its time constant.²⁶ This is because cells which have $G-R$ limited noise have the same frequency dependence of signal and noise. Thus the higher inherent sensitivity of a long time constant cell, coupled with the electronic compensation for extending its bandwidth, will give the optimum detector for faithful reproduction of a strong infrared signal. These conclusions have recently been verified by experiments at the NOL.

PART II. CORRELATION METHODS

A. Introduction

Applications of correlation methods to signal processing problems have been carried out at the NOL since 1950 in the fields of microseisms and underwater acoustics. Because of the large amounts of input data and the need to process the incoming data in real time, the primary method used has been polarity coincidence correlation. This refers to the process of cross-correlating extremely clipped time functions rather than the original time functions. These applications have involved the detection and localization of noise sources. The desired information is extracted by measuring cross-correlation functions between spaced detector outputs. The following sections give a short discussion of some basic considerations in applying correlation methods, a discussion of methods of instrumenting polarity coincidence correlation, and an example which illustrates its application to microseismic data analysis.

B. Basic Considerations

Consider a source of radiating energy which can be assumed to be a random function of space and time. An important observational problem is to determine when

the source is present (the detection problem), to estimate its space coordinates (the localization problem), and to estimate some of its statistical characteristics (the classification problem).

If we have a controlled source, such as one might have in a point-to-point communication system, and if we have a channel containing statistically independent additive noise, detection, localization, and classification can be accomplished by the observer if he cross-correlates a replica of the source, $II_1(u) = \sqrt{S_1}S(u)$, with the incoming signal plus noise $II_2(u) = \sqrt{S_2}S(u + \theta) + \sqrt{N_2}N(u)$. Assume that $S(u)$ and $N(u)$ have zero mean values and unit mean square values. By cross-correlation is meant the calculation of the function

$$A_T(t) = \frac{1}{T} \int_0^T H_1(u + t)H_2(u)du. \quad (3)$$

If the signal is present, the cross-correlator output is

$$A_T(t) = \sqrt{S_1 S_2} \rho(t - \theta) + \alpha_1(t, T). \quad (4)$$

This assumes that $S(t)$ is a stationary random function with auto-correlation function equal to $\rho(x)$. $\alpha_1(t, T)$ represents noise due to the finite averaging time T . Ideally, A_T is a function of a continuous time variable t . Since $\rho(t - \theta)$ assumes its maximum value at $t = \theta$, the localization parameter θ can be estimated by estimating the value of t for which A_T is a maximum. Detection can be accomplished by visual observation of $A_T(t)$ since if the signal is not present, the signal auto-correlation function cannot be present.

If we have an uncontrolled source, the observer can extract desired information by cross-correlating the outputs of spaced detectors. Consider, for example, two detectors II_1 and II_2 whose outputs can be written

$$\begin{aligned} H_1(u) &= \sqrt{S_1}S(u) + \sqrt{N_1}N_1(u), \\ \text{and} \\ H_2(u) &= \sqrt{S_2}S(u + \theta) + \sqrt{N_2}N_2(u). \end{aligned} \quad (5)$$

Assume that $S(u)$, $N_1(u)$ and $N_2(u)$ have zero mean values and unit mean-square values. Cross-correlation in this case refers to calculation of the function

$$A_T(t) = \frac{1}{T} \int_0^T H_1(u + t)H_2(u)du. \quad (6)$$

If the signal is present, the cross-correlator output is

$$A_T(t) = \sqrt{S_1 S_2} \rho(t - \theta) + \alpha_2(t, T). \quad (7)$$

This assumes that $S(t)$, $N_1(t)$, and $N_2(t)$ are statistically independent stationary random functions, and that $S(t)$ has auto-correlation function $\rho(x)$. $\alpha_2(t, T)$ represents noise due to the finite averaging time T . Detection and localization information can be obtained as above. The shape of A_T contains classification information.

Polarity coincidence cross-correlation refers to calculation of the function

$$B_T(t) = \frac{1}{T} \int_0^T \xi_1(u + t)\xi_2(u)du. \quad (8)$$

$\xi_i(u)$ represents the extremely clipped version of $II_i(u)$. Thus $\xi_i(u)$ is a random square wave which is equal to $+1$ or -1 depending upon whether $II_i(u)$ is positive or negative.

If the input signals and noises can be assumed to be stationary Gaussian processes, then this cross-correlator output is²⁷

$$B_T(t) = \frac{2}{\pi} \arcsin \left[\sqrt{\frac{S_2}{S_2 + N_2}} \rho(t - \theta) \right] + \beta_1(t, T) \quad (9)$$

for the controlled source case, and²⁷

$$\begin{aligned} B_T(t) &= \frac{2}{\pi} \arcsin \left[\sqrt{\frac{S_1 S_2}{(S_1 + N_1)(S_2 + N_2)}} \rho(t - \theta) \right] \\ &+ \beta_2(t, T) \end{aligned} \quad (10)$$

for the uncontrolled source case.

For large input signal-to-noise ratios, the mean value of $B_T(t)$ has a triangular peak at $t = \theta$, which is useful in estimating θ . For small input signal-to-noise ratios, the mean value of $B_T(t)$ is proportional to $\rho(t - \theta)$.

For comparison purposes, it is interesting to consider also the "ring modulator correlator." Cross-correlation in this case refers to calculation of the function

$$C_T(t) = \frac{1}{T} \int_0^T H_1(u)\xi_2(u + t)du. \quad (11)$$

Here only one input channel is extremely clipped. If the input signals and noises can be assumed to be stationary Gaussian processes, then this cross-correlator output is

$$C_T(t) = \sqrt{\frac{2}{\pi}} S_1 \rho(t - \theta) + \gamma_1(t, T) \quad (12)$$

for the controlled source case, and

$$C_T(t) = \sqrt{\frac{2}{\pi}} \frac{S_1 S_2}{S_2 + N_2} \rho(t - \theta) + \gamma_2(t, T) \quad (13)$$

for the uncontrolled source case.

An important feature of polarity coincidence correlation is the fact that we have a normalized correlation function. In ring modulator correlation the clipped channel is normalized. This feature eliminates the need for automatic gain controls to take care of short time variations in the mean-square value of the input signal plus noise. Since the problem is to measure the degree of similarity of two complex waveforms, and there is no need to reconstitute an original waveform as in some communication problems, the information loss due to extreme clipping need not be great. The following discussion supports this conclusion.

²⁷ J. J. Faran, Jr. and R. Hills, Jr., "Correlators for Signal Reception," Acoustics Res. Lab. Harvard Univ., Cambridge, Mass., Tech. Memo. No. 27; September 15, 1952.

If the input signals and noises can be assumed to be stationary and Gaussian with the same normalized correlation function $\rho(x)$, simple approximation formulas can be written for the output signal-to-noise ratios in the three types of correlators, when the input signal-to-noise ratios are small.^{27,28} By output signal-to-noise ratios is meant the ratios of the square of the mean correlator outputs to the mean square values of the corresponding noise terms α_i , β_i and γ_i for $t=\theta$. Under these assumptions, the output signal-to-noise ratio is equal to

$$K_i T \frac{S_1}{N_1} \quad (14)$$

for the controlled noise source case, and equal to

$$K_i T \frac{S_1 S_2}{N_1 N_2} \quad (15)$$

for the uncontrolled noise source case. The coefficients K_i are as follows.

For the multiplier correlator,

$$(K_A)^{-1} = 2 \int_0^T \left(1 - \frac{x}{T}\right) \rho^2(x) dx. \quad (16)$$

For the polarity coincidence correlator,

$$(K_B)^{-1} = 2 \int_0^T \left(1 - \frac{x}{T}\right) [\arcsin(\rho(x))]^2 dx. \quad (17)$$

For the ring modulator correlator,

$$(K_C)^{-1} = 2 \int_0^T \left(1 - \frac{x}{T}\right) \rho(x) \arcsin(\rho(x)) dx. \quad (18)$$

If the correlation function $\rho(x)$ is equal to $\exp(-ax)$, which corresponds to a low-pass type of signal power spectrum with a cutoff frequency of a radians, $K_A \cong a$, $K_B \cong (0.759)a$, and $K_C \cong (0.876)a$. These approximations apply for large values of $a^2 T$, and small values of $\exp(-aT)$.

A similar analysis can be done for sampling correlators with statistically independent, stationary, Gaussian input data.²⁹ In this case the multiplier correlator calculates

$$A_n(t) = \frac{1}{n} \sum_{j=1}^n I_1(u_j) I_2(u_j + t). \quad (19)$$

The polarity coincidence correlator calculates

$$B_n(t) = \frac{1}{n} \sum_{j=1}^n \xi_1(u_j) \xi_2(u_j + t). \quad (20)$$

The ring modulator correlator calculates

$$C_n(t) = \frac{1}{n} \sum_{j=1}^n I_1(u_j) \xi_2(u_j + t). \quad (21)$$

n is equal to the number of statistically independent products.

For such sampling correlators, the output signal-to-noise ratio for small input signal-to-noise ratios is equal to

$$M_i n \frac{S_1}{N_1} \quad (22)$$

for the controlled noise source case, and

$$M_i n \frac{S_1 S_2}{N_1 N_2} \quad (23)$$

for the uncontrolled noise source case. $M_A = 1$ for the multiplier correlator, $M_B = 4/\pi^2 \cong 0.405$ for the polarity coincidence correlator, and $M_C = 2/\pi \cong 0.636$ for the ring modulator correlator.

These considerations show that the processing gain relations in the three types of correlators are basically the same for small input signal-to-noise ratios. The simplicity of the polarity coincidence and ring modulator correlators is bought at the price of reduced processing gain. However, in many applications this lost processing gain can be bought back at the price of increased integration time and/or bandwidth.

C. Practical Considerations

Each of the three correlation methods involves the use of three primary processes: 1) the delay of one of the two inputs, 2) the multiplication of the delayed and the undelayed inputs, and 3) the time averaging of this product. For the multiplier correlator (19), the two inputs are analog voltages. For the ring modulator correlator (21), one of the inputs is an analog voltage while the other input has been quantized into one of two possible voltage levels (1 or -1). For the polarity coincidence correlator (20), both inputs have been so quantized. In going through these three primary processes, one calculates the value of the correlation function for a single value of its argument (delay time). In order to compute the function, this must be repeated for enough values of delay to characterize adequately the function. This means that either the inputs have to be repeatedly played back through the processing system with a different value of delay associated with each playback or else played simultaneously through many processing systems in parallel, each with a different delay associated with it. For the multiplier correlator of broadband signals, rapid processing becomes extremely expensive if not virtually impossible. This is due to the practical difficulty of obtaining a multiplier which will handle analog voltages over a wide dynamic range and bandwidth. By using extreme care, one can obtain analog multipliers which will operate over a range of about

²⁸ W. B. Davenport, Jr., R. A. Johnson, and D. Middleton, "Statistical errors in measurements on random time functions," *J. Appl. Phys.*, vol. 23, pp. 377-388; April, 1952.

²⁹ Y. W. Lee, T. P. Cheatham, and J. B. Wiesner, "The application of correlation functions in the detection of small signals in noise," *Proc. IRE*, vol. 38, pp. 1165-1171; October, 1950.

40 db with a bandwidth of 15 kc. However, the expense and complexity of such a multiplier virtually rules out the possibility of parallel processing. The limited bandwidth virtually excludes the use of any time compression scheme (time division multiplexing) for speeding up the processing except in the case of extremely narrow band input signals. It is for this reason that the multiplier correlator has never received wide usage in real-time correlation processing systems even though its usefulness has been known for many years.

From the standpoint of widespread use as a practical means of processing data, the ring modulator correlator is an older method than that of the polarity coincidence correlator. However, its main use until recent years has been in the crosscorrelation of an input signal with a discrete frequency component where the discrete component is the clipped signal. An example of this use is found in the NOL Narrow-Band Lock-in Analyzer.³⁰ The advantage of the ring modulator method of correlation is the relative simplicity with which the multiplication is handled. An advantage of the ring modulator over the polarity coincidence correlator, in some cases, is the lack of normalization in the integrated output of the ring modulator. This should provide a real advantage in circuits where the correlator output is used as a control voltage. A much wider use of this method of correlation is seen for the future.

Polarity coincidence correlation was first introduced as a practical means for processing data at NOL by H. E. Ellingson in 1950 in a study of microseisms. The major advantage of this type of correlation is the simplicity of the processing elements. For clarification, let us consider the implementation of the Multiplier Correlator in a digital fashion. We would first feed both input channels through analog-to-digital converters where samples of the analog information are frequent enough to preserve all of the original information. A delay is provided for one channel either before or after the conversion. The multiplication is then accomplished with a digital multiplier. The averaging may be accomplished by digital methods (counting) or by a digital-to-analog conversion of the multiplier output followed by analog averaging. The polarity coincidence correlator can be considered as a special case of the above where the sampled amplitudes are each characterized by single binary digits. A "1" represents a positive polarity while a "0" represents a negative polarity. The product of these two amplitudes is also characterized by a single binary digit. However, the product we are seeking is not the product of the two binary representations which would be "1" in case both polarities are positive and "0" otherwise, but is rather that process which is "1" when the polarities agree and "0" when they disagree. If a sample of one input is considered as E while the corresponding sample of the other input is F , then the desired

product is obtained by the implementation of the following binary logic: $(E \text{ AND } F) \text{ OR } (\text{NOT } E \text{ AND NOT } F)$. The implementation of the above logic is so simple that parallel processing now becomes practical in a processing system, *i.e.*, a large number of parallel circuits are feasible. The rate at which this logic may be accomplished (10 million times a second) is such that time compression (time division multiplexing) of the inputs is practical. Another advantage of the polarity representation of the input signals is the way they lend themselves to binary storage systems. For example, a shift register accepting information from an input so characterized becomes a tapped delay line. An example of the use of the shift register as a tapped delay line and a parallel system of processing is the Transistorized Shift Register Polarity Coincidence Correlator.³¹ This correlator utilizes a 20-element transistorized shift register and 20 parallel processing elements. Averaging is performed by analog integration. The outputs then provide a correlogram (an observed correlation function) for 20 values of delay. These outputs are multiplexed by a mechanical commutator which allows the correlogram to be displayed on a cathode ray oscilloscope.

The development which provided the most impetus to the use of polarity coincidence correlation was that of the DELTIC by Anderson.^{32,33} The DELTIC provides for the storage of sampled polarity information in a circulating delay line such that the entire stored output is available for processing between samples. Each new sample wipes out the oldest stored sample in the delay line. The output of such a delay line is called a moving time series (MTS) since each circulation represents the latest samples in time. If Δ is the width of the storage slot in the delay line and $n\Delta$ is the length of the line, where n is the storage capacity, then the above is accomplished by making the sampling interval equal to $(n+1)\Delta$. If samples from the MTS DELTIC are dumped into a circulating delay line memory of length $(n+1)\Delta$, then a stationary time series (STS) DELTIC is formed. It is so called because the output is stationary in time, *i.e.*, no new information is fed to the STS when the input is sampled. Fig. 6 shows a moving and a stationary time series DELTIC in block diagram form. When the outputs of the MTS DELTIC and the STS DELTIC are compared, the result is the comparison of the same pulse train shifted by one sampling period for each circulation of the DELTICs after the dumping process. If the dumping is programmed for every $n+1$ samples and if the logic of the polarity coincidence correlation is applied to the outputs of the DELTICs, then the result is a correlation function where $n+1$ points

³¹ C. R. Greene, Jr., "Measurement of Polarity Coincidence Correlation Functions Using Shift Register Elements," NAVORD Rep. No. 4396; October 19, 1956.

³² V. C. Anderson, "The DELTIC Correlator," Acoustics Res. Lab., Harvard Univ., Cambridge, Mass., Tech. Memo. No. 37; January, 1956.

³³ J. C. Munson and L. E. Barton, "Delay Line Time Compressor XT-1A," NAVORD Rep. No. 4244; March 7, 1956.

³⁰ A. Z. Robinson, "Narrow-Band Lock-in Analyzer and Slave Analyzer," NAVORD Rep. No. 4420; 1957.

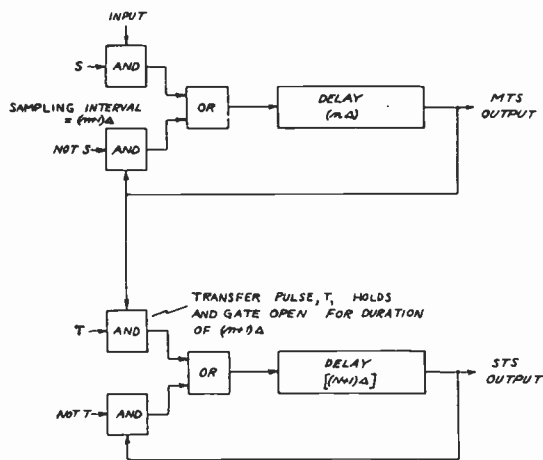


Fig. 6—Block diagram MTS and STS DELTICs.

have been calculated. A total of n comparisons is used for the calculation of each point. The correlation described here is autocorrelation. However, if the STS DELTIC and the MTS DELTIC are associated with different time series, the result is a cross-correlation function. What has been described here is a simple DELTIC system in which the output of the MTS DELTIC is a time compressed pulse train and the order is natural, *i.e.*, the compressed time series has the same order in time as the pulse train obtained from the sampling of the original time series. It is possible to have modifications of the above DELTIC (MTS) in which the stored order is not natural. An example is a sampling period, $T = (n+1)\Delta$, and a delay line length of $[m(n+1) - 1]\Delta$. The order of storing will then be 1, $m+1$, $2m+1$, \dots , $nm+1$, 2, $m+2$, $2m+2$, \dots , $nm+2$, 3, $m+3$, $2m+3$, \dots , $nm+n-1$, n , $n+m$, $n+2m$, \dots , $mn+n$. There are uses for DELTICs other than for correlation. For example, the MTS DELTIC may be used as a variable delay of an input pulse train. This is accomplished by sampling the output of the MTS DELTIC at increments of Δ after the sampling of the input time series. In the original DELTIC of length $n\Delta$ and input sampling period of $T = (n+1)\Delta$, the delay in the output of the DELTIC is $(n+1-r)T$ when the time between the sampling of the input signal and the sampling of the DELTIC output is $r\Delta$. Here increments of Δ in the sampling times of the input and output yield increments of T in delay. A DELTIC employing a line of length $[m(n+1) - 1]\Delta$ can be used in the same way to give a variable delay. The expression for the delay time in terms of the time between sampling of the input and output is not quite so simple. However, in general, increments of Δ in sampling the input and output will yield increments of mT of delay. The use of DELTICs in a correlator poses a serious problem, *i.e.*, that of achieving long integration times. As can be seen, the maximum amount of direct integration is the delay between samples multiplied by the number of stored samples in the line. In general this is not enough to provide a stable correlogram for low signal-to-noise ratios. The

result is a noisy correlogram produced at periods of $(n+1)T$. The problem is to pick out the correlogram from the noise—the post-integration problem.

The Dielectric Recorder^{34,35} has been successfully used as a post-integrator. It utilizes a rotating drum with a dielectric surface, the drum speed being synchronized with the period of the correlogram. An ion source is used to charge the dielectric surface through a slit. The ion source is controlled by the analog voltage which represents the correlogram. The rate at which the dielectric can charge, and consequently the integration time, is controlled by the slit width. While this method of post-integration has achieved the greatest success, it has two major disadvantages: contamination of the dielectric surface and lack of exact control of the integration time. The first disadvantage may be remedied by enclosing the drum in a controlled atmosphere.

There has been some work on a digital post integrator. This involves storing the correlogram in digital form and allowing the new correlogram being generated during each period to change the stored correlogram. The rate at which the new correlogram is allowed to change the stored correlogram determines the effective integration time. While this system is very complex, it can achieve any desired accuracy. With the new developments in core matrix type memories and magnetostrictive delay lines, it is expected that this is the most promising method for future development.

D. Applications

The earliest application of polarity coincidence correlation at the NOL occurred in 1950 in the processing of microseismic data. During the spring of that year a number of field stations were operated for the purpose of obtaining recordings of microseismic background.

Briefly stated, microseisms can be thought of as small earth displacements.³⁶ Low atmospheric pressure areas over the ocean serve in many cases as the exciting source.³⁷ They travel over land as a surface wave with the vertical component of displacement 90° out of phase with the horizontal component.³⁸ In general the recordings of microseisms show a strong periodic component.³⁹ The periods that predominate depend upon the water depth and the properties of the ocean bottom. Periods recorded on the Atlantic Coast are shorter than those recorded on the Pacific Coast.

³⁴ V. C. Anderson, "A Recording-Reproducing System Using a Dielectric Storage Medium," Scripps Inst. Oceanography, Marine Physical Lab., La Jolla, Calif.; January, 1955.

³⁵ B. Snavely, "Dielectric Integrator and Delay Mechanism XD-1A," NAVORD Rep. No. 4243; October 25, 1956.

³⁶ M. B. Dobrin, R. F. Simon, and P. L. Lawrence, "Rayleigh waves from small explosions," *Trans. Amer. Physical Union*, vol. 32, pp. 822-832; December, 1951.

³⁷ M. H. Gilmore, "Microseisms classified according to type of storm," *Bull. Seis. Soc. Amer.*, vol. 27, p. 473; October, 1946.

³⁸ J. L. Jones, C. B. Leslie, and L. E. Barton, "Acoustic characteristics of a lake bottom," *J. Acous. Soc. Amer.*, vol. 30, pp. 142-145; February, 1958.

³⁹ F. Press and M. Ewing, "A theory of microseisms with geologic applications," *Trans. Amer. Geophys. Union*, vol. 29, pp. 163-174; April, 1948.

The vertical component of earth displacement and the two horizontal components were recorded on Esterline-Angus (E-A) paper at each station. Also WWV time marks were recorded each second. A typical E-A recording is shown in Fig. 7. The traces appear distorted because of the curvilinear nature of the recording.

To characterize the microseismic background, a vertical component was correlated with a horizontal component as recorded at a single station. Also, corresponding components were correlated from the outputs of seismographs separated in distance from a fraction of a kilometer to approximately 1600 kilometers.

Since no correlation instrumentation was available for handling data recorded on E-A paper, the data was converted into binary form by observing the polarity of the seismograph output at each of the time marks and indicating the polarity with a plus sign if positive and a minus sign if negative. One hundred such observations were made from the recorded output of a seismograph located at one of the stations and can be thought of as a reference strip corresponding to the stationary time series in the DELTIC system.

Another similar strip of several hundred observations was made from the recorded output of a seismograph located at one of the other stations. This corresponds to the moving time series in the DELTIC system. The reference strip was placed along this second strip and the number of agreements in sign was observed from which a correlation coefficient r_k was computed from the formula

$$r_k = \frac{N_2 - N_1}{N_2 + N_1} \quad (24)$$

N_2 is the number of agreements and N_1 or $100 - N_2$ is the number of disagreements. Since 100 observations were made at the time marks for the reference strip, the integration time is 100 seconds. The reference strip was then moved a distance corresponding to one second of delay, and another correlation coefficient was computed. The variable k refers to the integral values of delay for which correlation coefficients were calculated.

Fig. 8 shows two correlograms that were obtained from data recorded on March 17, 1950, from stations located at Luverne, Minn.; Flint Hill, Va.; and the cave located near Front Royal, Va. The distance between the cave and Flint Hill is 20 km, while the distance between Luverne and the cave is 1600 km. The similarity between these correlograms is apparent when allowance is made for time displacements corresponding to the travel time of microseisms between stations.

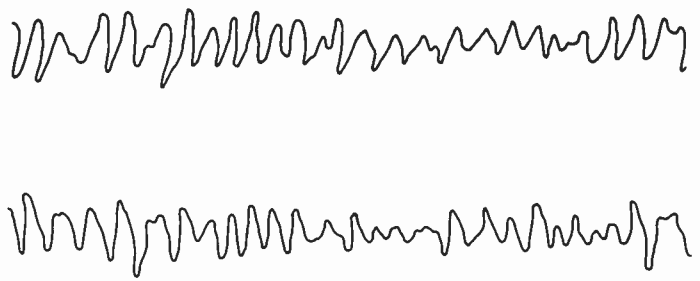


Fig. 7—Typical microseisms recorded on Esterline-Angus tape.

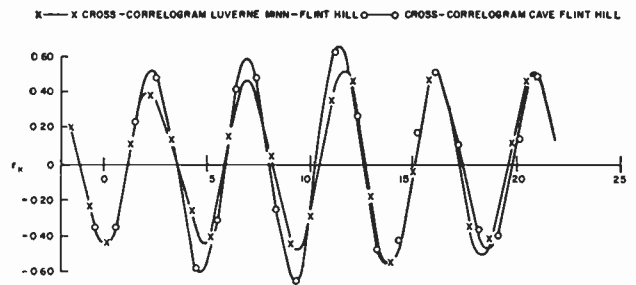


Fig. 8—Comparison of two cross-correlograms.

The ground displacement at Flint Hill, the cave, and at Luverne for the times considered was, respectively, 1.24, 1.22, and 0.54 microns. The comparison of the two correlograms thus indicates that the local interfering noise background level at Flint Hill was considerably greater than at Luverne.

In general the microseismic background content was large in comparison to displacements that took place due to local disturbances. For this reason it was possible to use a fairly small sample size. From the type of correlograms shown in Fig. 8 the velocity of microseisms can be obtained. With the long base line available, it has been possible to observe a change of direction of arrival of the microseisms in a relatively short period of time.

Numerous additional applications of polarity coincidence correlation have been made. The correlator has become a standard tool in the analysis of acoustic data. These applications have included signal detection, the analysis of water surface, and bottom reflections,³⁸ studies of the effects of transmission paths on the coherent characteristics of acoustic signals, and analyses of bottom topography data.

ACKNOWLEDGMENT

Part II-C was written by A. Z. Robinson. The application to microseismic data analysis was written by H. E. Ellington.

Charge Release of Several Ceramic Ferroelectrics Under Various Temperature and Stress Conditions*

L. W. DOREMUS†

Summary—Applications exist which require data on the large signal properties of ferroelectrics at room and elevated temperatures. This paper presents experimental results of measurements on the polarization change, piezoelectric coefficient, and pyroelectric charge release for several polycrystalline compounds. The piezoelectric properties are expressed as a function of stress (both statically and dynamically) up to the crushing strength of the material, while the sample temperature ranges from ambient to its Curie point. From mechanical, electrical, and temperature considerations, evaluations are offered for barium titanate, lead titanate-lead zirconate, and lead metaniobate.

INTRODUCTION

MECHANICAL and thermal stimuli alter the polarization of a ferroelectric, resulting in a charge release which can be gainfully employed in many electrical systems. Accordingly, experiments were conducted to obtain charge release data for ceramic ferroelectrics over wide temperature and stress ranges under both static and dynamic conditions.

PIEZOID

A ceramic ferroelectric exhibiting a net polarization, perpendicular to its electrodes, resulting from the past influence of an externally applied field (poling) will be referred to hereafter as a piezoid.

CHARGE DENSITY

A piezoid, due to its remanent polarization, binds a surface charge density σ_s' on its electrodes. The density of bound charge depends on the piezoid material, its temperature and state of stress, and the method of poling employed. The polarization of a piezoid can be altered by applying a stress (the piezoelectric effect) or by changing its temperature (the pyroelectric effect). When the polarization is altered by either of these two effects to some new value σ_s' a charge density $\sigma_1 = \sigma_s' - \sigma_1'$ is made available for practical use to either develop a potential or produce a current in some load. The sign of σ_1 is determined by the directional characteristics of the stimuli relative to the original poling direction as shown in Fig. 1.

* Original manuscript received by the IRE, November 26, 1958; revised manuscript received, February 27, 1959. The work performed for this paper will be submitted as part of the investigations being performed for a Master's thesis. This is being done with the permission of Prof. W. Sullivan, Head, Elec. Eng. Dept., Stevens Inst. Tech., Hoboken, N. J.

† Applied Res. Sec., Picatinny Arsenal, Dover, N. J.

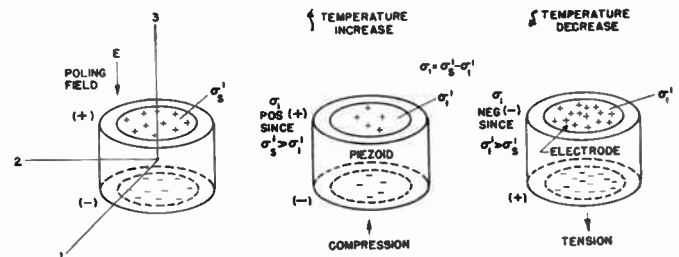


Fig. 1—Qualitative representation of charge release for mechanical and thermal stimuli relative to the original poling direction. σ_s' = charge density bound on electrode after poling. σ_1' = bound charge density after influence of temperature and stress. σ_1 = charge density released by temperature and stress influence.

CURIE TEMPERATURE CHARGE RELEASE

To determine the polarization σ_s' the piezoid was heated through its Curie temperature. In process of heating the piezoid, the total charge released was measured by a sensitive current integrator. The measured charge density σ_1 is numerically equal to the original polarization σ_s' since all bound charges are released by this method and σ_1' becomes zero. The average values of polarization for each type of piezoid investigated are given in line one of Fig. 2.

CHARGE RELEASE BY STATIC STRESSING

By stressing the piezoid up to its crushing strength and simultaneously monitoring its charge release, a continuous plot of polarization vs stress was obtained. The piezoid was mounted between parallel steel plates while its temperature was regulated at various levels and maximum stress was applied in approximately two seconds. A strain gauge calibrated in terms of the force applied to the piezoid, provided the horizontal signal on an *x-y* recorder. The charge developed by the piezoid under short-circuit conditions was presented simultaneously as the vertical signal. The results for several piezoid types are given in Figs. 3–8. The maximum polarization released by the application of stress for several piezoid types is given in line two of Fig. 2. By comparison with line one, it is seen that barium titanate is able to release all bound charge by stressing while lead titanate-lead zirconate releases about 64 per cent of its original polarization before it crushes.

The charge density measured on release of the stress was only a few per cent of that measured during the

application of stress, showing almost total degradation of polarization. It is important to note, however, that depolarizing by mechanical stresses of high magnitude is time dependent. It has been found in agreement with results previously presented [1] that large stress applied in the order of milliseconds has negligible depolarizing effects. Also, the polarization corresponding to the peak stress in a piezoid due to millisecond duration mechanical forces has the same magnitude as that obtained by statically stressing to the same stress level. These findings imply that a high-signal mechanical-to-electrical converter device such as the one described by Mason [2] could be designed using the virgin polarization values of a sample rather than the degraded value normally associated with static stressing.

In Figs. 3-8 the average piezoid coefficient \bar{d}_{33} is represented as the ratio of the coordinates of the charge-force curve and is irreversible for high static stresses.

For increasing frequency of applied mechanical stresses the coefficient becomes nearly reversible.

CHARGE RELEASE FOR DYNAMIC STRESSING

For dynamic testing the piezoid was mounted between two cylindrical steel rods which were insulated from each other and held in place by a connecting plastic tube. A light projectile fired from an air gun impacted the end of one rod at a high velocity, transmitting a high-magnitude, short-duration mechanical pulse to the piezoid. The piezoid was pulverized in approximately 4 μ sec. The total charge developed was measured, and the average values of charge density release for each type of piezoid are given in line three of Fig. 2. It is noted that the dynamic values are less than the static values in all cases and particularly in the case of lead metaniobate, the ratio of dynamic to static charge density release is small.

	Group I* Barium Titanate	Group II† Barium Titanate	Lead Titanate- Lead Zirconate	Lead Metaniobate	
Curie Point	59	64	251	—	Charge Density Release Microcoulomb per inch ²
Static	58	64	166	91	
Dynamic	44	43	150	20	
Theoretical Energy Density	6.3	7.9	83	>45	Joul per inch ³
Dielectric Constant	1225	1225	1675	400	Dimensionless
Curie Temperature	120	120	>350	570	Degrees Centigrade

* Over five years old.
† Less than six months old.

Fig. 2—Characteristics for several piezoids.

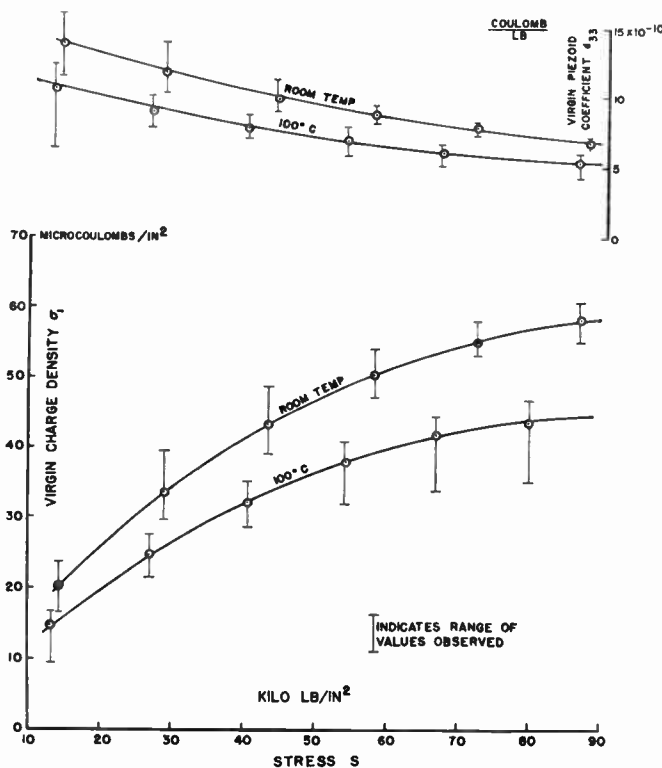


Fig. 3—Charge density release with stress for group I barium titanate piezoids at room and elevated temperatures.

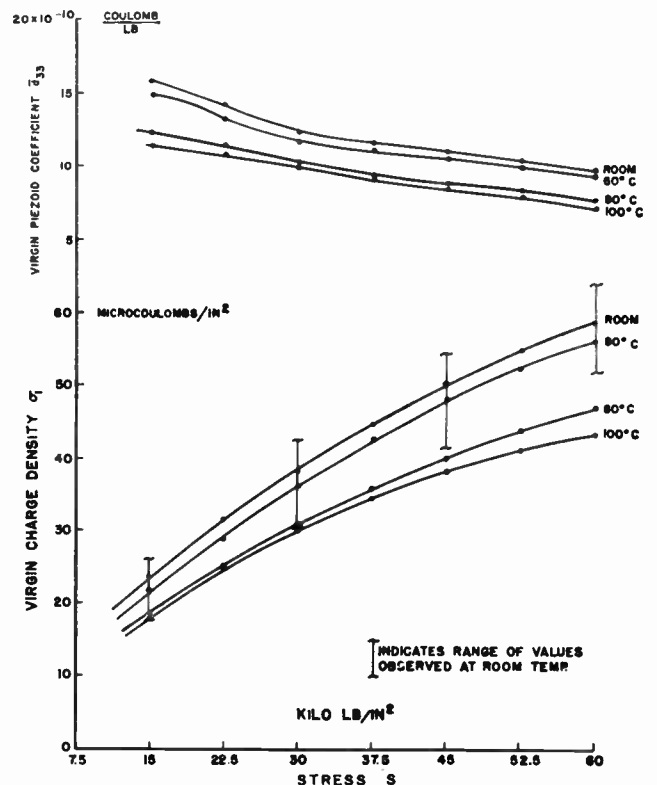


Fig. 4—Charge density release with stress for group II barium titanate piezoids at room and elevated temperatures.

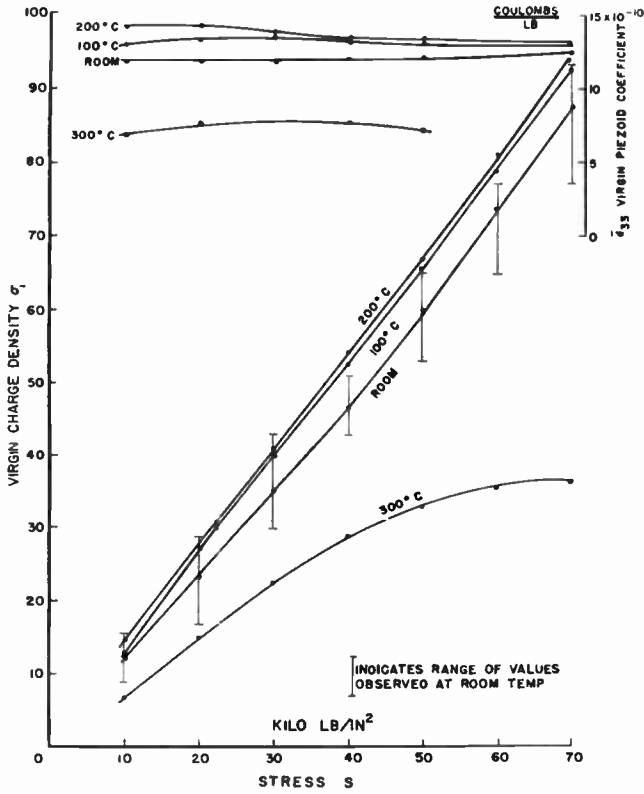


Fig. 5—Charge density release with stress for lead metaniobate piezoids at room and elevated temperatures.

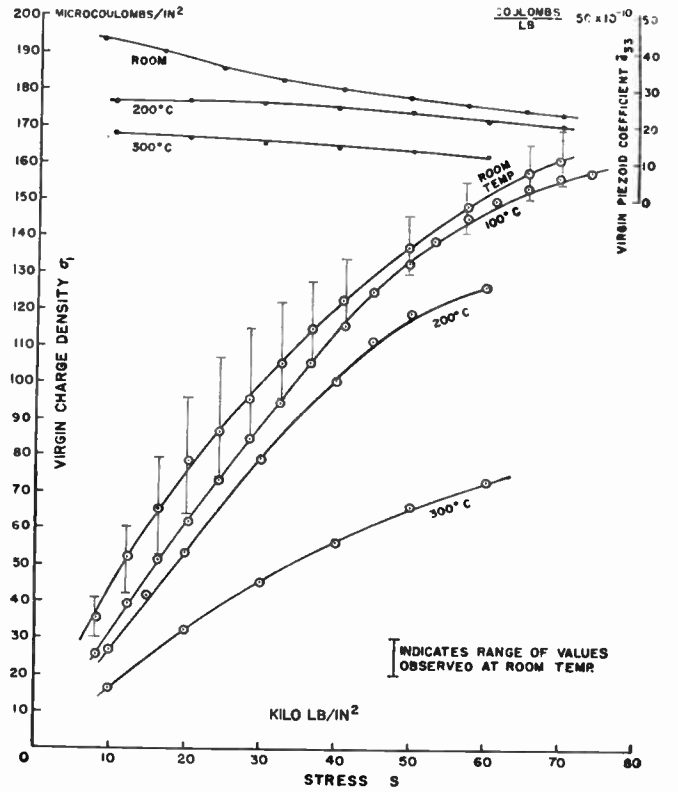


Fig. 6—Charge density release of lead titanate-lead zirconate piezoids at room and elevated temperatures.

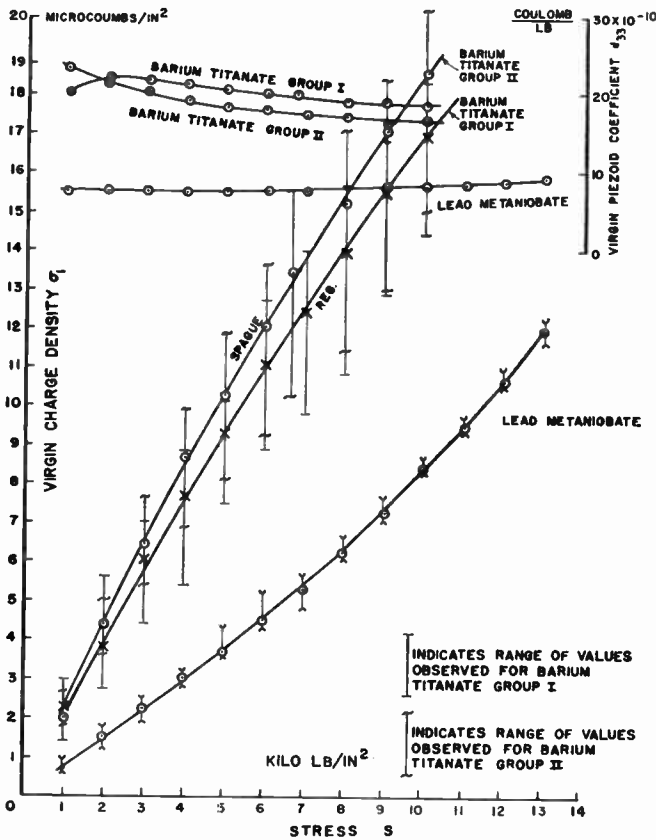


Fig. 7—Room temperature values of charge density release vs low stress for several types of piezoids.

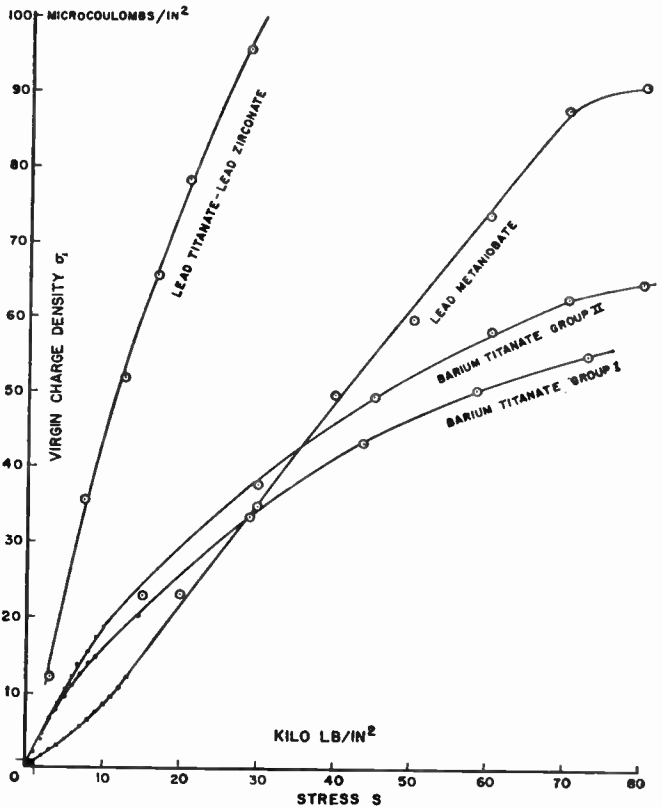


Fig. 8—Room temperature values of charge density release with stress for several types of piezoids.

ENERGY DENSITY

The energy density W_d which may be considered to reside in the volume V of a dielectric is

$$W_d = W/V = 1/2(\epsilon E^2),$$

but the electric field $E = \sigma_1/\epsilon$ where σ_1 is the free charge density on the electrode, thus giving $W_d = \frac{1}{2}(\sigma_1/\epsilon_0 K)$, where ϵ_0 is the permittivity of free space and K is the dielectric constant of the piezoid.

If all the bound charge density of a piezoid could be realized as free charge, then the energy density could theoretically assume the values as shown in line four of Fig. 2. It is realized that these values cannot be attained in practice; however, they do indicate the upper limits of energy densities that may be expected. The representative value of dielectric constant K for the piezoids tested is given in line five of Fig. 2.

CURIE TEMPERATURE

Each particular piezoid type is characterized by a Curie temperature above which it becomes nonferroelectric. It is seen in Figs. 3-6, however, that for practical purposes the magnitude of charge release relative to the values measured at room temperature drops significantly as the Curie temperature is approached. The Curie point for the various piezoids is indicated in line six of Fig. 2.

MATERIALS

It is not uncommon to find a broad range of characteristics for "similar" materials which results from one or more factors, such as firing process, additives, poling technique, aging method, and binders, to mention a few. Some piezoids investigated recently point out how such material factors produce different piezoid charac-

teristics. Since only a small number of these were investigated the results were not reported in the body of the paper. For example, samples of several lead titanate-lead zirconate "type" piezoids gave charge release values of over 200 microcoulombs per square inch for static pressures of 80,000 psi while some "doped" barium titanate piezoids had a Curie temperature of about 140°C and a charge density of 90 microcoulombs per inch² for Curie temperature charge release.

In view of the foregoing, the data presented in body of this paper should be regarded as an indication of the trend and variability of the characteristics for several commercially available piezoids. Material sources were Sprague BaTiO₃, Clevite Pb(Ti Zn)O₃, type PZT-SA, General Electric Pb(NbO₃)₂, and Solar SS-9, doped BaTiO₃.

It is hoped that manufacturers will be more conscious of providing large signal properties of their "modified" types of materials or any newly developed piezoids as well as the standard small signal parameters normally offered.

ACKNOWLEDGMENT

The author acknowledges gratefully the advice and assistance received from K. D. George, A. Graf, and G. Demitrack of the Feltman Research Engineering Laboratories, Picatinny Arsenal, and Professor W. Sullivan of Stevens Institute of Technology.

BIBLIOGRAPHY

- [1] Glenco Corp., "Investigation of Materials for Electrical Energy Sources," Tech. Reps. 1954-55, Contract ORDANCE (P)-1059.
- [2] W. P. Mason, "Energy conversion in the solid state," *Proc. Seminar on Advanced Energy Sources and Conversion Techniques*; November 3-7, 1958 (to be published).
- [3] W. J. Merz, "Spontaneous polarization of BaTiO₃ single domain crystals," and "Electrical and optical behavior of BaTiO₃," *Phys. Rev.*, vol. 76, pp. 459, 1221; 1949.

Some Special and Unusual Applications of Barium Titanate in the Underwater Sound Hydro-Engineering and Medical Fields*

J. D. WALLACE† AND J. R. BROWN, JR.†

Summary—The advent of ferroelectric materials, especially in the ceramic state, has worked a veritable revolution in the design and production of many equipments. The impact on certain military arts has been striking; the field of accelerometers has profited well as has the sonar art. The experience of the Naval Air Undersea Warfare effort is outlined as well as several developments of the National Bureau of Standards. The instances cited are examples of widespread military applications. The development of microminiature hydrophones and their use in cardiac diagnosis is also described.

HISTORY

IN 1917, Rochelle salt was noted to have certain anomalies in its dielectric behavior,¹ most notable of which was the existence of a ferroelectric hysteresis loop. It was also noticed to have a rapid change in its piezoelectric activity at a given temperature. This established Rochelle salt as a material with peculiar properties and constituted the establishment of the ferroelectric class of materials.

In 1935, potassium dihydrogen phosphate and later ammonium dihydrogen phosphate were discovered to have a similar dielectric behavior, that is, ferroelectric hysteresis loops and again these crystals were piezoelectric.² In fact, the main piezoelectric requirements up to and including World War II were borne by these two crystals, with special requirements being met by tourmaline and quartz. Quartz was used almost exclusively in the frequency control field.

Shortly before World War II, anomalies were noticed in the dielectric characteristics of an extremely high dielectric constant material, barium titanate.³ Considerable study of both crystalline and polycrystalline barium titanate indicated the existence of ferroelectric hysteresis loops over a fairly wide temperature range.⁴ Shortly after World War II, researchers in both this country and Russia noticed that the piezoelectric effect could be induced in polycrystalline barium titanate ceramic by external polarization.

* Original manuscript received by the IRE, December 1, 1958.

† U. S. Naval Air Dev. Center, Johnsville, Pa.

¹ J. Valasek, "Piezo-electric and allied phenomena in Rochelle salt," *Phys. Rev.*, vol. 17, pp. 475-481; 1921.

² G. Busch and P. Scherrer, "Eine neue seignette-elektrische Substanz," *Naturwiss.*, vol. 23, p. 737; October 25, 1935.

³ E. Wainer and A. N. Salomon, Titanium Alloy Manufacturing Co., Elec. Rep. No. 8, 1942; Elec. Rep. 9 and 10, 1943.

⁴ A. von Hippel and co-workers, "High Dielectric Ceramics," Lab. for Insulation Res., Mass. Inst. Tech., Cambridge, Mass., Rep. No. 500; August, 1944.

— "Titania Ceramics II," Lab. for Insulation Res., Mass. Inst. Tech., Cambridge, Mass., Rep. No. 540; October, 1945.

CURRENT USAGE

Sonar

The history of modern naval sonar transducer technology can be traced in the developmental steps from the primitive listening tubes to the modern ferroelectric ceramic transducer.

With the advent of the use of Rochelle salt, the first discovered ferroelectric material, the refined art of sonar transduction was launched. Rochelle salt proved the mainstay through the primary developments of sonars into World War II and served well in early undersea warfare. However, during World War II, Rochelle salt gave way to another ferroelectric crystal, ammonium dihydrogen phosphate. ADP together with magnetostrictive nickel were the principal transducing materials. After the war, magnetostrictive materials were augmented by the introduction of the cobalt alloys and somewhat later by the introduction of relatively hard magnetostrictive ferrites.

Although barium titanate had been noted for its peculiar characteristics, especially its transducing properties, it was not until 1948 that this material, in ceramic form, was used in the sonar field. Fig. 1 shows the hydrophone which has been used in all the nondirectional AN/SSQ-2 radio sonobuoys since its origin in 1948. In the sonobuoy application, barium titanate was indeed fortunate in its initial usage, for here productivity, manufacturability and cost were weighed equally with performance. The AN/SSQ-2 sonobuoy with its three-tube barium titanate hydrophone is now a familiar sight throughout the free world.

Fig. 2 is a four-element directional array which was designed to replace a larger and more expensive crystal unit. This hydrophone array is made up of four single tubes of barium titanate which are about 6 inches long and $\frac{3}{4}$ inch in outside diameter.

Fig. 3 shows a tubular element which has been designed as a replacement for a packet of crystals or a magnetostrictive element. The element is simply a barium titanate tube with loading on either end; the piston type radiation is taken off the upper end. While the elements illustrated utilize brass pieces for the loading, units have been designed with other materials as masses. The unit vibrates in half wavelength resonance with a node at approximately the midpoint of the tube.

Fig. 4 illustrates a cavity type directional transducer



Fig. 1—AN/SSQ-2 sonobuoy with hydrophone.

which is made up of a group of barium titanate tubes. In this case, the radiation is taken off the inside of the tube and then directed out the mouth of the tube.⁵

Hydraulic Pump and Servo

There is a possibility of providing a hydraulic servo system based upon the piezoelectric effect. Such a system would utilize a piezoelectric pump and fast-acting valves. To this end, a piezoelectric pump⁶ has been under development (Fig. 5). The pump shown is a single-acting pump, *i.e.*, half cycle. The operation results from expanding the radius of the inner tube while contracting the radius of the outer tube. This action reduces the volume between the tubes, and when we alternately ex-

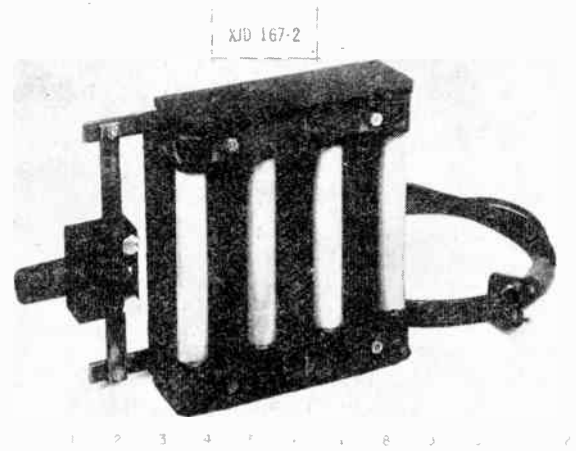


Fig. 2—XJD-167-2 four-element array.

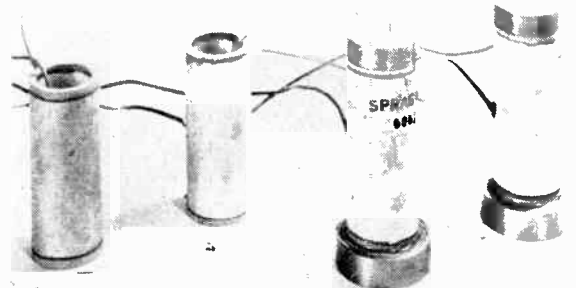


Fig. 3—Tubular sonar transducer elements.

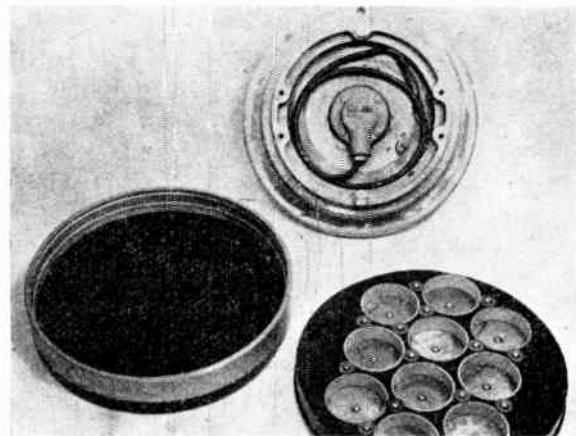


Fig. 4—Cavity type transducer.

pand the outer and contract the inner tube, we increase the volume—and have a pump. An experimental model of this pump has been made and has been relatively successful. Useful pumping operation depends upon valves. Two types of valves (Fig. 6) have been developed. One utilizes a tube of barium titanate operating in the longitudinal direction driving a mechanical transformer, which is the solid horn-like appendage. While this type has been fairly successful, it is rather large. The pump illustrated operated at about 15,000 cps; longitudinal

⁵ Produced by Texas Instruments, Inc., Dallas, Tex.

⁶ Originally suggested by C.L. Stec of the U. S. Navy Bur. of Ships.

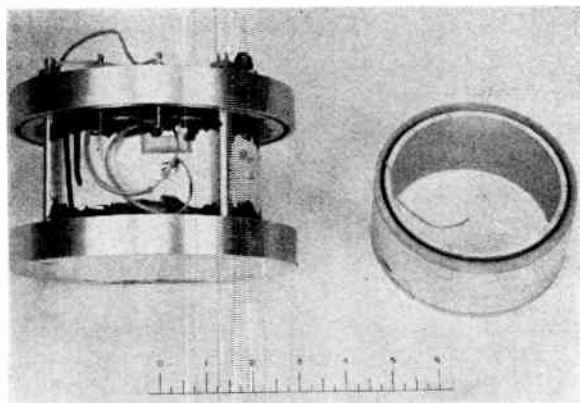


Fig. 5—Ceramic hydraulic pump.

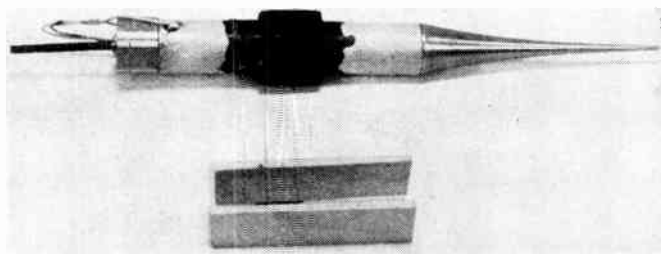


Fig. 6—Pump valves.

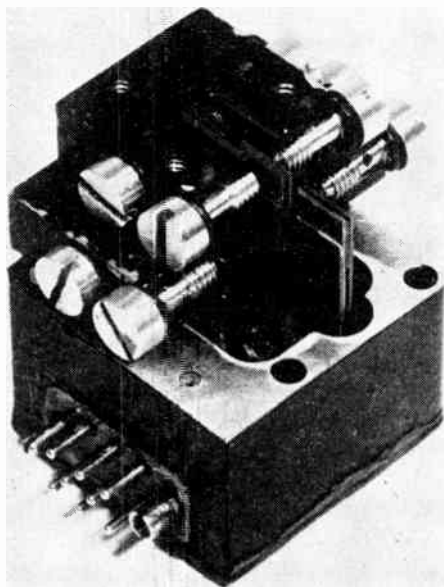


Fig. 7—Control valves.

pumping valves operating at that frequency turned out to be about 9 inches long. To alleviate this difficulty, a piezoelectric ceramic tuning fork valve has been developed. This is an element which operates as a driven tuning fork. A 15,000-cps unit of this type is between $\frac{1}{2}$ and 1 inch long. Both of the valves have been used with the pump and pressures of the order of 50 psi at flows of about 100 cubic inches per minute have been achieved. The heart of any hydraulic servo system is, of course,

the control valve. Fig. 7 illustrates the piezoelectric control valve developed by the Dalmo Victor Co. for the hydraulic servo. It embodies a vane type of piezoelectric element which varies the flow from two orifices. A servo system has been built which has a maximum servo error of about $1\frac{1}{2}$ minutes. The unit has no perceptible or measurable hunt.

Accelerometers

Inertial guidance systems used in our large missiles are dependent on correction signals from accelerometers; that is, the accelerometer signals indicate deviation from the programmed path.

Fig. 8 indicates the general type of accelerometers developed and made by the National Bureau of Standards.⁷ Accelerometers are also used in connection with vibrating machines and provide the usual instrumentation in system vibration studies. Accelerometers have to be calibrated; such is accomplished using pendula, loudspeaker shakers, etc. Fig. 9 shows several barium titanate calibration shakers recently developed by the National Bureau of Standards.⁸ In these units, the barium titanate tubes or discs vibrate and drive or shake the accelerometers under calibration; such units provide calibrations in excess of 10,000 cps.

Medical Instrument

Out of the military work of the Naval Air Dev. Center at Johnsville, Pa., has come a rather unique development—the intracardiac phonocatheter.⁹ This is a small underwater or underblood transducer which is inserted into a vein or artery and threaded directly into the various chambers of the heart. At this writing the procedure has been done on several hundred human patients with very encouraging results.¹⁰ The development work has been done in conjunction with the Philadelphia General Hospital Div. of Cardiology. The technique is now being introduced into hospitals throughout this country and Europe. Fig. 10 illustrates the three types of phonocatheters produced to date. The uppermost catheter was the first ever made and has an outside diameter of 105 thousandths of an inch. The $\frac{1}{2}$ inch barium titanate tube at the tip was 80 thousandths OD and 40 thousandths ID. The middle unit is the double lumen catheter. This 0.110 inch OD catheter has two bores. One is simply a hole from which blood can be sampled to measure oxygen content or pressure. The other bore

⁷ T. A. Perls and C. W. Kissinger, "A Barium Titanate Accelerometer with Wide Frequency and Acceleration Ranges," Natl. Bur. of Standards Rep. 2390; April, 1953.

⁸ S. Edelman, E. Jones, E. R. Smith, B. D. Simmons, and R. C. Braunberg, "Self-Noise of Projectiles," Natl. Bur. of Standards Rep. 3465; July 23, 1954.

⁹ J. D. Wallace, J. R. Brown, Jr., D. H. Lewis, and D. H. Dietz, "Acoustic mapping within the heart," *J. Acoust. Soc. Amer.*, vol. 29, pp. 9-15; January, 1957.

¹⁰ D. H. Lewis, G. W. Deitz, J. D. Wallace, and J. R. Brown, Jr., "Intracardiac phonocardiography in man," *Circulation*, vol. 16, pp. 764-775; November, 1957.

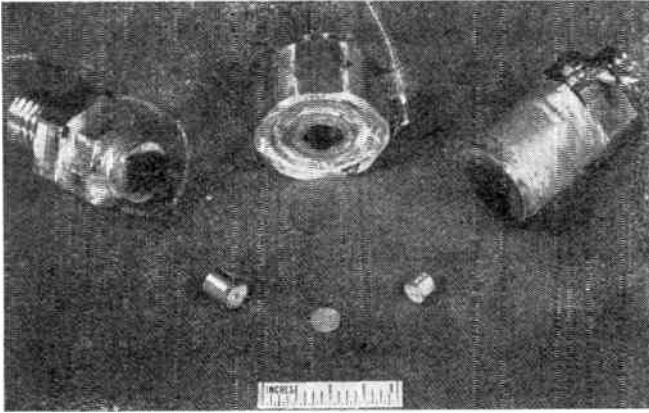


Fig. 8—NBS accelerometers.

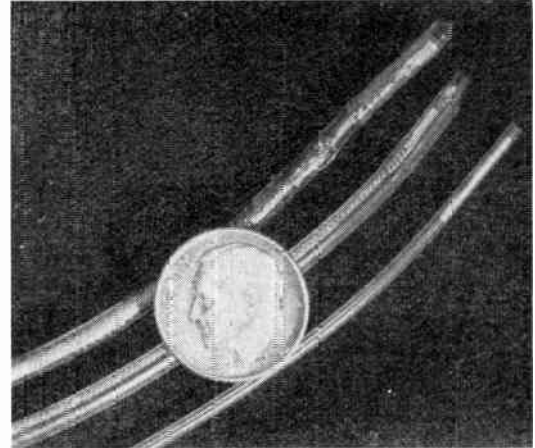


Fig. 10—Intracardiac phonocatheters.

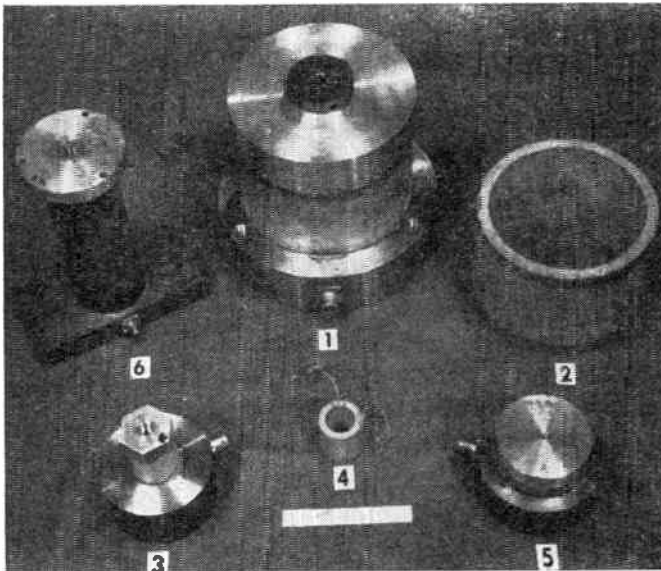


Fig. 9—NBS ceramic shakers.

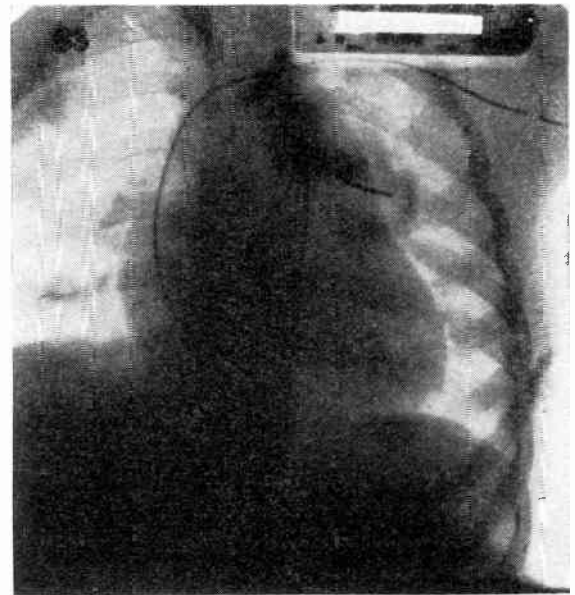


Fig. 11—X ray of intracardiac phonocatheter in cardiac system of 4-year old girl.

contains the phonocatheter; in this case the $\frac{1}{2}$ inch ceramic tube is 40 thousandths OD and 20 thousandths ID. The bottom catheter is the type currently in use in babies and children. The single lumen catheter has an OD of 60 thousandths using the same piezoelectric ceramic tube as the double lumen. The children's catheter has been inserted into living hearts of as small as a 5-pound, 11-week-old premature baby. Fig. 11 is an X ray taken of a catheter in the heart system of a 4-year-old girl. Note the catheter which has been threaded through a vein in the left arm, through the right atrium, through the valve and right ventricle, through the pulmonic valve and out into the pulmonary artery where the phone tip lies at the base of the left lung lobe. The sounds picked up by the phonocatheter have served in many cases to assure the diagnosis and in several provided the basic diagnosis. In this instance, the defense dollar and effort have served the betterment of mankind.

CONCLUSIONS

The foregoing has indicated the usefulness and adaptability of ferroelectric materials as applied to military problems. The breadth of application across diverse fields has served to enhance military capabilities to a great degree. The medical application has future clinical significance since the devices described are being readied as the physiological instrumentation for the biological satellites.

ACKNOWLEDGMENT

The authors are pleased to acknowledge the assistance of M. F. Pressler for his basic contribution of the ceramics, and the contributions of D. J. Repici, A. B. Alexander, and W. A. Bertles.

Some Representative Electronic Subsystem Development Problems in Naval Ordnance*

E. H. BEACH†, MEMBER, IRE, M. J. PARKER†, SENIOR MEMBER, IRE, P. YAFFEE†, SENIOR MEMBER, IRE, AND R. B. KNOWLES†, MEMBER, IRE

Summary—This paper discusses the problems facing the electronic research and development engineer in designing modern naval ordnance equipment. Four distinct areas are presented and are illustrated by the following particular developments: underwater ordnance, missile carried ordnance equipments, equipment intended for field system evaluation and training, and test sets for subsystem testing. Technical details are given on the meteorological rocket HASP, the Miss Distance Measurement System, AN/USQ-11, and applications of the recently developed Solion.

I. INTRODUCTION

NAVAL ordnance requirements for Weapon System performance have traditionally been such as to challenge the most ingenious designer. The prime motive for excellence is not for profit, as in commercial design, but for national survival itself. The electronic research and development engineer working on naval weapon systems is constantly searching for new components, new techniques, and new concepts to meet the difficult technical and operational objectives. Each subsystem application is a unique problem, and the generalities of electronic design criteria so earnestly propounded by many technical experts may not necessarily apply. For example, gun-fired proximity fuzes are required to operate with a high degree of reliability for a matter of seconds, while mine firing mechanisms may operate unattended for years. Conventional concepts of conservative circuit designs may be somewhat academic when the available volume and usable components are limiting factors. Conversely, many shipboard ordnance equipments feature ultraconservative design not only to allow for a minimum operation and maintenance staff, but also to allow for more reliable operation under transient battle extremes. Subsystem manufacturing cost is important but may be a fourth order consideration in some applications, such as fuzing nuclear weapons where safety and reliability are paramount.

In this paper, four distinct types of naval electronic subsystem developments are presented. The first subsystem deals with the design problems of underwater influence mechanisms, the second subsystem illustrates an example of the gun and missile mechanism design area, the third subsystem presents the design of a missile system evaluation and operational training device, and the last subsystem outlines the problems of designing rugged, simple, field test sets with which the fleet can keep its electronic equipment in operational readiness.

* Original manuscript received by the IRE, December, 3, 1958; revised manuscript received March 2, 1959.

† U. S. Naval Ordnance Lab., Silver Spring, Md.

It is hoped that these discussions and examples, necessarily presented within the boundaries of tight security regulations, will further the appreciation of the problems involved in the development of modern naval ordnance electronics.

II. UNDERWATER INFLUENCE MECHANISMS¹

Underwater influence mechanisms are the intelligence devices or "brains" used in underwater weapons such as mines, torpedoes, and other antisubmarine weapons to detect the presence of an enemy ship or submarine by means of the target's influence (*e.g.*, acoustic, magnetic, pressure, etc.) to detonate a destructive warhead or charge if the influence is legitimate, and to resist responding if the influence field is generated by countermeasures or by natural earth perturbations. A very simple influence mechanism would be an induction coil or search coil connected to a highly sensitive relay. The presence of a ship distorts the ambient magnetic field of the earth; the time rate of change of this field generates a voltage to which the relay responds. Needless to say, such a simple device is not sufficiently intelligent for modern warfare.

The underwater weapons used by the Navy today are highly engineered devices designed to meet a very severe set of operational requirements and to function under the most unusual environmental conditions. For the design engineer and scientist, these weapons provide a most interesting and stimulating challenge.

Design Goals and Environmental Factors

A qualitative understanding of the design goals and environmental factors for underwater weapons may be obtained from the following general statements.

1) These weapons are often launched by the most modern aircraft and submarines. The safety of these weapons must be commensurate with the cost of such launching vehicles together with their trained crews—multimillion dollar investments in many cases.

2) The weapon performance must be highly reliable over the expected life of the weapon. For certain weapons such as mines, life may be measured in terms of months or years, and satisfactory performance is desired without servicing or other maintenance.

3) Weapon response must be such that it destroys or damages the intended or desired targets but resists

¹ This section prepared by E. H. Beach.

insofar as possible false man-made or perturbation stimuli.

4) The weapons will, in general, be transported without the benefit of special packaging other than crates or cartons by almost all modern transportation conveyances and subjected to the various modes of vibration which may result; the launching of these weapons may be from submarine tubes at some depth, as well as from the highest altitudes of modern aircraft, the latter sometimes being accomplished without the benefit of retardation devices. Thus, the shock handling capability of these weapons must be very great.

5) These weapons must be capable of withstanding the temperature extremes of both the desert sun and the arctic wastes.

6) The final environment of all of these weapons is the water. It is to be expected, therefore, that very often the humidity inside these weapons may approach high levels unless special preventive precautions are taken. It should be appreciated that high humidity can cause serious problems for moving mechanical parts, electrical contacts, and high-impedance electronic circuits.

The Importance of the Influence Mechanism

In almost all cases, the influence firing mechanism is the weapon component which provides the target discerning and discriminating intelligence, which is the principal power consumer and thus the limiter on life, which because of complexity, limits the weapon reliability, and which is most prone of all components to damage by shock, vibration, temperature, and humidity extremes.

In general, a superior weapon requires an influence mechanism of superior design. Superior influence mechanisms are designed by using components which require less power, take less space, eliminate high-voltage power sources, take more shock and vibration, and take greater temperature and humidity extremes. Transistors, Solions, magnetic amplifiers, ring modulators, Zener-diode regulators, and magnetic memory cores constitute the component family which is rapidly replacing filamentary tubes, cold cathode gas tubes, meter relays, telephone-type relays, motor driven switches and escapement-type timers. The Solion is one of the newer, more interesting components which will be described in greater detail.

The Solion

The Solion^{2,3} is an electrochemical family of devices developed by and for the Navy. The Solion principle may be used in a number of different ways to produce pressure or heat sensing detectors, integrators, deriva-

tive calculators, and product taking devices. Fig. 1 depicts a very simple Solion pressure detector.

The cell in Fig. 1 is equipped with two platinum anodes and a platinum orifice-type cathode. The electrolyte is composed of a high concentration of potassium iodide (KI) and a low concentration of iodine (I₂). The reduction of iodine at the cathode is the controlling reaction due to the diluteness of the iodine in the solution and the rate at which it diffuses to the cathode. The following electrode reactions occur [(I₃)⁻ is equivalent to writing (I₂+I⁻)]:

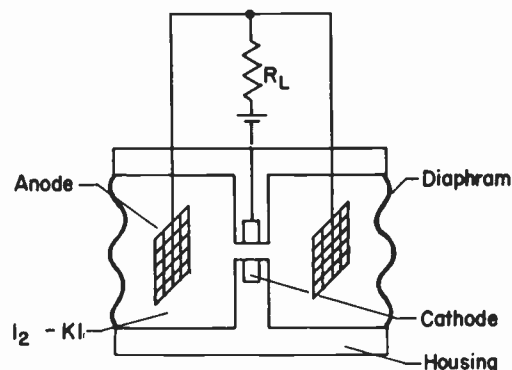
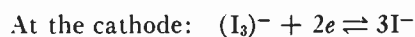


Fig. 1—Solion pressure detector.

With the cell on bias as shown in Fig. 1, the iodine in the vicinity of the cathode becomes reduced, and the background current level that is established is that due only to the diffusion of iodine from the anode regions to the cathode area. Fig. 2 shows a voltage-current plot for several different values of triiodide (*i.e.*, iodine) concentration C for quiescent conditions with the diaphragms undisturbed. Note that for voltages approximately between 0.2 and 1.0 v a plateau region exists in which the current is essentially voltage independent and is limited by diffusion effects. However, when a pressure differential is established across the two diaphragms, hydraulic flow of fluid rich in iodine is produced through the cathode, and the current level rises, being no longer diffusion limited. The relationship between the acoustical pressure applied to the diaphragms and the current flow may be shown to have the form, $I = kP^n$ or $I = k \log P$, depending upon the cathode geometry. When a low-frequency sinusoidal drive is impressed on the cell diaphragms for a detector of the linear type ($n = 1.0$), the current-voltage characteristics for different drives appear as in Fig. 3 and the current-drive curve is like that shown in Fig. 4.

For low-frequency applications, Solions offer many interesting possibilities. These devices are attractive because of their low-power requirements, because of their operation from low-voltage power supplies, and because of the ease with which they may be matched to transistor circuitry.

² R. M. Hurd and R. N. Lane, "Principles of very low power electrochemical control devices," *J. Electrochem. Soc.*, vol. 104; December, 1957.

³ R. H. Reed, *et al.*, "The Solion," *Yale Sci. Mag.*, vol. 32; February, 1958.

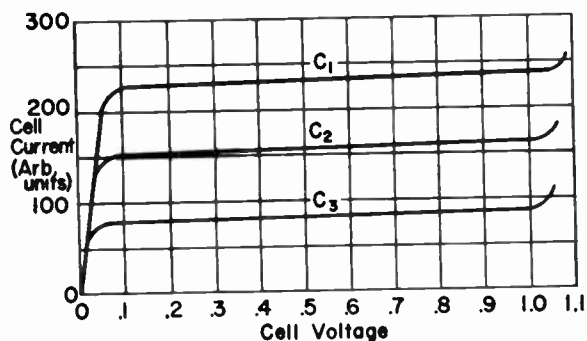


Fig. 2—Current-voltage plot with concentration as parameter.

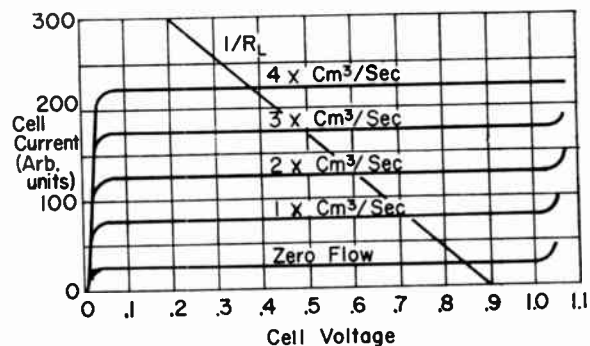


Fig. 3—Current-drive curves for varying sinusoidal drives.

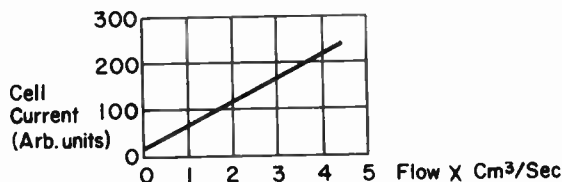


Fig. 4—Current-drive curve.

Solion Battery

In order for the Solion described to be a Redox system, that is, a reversible system in which gassing does not occur, it is necessary that the biasing potential be kept below the proper level (for the iodine-iodide cell below 1.0 v). Since Leclanché-type batteries have potentials in excess of 1.0 v a new battery system having a potential of 0.9 v has been developed.⁴ This battery is of the chargeable type using silver and lead-oxide electrodes with potassium hydroxide electrolyte. One of the principal advantages of this battery is long storage life (uncharged) of up to ten years. Fig. 5 gives a cross-sectioned sketch of the cell. Upon charging at 1.12 ± 0.03 v, the silver is converted to silver oxide (Ag_2O) and the lead oxide to metallic lead.

III. ROCKET-SONDE METEOROLOGICAL EQUIPMENT⁵

The Navy requires operational equipment in order to make regular meteorological observations to alti-

tudes in excess of 100,000 feet. At present, balloons are used for synoptic observations up to 70,000 feet. In order to reach higher altitudes with any degree of reliability larger balloons are necessary. The operational problems imposed by the use of balloons, and the aggravation of these problems if larger balloons are used, has led to the consideration of other means for conveying instruments aloft. Aside from its expense, a rocket is an almost ideal vehicle for this purpose. The Army T-2006 Loki, a small rocket originally developed as an anti-aircraft weapon, has required only slight modifications for launching from Navy semi-automatic 5-inch guns. This meteorological rocket is known as HASP (High Altitude Sounding Projectile). The HASP is easily handled and its cost is moderate enough so that it holds promise of being usable on a routine basis. Actually there are several members of the HASP family. The first, which is in use, is a chaff-loaded 5-inch projectile. Radar tracking of the chaff, ejected at altitudes to 50,000 feet results in wind velocity and wind direction determination. The second is the rocket propelled HASP which carries chaff above 100,000 feet, and the third, a rocket, carries a reflecting parachute above 100,000 feet. The fourth, also a rocket, is the instrumented round.

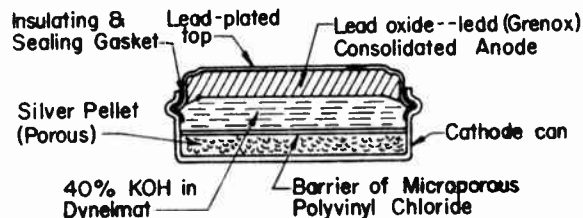


Fig. 5—Cell cross-section.

The Loki is a single stage solid propellant rocket. With the 3-inch diameter Phase I Booster, burnout occurs at 0.8 second. The motor then separates from the smaller forward portion of the assembly, called the dart, or head, which coasts up to the top of its trajectory, somewhat higher than 100,000 feet. The dart is 40 inches long, $1\frac{3}{8}$ inches outside diameter, exclusive of the fins, and weighs 6.3 pounds. As modified for meteorological purposes, the dart will carry instruments to near the top of its trajectory where they will be ejected. The instruments on their descent sense temperature, and in latter versions pressure, density, and humidity. These sensors are combined with telemetering equipment into an assembly, hereafter referred to as the payload, which also includes a parachute or balloon for slowing the fall. Radar tracking of the metallized parachute or balloon permits determination of wind velocity, wind direction, and altitude of the payload.

The first instrumented HASP to be produced will have 100,000 feet altitude capability and will measure temperature only, in that a suitable pressure, density,

⁴ H. B. Reed and M. B. Goldberg, "A new rechargeable dry cell," Extended Abstracts of the Battery Div. of the Electrochemical Soc., Fall Meeting, p. 33; October 7-10, 1957.

⁵ This section prepared by M. J. Parker.

or humidity sensor has not been developed. This HASP is referred to as the "telemetry only" system. The transponder system is discussed later. The altitude, wind velocity, and wind direction will be determined by tracking the metallized parachute by radar. The temperature sensor is a small bead thermistor with a bead diameter of 0.040 inch and 0.004 inch diameter leads. A survey of work performed by various activities, supported by laboratory research performed at the Naval Ordnance Laboratory indicates that this thermistor, if coated with a suitable reflective material such as lead carbonate or aluminum, will measure the true air temperature within $\pm 1^\circ\text{F}$ up to about the 100,000-foot altitude. A somewhat smaller bead thermistor should be satisfactory up to 150,000 feet.

The circuit is similar to that used in balloon sondes such as the AMT/11, but is miniaturized to fit in the cylindrical space measuring $1\frac{1}{4}$ inch diameter and 10 inches long, which is available in the HASP head. Such environmental factors as 300 g acceleration, ascent temperatures of the dart to 700°F , descent temperatures of the payload to -70°F , in addition to military standard requirements, complicate the design.

Transistors are used instead of tubes, except for the 403 mc/sec oscillator which uses a subminiature triode. The power supply is provided by a 6 v sealed nickel-cadmium battery which measures 1 inch in diameter by $2\frac{1}{2}$ inches long.

The circuit is designed for use in conjunction with the shipboard AN/SMQ Radiosonde Receptor, presently used with balloon borne meteorological sondes. This means that no additional shipboard receiving equipment will be necessary for HASP. The AN/SMQ consists of a 403 mc/sec receiver (tunable from 390 to 410 mc/sec, this frequency is assigned for meteorological use), a "pulse-integrating" frequency meter, and a slide wire self-balancing type recorder. The receiver bandwidth is optimized for 200 μsec pulses, and the frequency meter and recorder are designed to detect and record pulse repetition rates ranging from 10 to 200 pulses per second.

The resistance of the temperature sensor controls the frequency of a blocking oscillator. For calibration purposes, a fixed precision resistor is switched periodically into the blocking oscillator circuit in place of the thermistor by an electronic switch. The pulses generated by the blocking oscillator are amplified and applied via a step-up pulse transformer to plate modulate the transmitting oscillator.

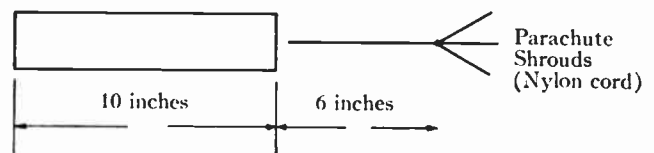
Referring to the circuit diagram, Fig. 6, the transistors TR1 and TR2, and associated components, comprise a multivibrator which performs the switching function. The diode D1 (or D2) effectively isolates the feedback capacitor C1 (or C2) from the collector of the associated transistor TR1 (or TR2) during the "off" period, thus allowing the collector to go almost to ground potential. The collectors of transistors TR1 and TR2 switch alternately from ground potential to

almost +4 v. The potential at the anodes of diodes D3 and D4 always remains within these limits, so that when the collector of TR1 is at ground potential, and the collector of TR2 is approximately +4 v, D3 will be conducting and D4 will be nonconducting. The thermistor R7 will, therefore, control the frequency of the blocking oscillator during this period. During the other half of the blocking oscillator cycle, the situation is reversed, and the calibrating resistor R8 is connected into the blocking oscillator circuit.

The blocking oscillator consists of transistor TR3 and associated components and is designed to develop a pulse of about 200 μsec duration. The diode D5 prevents reverse conduction or leakage current from flowing through the base of the transistor during the "off" period. This current is nonlinear and varies with temperature, and would affect the rate of discharge of the condenser during the "off" or timing period.

The period of the blocking oscillator is a linear function of the resistance of the sensor. Temperature effects and variation of the supply voltage merely cause slight changes in the slope of the period vs resistance curve which always passes through the origin. By comparing the frequencies produced by the thermistor resistance and the calibrating resistor, the temperature of the thermistor, which is the temperature of the atmosphere, can be accurately determined.

The pulses generated by the blocking oscillator are amplified by the buffer stage TR4, power amplifier TR5, and applied to the plate of the transmitter tube via the 1:25 step-up transformer. Laboratory tests show that a peak power output of 2 w can be obtained from the transmitter at a plate efficiency of about 30 per cent. The power output under actual pulse operation has not yet been measured. The antenna system consists of the container of the instrument package which has an equivalent electrical length of approximately $\lambda/2$, and a length of wire approximately $\lambda/4$ long, which also will act as the supporting cable, being attached at the upper end to the parachute shrouds. The antenna pattern is approximately that of a $\lambda/2$ dipole, but leans



downward slightly. The shipboard receiving antenna is a $\lambda/4$ vertical antenna mounted on a circular "ground plane" 1 wavelength in diameter. This gives an upward leaning pattern which combines with the transmitting antenna pattern to give good signal transfer up to elevation angles of about 65° - 70° .

One of the limitations of the meteorological system using the instrumented payload described above, is the requirement for the radar to acquire the 7-foot metallized parachute at 100,000 feet altitude, and then to track this target, as it follows the winds, to possibly

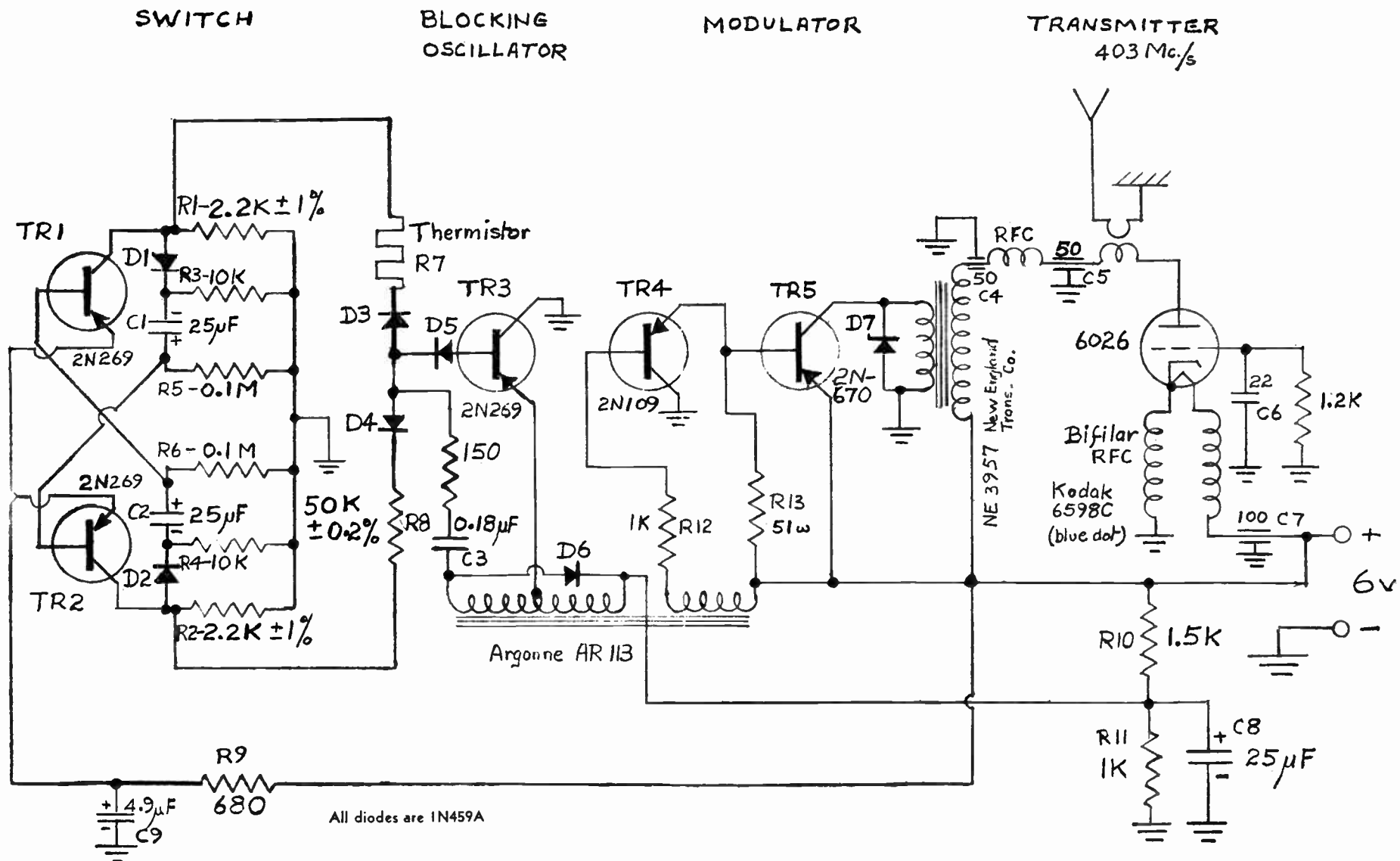


Fig. 6—Circuit diagram of "telemetering only"—HASP airborne instrumentation.

200,000 feet. In order to overcome this requirement and to extend the altitude capability of the HASP to 200,000 feet, a transponder version of HASP has been breadboarded. This will operate as follows.

A transponder in the payload receives the signals from an X-band radar and repeats them on a 403-mc carrier. The repeated pulses are of constant height but each has a length between 1 and 10 μ sec proportional to the received strength of the corresponding X-band pulse. In the shipboard radar, the beam is slightly inclined with respect to the antenna axis and rotates about the axis at 30 rps. Thus, if the payload is close to but not exactly on the axis of the antenna, the envelope of the received pulse heights is amplitude modulated at 30 cps, and this modulation appears as pulse width modulation in the repeated pulses. This modulation is illustrated by the pulses labelled *R* in Fig. 7. From the amplitude of this modulation and its phase with respect to the rotation of the beam the position of the payload with respect to the axis of the radar antenna can be determined.

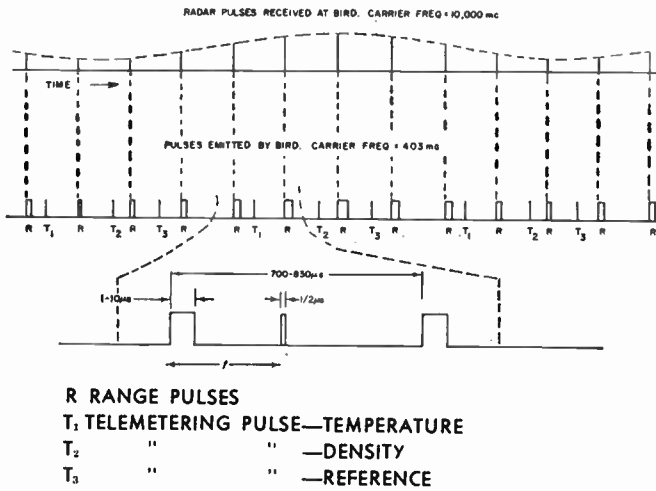


Fig. 7—Modulation of the HASP airborne transponder.

An AGC on the video amplifier within the payload will be required in order that this system of tracking may function over a wide range of distances. The time constants of the AGC must be sufficiently long so that the 30-cps modulation is not suppressed yet sufficiently short to compensate for signal strength variations that may result from changes in altitude of the payload.

The shipboard radar emits between about 1200 and 1400 pulses per second. Since the pulses emitted in reply by the payload at 403 mc will be at most 10 μ sec wide, there is an interval of nearly 700 μ sec associated with each pulse which may be used for telemetering. It is proposed to insert single one-half microsecond pulses into these spaces and to use the position of the pulses to convey temperature and density information. This is illustrated in Fig. 7, where the time interval *t* conveys information from a single temperature measurement.

These telemetering pulses would be divided into groups by periodically omitting one or more pulses. A specific position in each group would then be assigned for telemetering temperature while another position would be assigned to density. Other positions in the group would be used for reference pulses, by which effects due to gradual changes in the components of the telemetering system could be eliminated.

A block diagram illustrating the proposed electronic system in the payload is shown in Fig. 8. The outputs of the temperature and density sensors and the reference signals are connected through individual gates to the input of the amplifier supplying the bias of the phantatron. The gates are controlled by a stepping circuit so that at any instant only one gate is open. With each step, one gate is closed and another opened so that the various inputs are sequentially connected and each in turn determines the bias applied to the phantatron, and hence, the time delay generated by it. With each output pulse from the phantatron, the stepping circuit is advanced to the next position. The stepping circuit also controls the flow of pulses from the phantatron to the 403-mc transmitter, and for certain of its positions, interrupts the pulses in order to provide synchronizing information to the shipboard equipment.

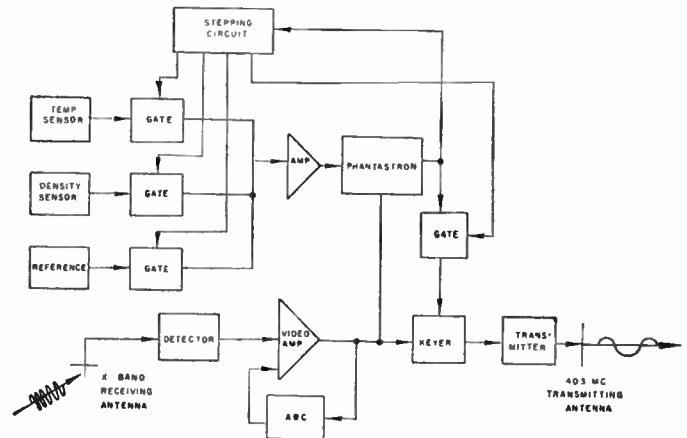


Fig. 8—HASP airborne transponder block diagram.

A block diagram of the shipboard electronic equipment is shown in Fig. 9. The 403-mc pulses from the payload are received by a nondirectional receiver which supplies signals to the radar by which the range, bearing, and elevation of the payload are determined. For this operation, the radar is modified so that the inputs of the IF amplifier and of the train and elevation error detector chassis are disconnected from their normal sources and are connected instead to special output circuits from the 403-mc receiver. The receiver separates the half microsecond telemetering pulses from the pulses conveying range information and then uses the leading edge of the range pulses to generate quarter microsecond pulses on a 59-mc carrier which are applied to the input of the radar's IF amplifier. The receiver

also detects the 30-cps pulse width modulation of the range pulses and applies the resulting signal to the radar's train and elevation error detector. In addition, the receiver includes equipment for separating the several channels of information contained in the telemetering pulses and delivering the signal from each channel to its respective output terminals in a form usable by an indicating device or a data computer.

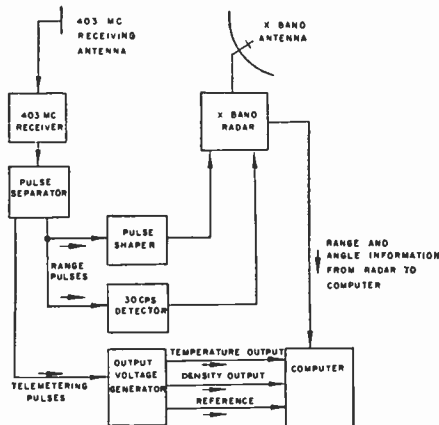


Fig. 9—Block diagram of shipboard equipment for HASP transponder system.

The limiting accuracy, as well as the capacity of this telemetering system, are more than adequate for the present application. The total interval into which the telemetering pulse might be inserted is about 700 μ sec wide. If only half of this is actually available, and if the time interval between the leading edge of the range pulse and the telemetering pulse can be read only to the nearest microsecond, which can be readily achieved, there will be available about 400 graduations. This corresponds to indicating temperatures over a 100° range to the nearest quarter degree, and is also far beyond any meaningful accuracy in the reading of any density equipment likely to be available in the near future. If there are eight range pulses associated with a complete group of telemetering signals, information from each sensor is received more than 150 times per second or more than once for each foot of fall of the payload. The data handling system on the ship should be designed to take advantage of this redundancy. Thus if any interval does not contain the proper number of pulses, the telemetered data obtained during that interval should be automatically rejected. If any group of pulses is not the right length in time then all the data in the group should be automatically rejected. Thus, only valid data will be passed on to the computer. It is possible that a system which remembers the immediately preceding valid data will be required in order to assure a continuous flow of information to the computer during periods when satisfactory data are not being received from the payload.

IV. MISSILE-TARGET MISS MEASUREMENT⁶

Since man began to use modern weapons, the need has existed for means of measuring their effectiveness. The short range weapon, from the thrown stone to the bombardment and anti-aircraft artillery projectile, has presented no great problem since visual observation was an effective way of assessing the result of the action. As weapon development has advanced, however, visual observation has ceased to be feasible. Aircraft have become faster, making detection at longer ranges necessary. High-velocity, long-range missiles have been developed to intercept these aircraft at ranges sufficiently distant to insure survival of the launching ship. As the development of these missile systems has progressed, the need for accurate methods of system evaluation against typical targets under tactical conditions has become acute. This need has long been recognized by all activities concerned with the development and evaluation of missile systems and fire control equipment, as well as those responsible for the proficiency of crews in the use of these systems.

The widespread recognition of the need has been paralleled by a correspondingly widespread application of effort to develop methods for measuring the miss distance, relative velocity, and trajectory during an intercept of a missile and its target. Achievement of a satisfactory solution is complicated in that so many different types of missiles and targets, each with its peculiar operating characteristics, must be considered. Thus, the desired miss distance and relative velocity measurement accuracies, the desired miss distance range, the method and rapidity of data analysis, and the type of target and missile instrumentation, become direct functions of the specific intercept problem to be solved.

In order to detect the intercept between a missile and a target, a method of sensing must be used. The first attempt to extend intercept instrumentation beyond the physical limitations of tow-targets was made during World War II, making use of magnetic sensing. By magnetizing a projectile prior to firing, its passage could be detected by a magnetic detector in the target. Although some effort was expended in a study of this sensing method, its inherent limitations resulted in its being abandoned in favor of more promising methods. Since that time, the sensing methods which have been investigated have included electrostatic, acoustic, radio-active, optical, and electromagnetic. The greatest developmental effort has been spent on electromagnetic sensing systems employing FM radar, phase comparison, CW radar, Doppler, and ringing frequency techniques, as well as combinations of these techniques. The physical principles involved in the first four systems are available in the literature.⁷ The fifth technique makes use of an

⁶ This section prepared by P. Yaffee.

⁷ See for example, M. A. W. Ismail, "A precise new system of FM radar," *Proc. IRE*, vol. 44, pp. 1140-1145; September, 1956.

electronic transponder in both the missile and the target to set up an oscillating space-coupled feedback loop when the two are close enough. Since the frequency of oscillation is a function of the loop delay, which is proportional to the space delay, the oscillation frequency is a measure of the instantaneous distance between missile and target. The space delay at the point of closest approach will be a minimum and the ringing frequency a maximum. Calibration of the system provides the miss distance at closest approach.

In addition to classification on the basis of operating principle, electromagnetic sensing systems have generally been classed as active or passive, active methods requiring a source of radiation in the missile, passive methods requiring none. It is apparent that a passive system is highly desirable. However, the inherently low signal-to-noise ratios achievable with these systems limit the measurable miss distance range to a figure less than that required for evaluating the performance of most missile systems. For that reason, the major current effort on miss distance measurement using electromagnetic principles is concentrated on active methods.

Because the results of the efforts of both government and industrial laboratories to that date had been disappointing, the Navy Department, Bureau of Ordnance, in late 1953 requested the Naval Ordnance Laboratory, White Oak, Md., to undertake the development of a miss distance measuring system for missile system, gunnery, and fire control evaluation. The result of this development is the AN/USQ-11 Miss Distance Measuring System (MDMS).

AN/USQ-11 SYSTEM

Purpose and Principle

The AN/USQ-11 MDMS is an active electromagnetic system which operates in the VHF region and uses the Doppler principle as its operating basis. It is capable of measuring the minimum slant range (miss distance) between a missile, rocket, or projectile and its airborne target, their intercept velocity, and their instant of closest approach, each with a high degree of accuracy. The system has become a unit of standard range instrumentation at the White Sands Proving Ground, N. M., and is in use aboard ship and at the test ranges of the Navy, Army, and Air Force. It is being used effectively in the evaluation of weapon systems and is a valuable training aid in the instruction and scoring of missile crews.

The MDMS makes use of a radiating source in the missile to transmit a signal at a given frequency f_m . This signal, as observed at a receiving antenna on the target aircraft, will be higher than the transmitted signal as the missile approaches the target, equal to the transmitted signal at the point of closest approach, and lower, as the missile departs. This apparent frequency difference f_{ma} is, of course, the familiar Doppler shift. It is proportional to the relative velocity between the

source and the receiver and is equal at any instant to the magnitude of the source velocity vector component, which is directed toward the receiver, divided by the wavelength of the radiated signal.⁸ A plot of f_{ma} vs time during the intercept will appear as Fig. 10. Two important pieces of information are contained in this curve for any given intercept situation. The first, the relative or intercept velocity, is a function of the distance between the asymptotic limits of $\pm V_r/\lambda_m$, where V_r is the magnitude of the relative velocity vector, and λ_m is the wavelength of the radiated signal. The second, the minimum slant range or miss distance at closest approach, is a function of the maximum slope of the curve at closest approach and the relative velocity. Derivation of the relative velocity and miss distance are thus dependent upon the determination of the asymptotes and the maximum slope S of the Doppler curve,

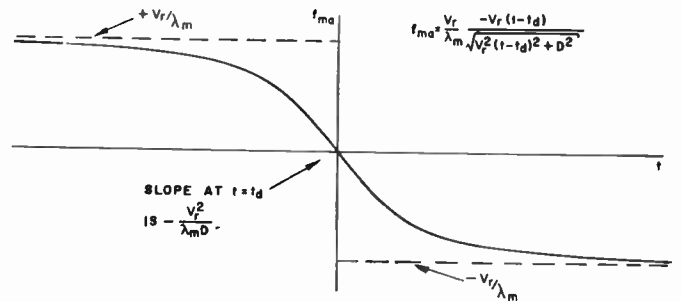


Fig. 10—Doppler frequency of signal received at the target.

The AN/USQ-11 MDMS measures the change of Doppler frequency during the intercept and provides a recorded trace of the form of Fig. 10. The data implicit in this plot is extracted by a mechanized analysis technique which solves the following two expressions:⁹

$$V_r = \frac{\Delta f_{ma} \lambda_m}{2} \tag{1}$$

and

$$D = \frac{V_r^2}{\lambda_m S} \tag{2}$$

where V_r , λ_m , and S have been defined previously, Δf_{ma} is the total Doppler frequency excursion during intercept, and D is the miss distance at closest approach.

General System Description

The MDMS is composed of three components whose location has been shown in Fig. 11. They are a simple transmitter in the missile, a transponder with its receiving and transmitting antennas on the drone or towed-target, and a receiver-recorder console with its

⁸ L. Page, "Introduction to Theoretical Physics," D. Van Nostrand Co., Inc., New York, N. Y., 2nd ed., p. 618; 1935.

⁹ P. Yaffee, "The AN/USQ-11 Miss Distance Measuring System," Proc. East Coast Conf. on Aeronautical and Navigational Electronics, Baltimore, Md., pp. 71-80; October 27-28, 1958.

antenna on the ship or at the ground monitor station.

The missile transmitter signal, with its missile-to-target Doppler shift, is received at the target's receiving antenna. The transponder acts as a beacon, which is excited by this signal during the intercept between the missile and target. It merely changes the input frequency by heterodyning with a self-contained crystal oscillator and amplifies the desired sideband to the proper level for retransmission to the shipboard receiver. The latter receives two input signals; one, from the target transponder, the other direct, from the missile transmitter. The basic function of the shipboard receiver is to heterodyne these two input signals and to extract the missile-to-target Doppler shift during intercept for recording and analysis. A detailed discussion of the theory of operation is contained in the previously mentioned paper.⁹ In the frequency conversion process, the direct missile signal is used as a reference oscillator and is eliminated. This is an important feature of the system since the effect of frequency drift in this signal is also eliminated. Thus, the only frequency stability requirement placed on the missile transmitter is that it remain within the system passband during the operating period. This is easily achieved.

corresponding reduction in the accuracy of miss distance measurement. A study was thus undertaken to determine qualitatively the seriousness of the effect.

An analysis based upon the use of the entire aircraft structure as a receiving antenna becomes almost hopeless because of the extremely complex nature of the electromagnetic field in the vicinity of the structure. It was thus necessary to make the simplifying assumption that the aircraft could be approximated by a simple antenna system. Under this assumption, the study compared the equations for miss distance derived by using two approximations for the aircraft antenna. These were: first, the aircraft could be represented electromagnetically by a single isotropic antenna; and second, it could be represented more realistically by a long thin linear antenna oriented along the aircraft fuselage and an odd number of half-wavelengths long. The analysis of the thin linear antenna is simplified by assuming a sinusoidal current distribution. In this case, the electromagnetic field as a radiator would be the same if the linear antenna were replaced by three isotropic sources, two at the end points and one at the center of the antenna.¹⁰ Further, for antennas equal in length to an odd number of half-wavelengths, the contribution of the center isotropic source vanishes, and the linear antenna can be replaced by two isotropic sources at the ends. Thus, by specifying a thin linear radiating antenna, an odd number of half-wavelengths long for the aircraft target, the latter may be represented by two isotropic radiators located at the ends of the fuselage and, by reciprocity, by two isotropic receiving elements similarly located. For the single isotropic antenna, the equiphase surfaces are concentric spheres about the aircraft; for the thin linear antenna, the spheres become distorted into confocal ellipsoids of revolution.

A comparison of the miss distance equations obtained under the two approximations outlined above resulted in an expression for the error existing in a value of miss distance as a result of assuming the target to be an isotropic antenna rather than a thin linear antenna. This error, of course, will not be the total error caused by assuming the target aircraft can be replaced by a simple antenna. It will, however, be an expression which indicates the approximate lower limit of the total error.

The conclusion reached as a result of the above study was that the use of the entire target aircraft as a receiving antenna introduced errors which could not be tolerated. The required miss distance measurement accuracy could only be achieved by a receiving antenna which was sufficiently decoupled from the aircraft structure to be essentially a point source. Such an antenna was developed for the F6F-5K target drone (Grumman Hellcat), using the Naval Ordnance Laboratory's model antenna range as a test facility. It consisted essentially of a half-wave dipole mounted one-quarter wavelength

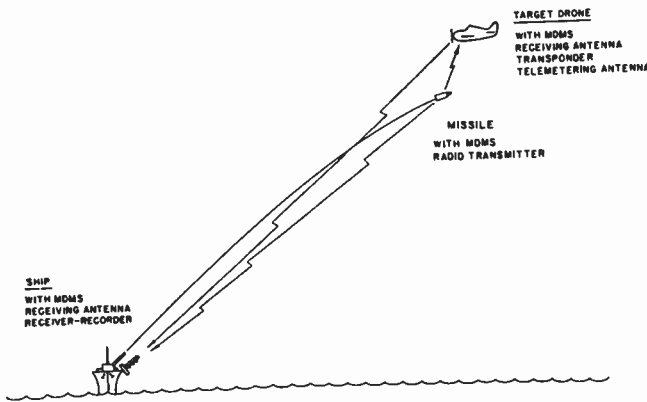


Fig. 11—Primary MDMS components and their location.

The effect of other Doppler frequency components, which are generated in the system due to the relative velocity between the missile and ship and between the target and ship, is also discussed in the previous paper.

System Development—Critical Areas

Early in the development of the MDMS, it became apparent that certain critical areas existed which would directly affect the operation and accuracy of the system. Foremost among these was the question of whether a receiving antenna installed on an aircraft target could be made to have spherical equiphase surfaces. Since the instantaneous Doppler frequency is derived from the missile's rate of crossing of the receiving antenna equiphase surfaces, any appreciable departure of these surfaces from spherical would result in a distortion of the recorded Doppler frequency vs time curve, with a

¹⁰ E. C. Jordan, "Electromagnetic Waves and Radiating Systems," Prentice-Hall, New York, N. Y., pp. 320-325; 1950.

from the wingtip of the plane. Since that time, receiving antennas of various configurations for several types of jet and propeller-driven drones and for towed-targets have been developed. All have satisfied the requirement of being well decoupled from the structure of the target.

In general, during the development phase, conventional RF circuit techniques were used on the electronics of the system. The major effort was concentrated on developing equipment which would be adequate to prove-in the system concept, rather than be sophisticated. Particular attention, however, was paid to the problems of receiver bandwidths, crystal oscillator stability, and discriminator output filtering. In the previous paper, it was shown that the Doppler shift during intercept appears at the shipboard receiver discriminator as a variation about the final IF frequency.⁹ The optimum band-pass for the receiver is a function of the signal-to-noise ratio, the maximum Doppler swing expected, the amount of interference rejection desired, etc. It is thus desirable to limit the band-pass to the minimum which is compatible with the expected missile-target closing rate. At the same time, the use of a narrow band-pass imposes severe stability restrictions on the oscillators of the system. It has been mentioned above that the missile signal undergoes a frequency conversion in the target transponder by mixing with a crystal oscillator. It has also been noted that the missile signal itself is eliminated in the shipboard receiver conversion process. The outcome of this conversion is thus a composite frequency equal to the target oscillator frequency plus the Dopplers existing in the system during the particular intercept. In order to obtain the final IF center frequency, the receiver contains a crystal oscillator which is tuned below the transponder crystal frequency by an amount equal to this final IF. The over-all stability of the system is thus a function of these two crystal oscillators. Further, the difference between these crystal frequencies must be such that, taking into account the two frequency biases due to target-to-ship and missile-to-ship Doppler, the composite frequency is sufficiently within the final IF band-pass to permit the total missile-target Doppler swing to be accommodated.

For noise limiting purposes, the output of the discriminator was passed through a low-pass filter, designed to pass all significant frequency components of the Doppler wave form without introducing appreciable distortion. This filter was designed to conform to the results of a study by Ender.¹¹

The AN/USQ-11 MDMS provides a recorded Doppler trace as its output. It is then necessary to perform an analysis of the record to extract the miss distance and relative velocity data. Several analysis techniques were investigated during the system development phase for the purpose of finding the most accurate, rapid, and

practical method. This investigation led to the design of the ideal Doppler curve plotter of Fig. 12. This device draws a curve of the equation

$$y = \frac{-ax}{\sqrt{x^2 + b^2}}, \quad (3)$$

which is of the same shape as Fig. 10. The latter is a curve of the equation

$$f_{ma} = \frac{V_r}{\lambda_m} \frac{-V_r(t - t_d)}{\sqrt{V_r^2(t - t_d)^2 + D^2}}, \quad (4)$$

where the factors are seen to correspond to those of (3).

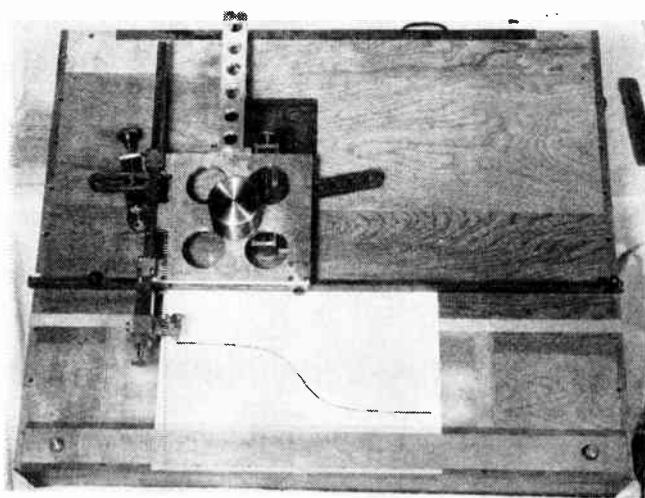


Fig. 12—Doppler curve plotter.

The operator superposes an ideal Doppler frequency curve on the curve recorded in the field. The equivalent miss distance and relative velocity parameters, b and a respectively, are adjustable on the curve plotter. The operator adjusts these parameters until the curve drawn by the plotter identically matches the recorded curve. By reading the parameter values from scales provided on the plotter and multiplying these values by frequency and time calibration constants obtained from the recorded trace, the actual values of miss distance and relative velocity at closest approach are obtained.

Although the mechanical curve plotter technique was an acceptable method of analysis, the fact that it was essentially a point-to-point method which did not permit the operator to view the entire ideal curve at once (*i.e.*, without actually drawing and erasing a poor choice) led to a search for a better method. Additional reasons were the need for computation and the time required to do the analysis.

Further work on analysis methods resulted in the optical comparison system of Fig. 13. This device, affectionately named "Mickey Mouse," is a modified Eastman Kodak Recordak, with the ground-glass viewing screen replaced by a plastic plate. Mounted on this plate are supply and takeup reels for the recorded

¹¹ R. C. Ender, "The Effects on a Doppler Waveform of Transmission Through a Filter Network," U. S. Naval Ordnance Lab., White Oak, Md., NAVORD Rep. No. 4098; September 16, 1955.

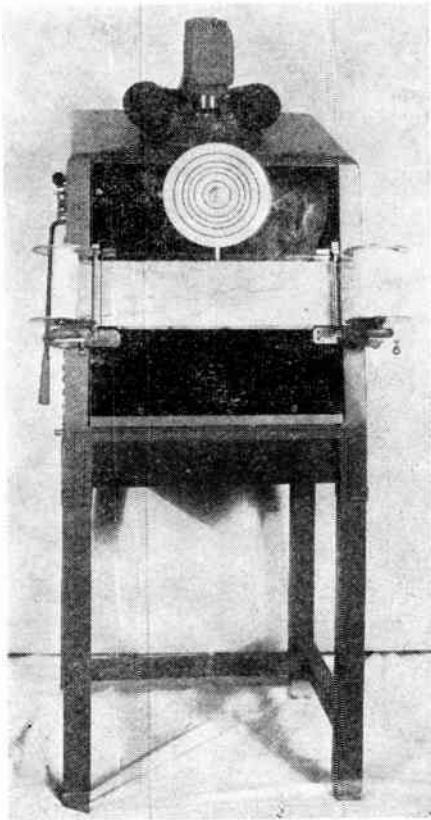


Fig. 13—Optical comparison data analyzer.

traces and a modified circular slide rule. With the exception of the area behind the recorded trace, the plate has been made opaque. A set of ideal Doppler curves, generated by the curve plotter of Fig. 12 and photographically reproduced on 35-mm film, is mounted on the Recordak projector unit. The ideal curves are individually projected from behind onto the front plate and are viewed through the translucent record tape. The proper ideal curve and Recordak magnification are chosen to achieve exact matching with the recorded Doppler trace. The modified slide rule and converter disk permit direct insertion of all record parameters and the immediate read-out of the miss distance and relative velocity from the slide rule scales. The total analysis time including matching and read-out is approximately two minutes.

MDMS—Developmental Prototype

The development phase on the MDMS resulted in a set of equipment which was adequate to prove-in the system concept and demonstrate feasibility. This was successfully accomplished and the accuracy of the system determined by comparison with optical standards. The measurement accuracy of the system was shown to exceed that of the development specification. Based upon the excellent results obtained in the use of the developmental MDMS at missile ranges and at sea, the Bureau of Ordnance in March, 1957, negotiated a contract with Aircraft Armaments, Inc., Cockeysville, Md., for production engineering of the system. Tech-

nical administration of this contract was assigned to the Naval Ordnance Laboratory. The current system to be described below is the outcome of this contract.

PRODUCTION ENGINEERED AN/USQ-11 SYSTEM

General Details

The purpose of the production-engineering work on the AN/USQ-11 System was to repackage the developmental prototype for compliance with MIL specifications, reliability, and quantity production. A corollary purpose was the miniaturization of the airborne equipment to make it adaptable to as many different types of targets as possible. Re-engineering for production was required with no deterioration in system performance.

Many difficulties are encountered in the transition of electronic equipment of this type from the developmental prototype to the production prototype. Reduction in the size of components and subsystems often results in heat dissipation problems with related frequency stability and reliability problems. Care is required in the design of the high-frequency circuitry to obtain the best possible signal-to-noise ratios within the limits imposed by MIL standard components. For these and similar reasons, the effect of changes in electrical and mechanical design on the system performance must be carefully considered during the early stages of production engineering.

System Components—Missile

As has been indicated above, the AN/USQ-11 Miss Distance Measuring System has three components. The first of these is the missile transmitter. Missiles equipped with a standard telemetry system do not require an additional signal source to operate the MDMS, since the telemetering carrier will provide the primary frequency source. Missiles and projectiles, which do not contain such a transmitter, must be provided with a signal source. Since the number of missiles and projectiles, each with its particular configuration and space arrangement, is so large, no attempt will be made to describe a typical transmitter. Suffice it to say that only the simplest oscillator and power supply are required.

Although the system concept has no basic limitation frequency-wise, the production-engineered system is designed to operate at any selected frequency within the 215- to 235-mc telemetry band. The system input bandwidth of 4 mc makes it possible to accept several different telemetry frequencies during any given target aircraft mission. This provides for versatility of operation and economy of drone flight time.

System Components—Target

The second component of the system is the airborne transponder with its receiving and transmitting antennas. The receiving antenna is tailored to the specific

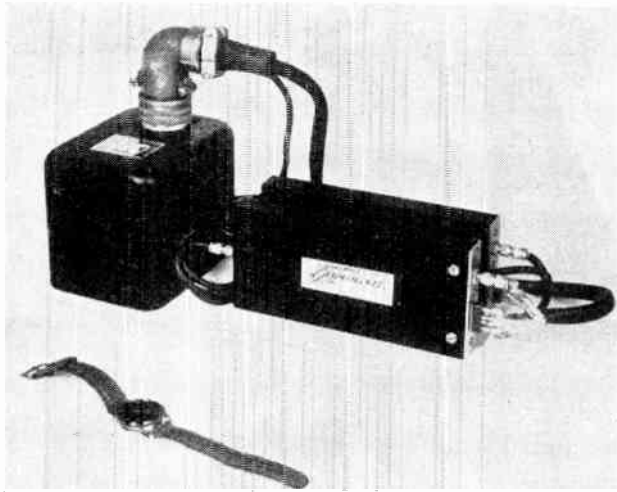


Fig. 14—MDMS airborne transponder.

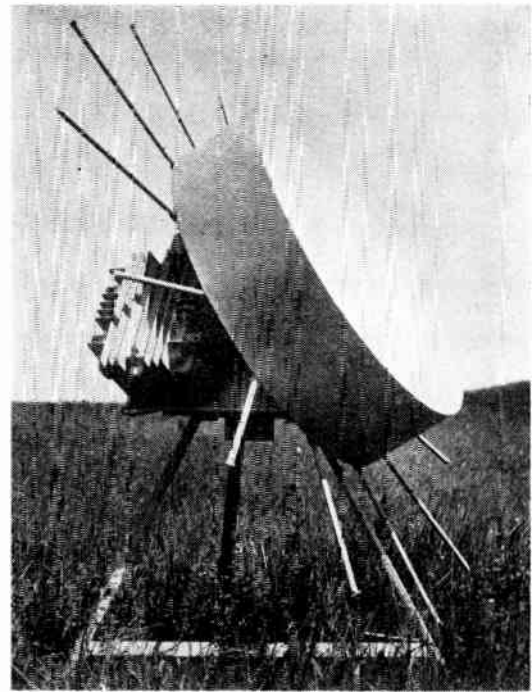


Fig. 16—MDMS shipboard receiving antenna.

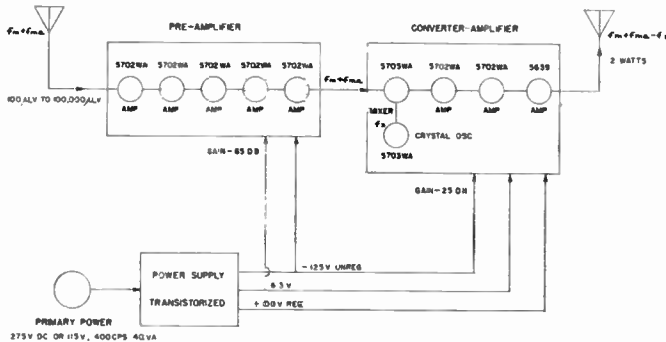


Fig. 15—MDMS airborne transponder.



Fig. 17—MDMS shipboard antenna—spiral element.

target being instrumented. A typical production-engineered receiving antenna for the F9F-6K (Grumman Cougar) drone is a coaxial dipole mounted on the nose.

The transmitting antenna used on the target is a simple stub tuned to the output frequency of the transponder. It is installed at a location such that essentially omnidirectional coverage is achieved.

The production-engineered transponder has been sub-miniaturized and sub-divided into three packages as shown in Fig. 14 and Fig. 15. The complete transponder is designed to provide 2 watts output for an input signal level of 1000 microvolts. The saturation level of 1000 μ v was chosen to give full output for all missile-to-target slant ranges and geometrical conditions of interest.

All airborne components of the MDMS have been designed to meet MIL-E-5400B and MIL-E-5272A specifications and are operable to an altitude of 60,000 feet. The production-engineering program has resulted in a twenty-to-one saving in volume, a seven-to-one saving in weight, and has produced a much more versatile airborne transponder installation.

System Components—Shipboard

The shipboard components of the Miss Distance Measuring System include an antenna assembly and a receiver-recorder console. The antenna assembly is

shown in Fig. 16. It has three functional parts; an antenna, a pre-amplifier, and a pre-amplifier power supply. The antenna is a circularly-polarized logarithmic spiral with the following characteristics:

Gain	6 db
Beamwidth	80°
VSWR	2.0 to 1
Axial Ratio	3 db approximately

The spiral element shown in Fig. 17 is fabricated by the use of printed circuit techniques. It is mounted approximately 12 inches from the ground plane reflector by means of foam plastic. The antenna is fiberglass-wrapped and resin-impregnated for strength and weather-proofing. The feed point for the antenna is seen at the center of the spiral element. The antenna is designed to accept both the signal direct from the missile trans-

mitter and that from the target transponder. The pre-amplifier performs the initial conversions on the two receiver signal inputs. It permits the transmission of signals over long cable runs to be done at lower frequencies, thus reducing the effect of cable attenuation. This is of importance in shipboard installations where the antenna may be several hundred feet from the receiver. Of particular interest is the location of system self-test circuitry in the pre-amplifier unit. This permits the receiver-recorder console alone, the complete shipboard installation alone, or the entire MDMS system, including target drone equipment, to be tested immediately prior to a missile launching. The test circuitry is remotely controlled from the receiver-recorder console. This test capability had been present in the developmental MDMS. However, location of the test circuitry in the receiver console had made it necessary to provide a second helical antenna for transmission of test signals to the aircraft. Location of the test circuitry in the pre-amplifier and the use of a coaxial relay switching network eliminated the need for a second shipboard antenna in the product-engineered system.

Operating controls for the MDMS are located on the receiver-recorder console of Fig. 18. This unit contains the receiver circuitry required to accept the output signals from the pre-amplifier, operate on them, and to extract the missile-target Doppler. Of particular interest in this unit is the calibrator, which provides crystal-controlled standard frequencies for calibration of the receiver's discriminator. This serves to calibrate the frequency axis of the recorded trace so that the total Doppler swing between asymptotes may be measured. Manual, semi-automatic, or automatic calibration modes are provided for use during test and operation. For the latter, automatic calibration is accomplished upon the actuation of the recorder paper drive switch.

The use of a Visicorder direct-writing oscillograph recorder provides a Doppler trace, without the need for photographic processing, immediately after the intercept has occurred. The data contained in the trace may be extracted and are available within minutes. A timing generator provides marker pips at 0.01 second intervals for calibration of the time axis of the recorded trace. In addition, a reference mark is generated for transmission to the ship's telemetry recording station. This permits time correlation between the MDMS record and the telemetry record.

The shipboard components have been designed to meet MIL-E-16400B specifications.

V. TEST SET ELECTRONIC DEVELOPMENT¹²

The special considerations applicable to naval ordnance extend beyond the design of the weapon itself. The design techniques applied to equipment to checkout or nondestructively test certain specialized weapons are

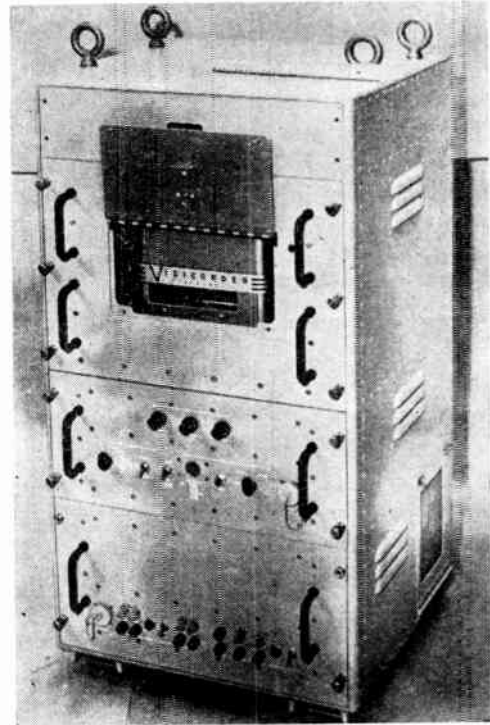


Fig. 18—MDMS shipboard receiver—recorder console.

perhaps even more unique. The test equipment referred to is that designed and provided as part of the total weapon system for use by the procuring and operating arms of the Navy in employing a given weapon. Although challenging and ingenious designs are required for developmental testing throughout the research and developmental phase of an ordnance development, a more specialized field is presented in consideration of suitable equipment to be used by nondevelopment engineering personnel in logistic handling of the production weapon. Experience gained in World War II and since, has indicated the great importance of simultaneous design of this military test equipment and the weapon hardware—that is, design of the weapon system together with the testing equipment. Only in this manner can the special needs of testing be provided in the weapon design, and conversely, the availability of testing knowledge used to assist in simplifying the weapon component details.

The purpose of such naval ordnance test equipment is to insure that characteristics originally designed into the weapon are retained in sufficient detail at the time of final use to provide the needed military performance. Implicit in this statement is the fact that reliability cannot be tested into a weapon—reliability must be designed into a weapon. But testing can be used to insure that the designed reliability is retained up to the end military use of the ordnance.

With the present-day complexity of weapons, it has been noted that the dollar cost of maintenance and testing of the average weapon is more than ten times the original cost of the weapon. For this reason alone, early

¹² This section prepared by R. B. Knowles.

consideration and an optimum plan for the test equipment program is most necessary.

Unlike many military electronic uses, certain specialized naval ordnance weapons are characterized by a) relatively small numbers of a given type in fleet use; b) inability to give full operational exercise of the weapon; c) relatively high cost of operational support and delivery to cost of weapon; d) unavailability of industrial technical representatives at the checkout and use point; and e) the extreme variation, geographically and environmentally, in the characteristics of the weapon staging point.

Although the individual pattern for each specific weapon can be defined only in terms of the actual weapon's needs, the general chronological application of Navy weapon test equipment can be grouped into the following categories:

1) **Production Test Equipment:** This equipment normally encompasses that which is provided by a given fabrication contractor to insure that the weapon hardware being manufactured meets the contractor's self-established goals. This equipment is normally not specified in Navy contracts and the total provision and use becomes the responsibility of the hardware fabricator.

2) **Acceptance Test Equipment:** This equipment is that used by the on-the-spot representatives of the procuring agency (usually the Inspector of Naval Materiel) and is used to measure the quantities or performance specifically called for in the procurement document (Ordnance Specification). As a result of the acceptance tests the ordnance material is accepted by the government for transfer from the civilian manufacturer to the naval storage or stockpile entry point.

3) **Stockpile Entry Tests:** These tests are made by the personnel of the ordnance depot to determine general suitability of the ordnance for stockpiling purposes. Tests are intended to be broadly representative of the acceptance tests, though generally to less rigid standard and less complete.

4) **Component Ready Tests:** These tests are made on material as it is withdrawn from general stockpile for issue to an endpoint user in response to a specific need. This test is usually the last test made on an individual part of a weapon, is made by relatively unskilled personnel, and is made on components after having experienced a major portion of their storage life. As a consequence, the prime importance is to determine that such storage has not resulted in abnormal deterioration of the component and that the component can be expected to provide satisfactory service if combined into a complete weapon and used in the relatively near future.

5) **Assembly Ready Tests:** This test on a weapon is made on the complete assembly of the individual components by the assembly activity. Since the individual components have recently been individually tested its principal goal is to assure that the individual parts have

been properly assembled, and that all components are properly installed.

6) **System Check-out Tests:** System tests are frequently combined with the assembly ready tests. The purpose of the Systems assembly tests is to determine, where necessary, that over-all compatibility of adjustments, sensitivities, response times, etc., are within limits. Such specific tests are necessary where individual component limits cannot be maintained with sufficient accuracy to insure the desired over-all weapon performance.

7) **Monitor and Safety Tests:** This set of tests is made with the sole purpose of guaranteeing safety to handling, loading, or using activity of the weapon. It is generally made by nontechnical personnel and is not concerned with predicting reliability of the weapon.

It is readily apparent that from the test set designer's point of view, a major difference exists between the requirements of steps 1-3 and those of steps 4-7. Test equipment to meet the needs of steps 1-3 is clearly of a type that is amenable to use of laboratory-type equipment (perhaps assembled for cohesion as a unified group) and intended for operation by well trained technicians under surveillance of qualified engineers. Problems of calibration and maintenance can be solved by the further use of standard laboratory type instrumentation and procedures. Where feasible, it has been found most economical in naval ordnance to define the test procedures required in steps 1-3 by specifications. These specifications define what and how accurate the various measurements shall be and permit the choice of equipment, maintenance, and calibration to be provided by the weapon fabricator under the surveillance of the applicable inspector of naval materiel. Fig. 19 illustrates a typical assembly of this type of equipment.

Steps 4-7, however, have required a different approach to the problem. In these areas, the testing needs must be met through use of naval personnel of only limited technical background in the specific weapon, and with only a minimum time consumed by the testing operation, since the desired function is one of getting the weapon into deployment as expeditiously as possible.

This latter need is met through the provision of a set of field test sets for each of the weapon programs. For most weapon programs, their development is undertaken simultaneously with the development of the weapon itself. Experience has shown that this early start is essential if the basic program needs are to be met with well designed test sets that meet the weapon needs, yet in themselves fulfill the rigid design goals found necessary for rugged field usage by inexperienced personnel.

Fig. 20 illustrates one such set of field test sets for a single weapon program. Six of the individual units provide capability for performing Component Ready Tests (item 4 tests) on weapon components, while a

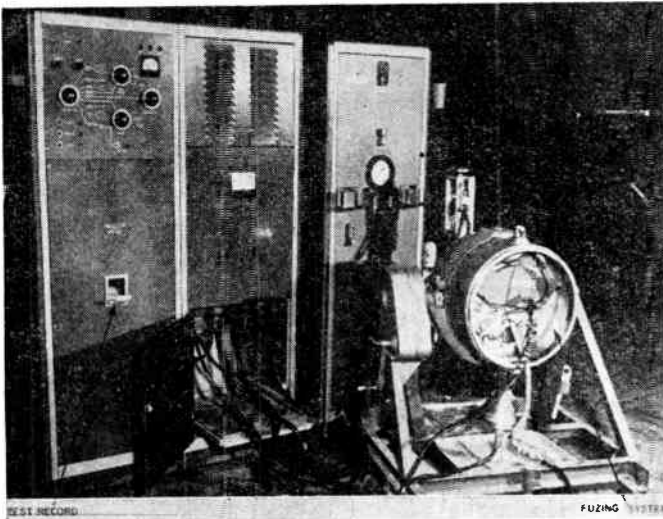


Fig. 19—Production—acceptance type test equipment—mk 101 fuzing system.



Fig. 20—Set of field test sets for one Navy weapon system.

seventh unit when combined with one of the previous six units (Hydrostat Test Set) provides capability for Assembly Ready and System Checkout Tests (item 5 and item 6 tests). The eighth unit, the Monitor Test Set, is used at the final stage of weapon use and provides confirmation on the weapons safety status (item 7 tests).

This equipment is typical of that developed for most programs and incorporates special features to meet specific design goals. The significance of each of these design criteria, and some of the general methods of providing for each is discussed in the following:

Portability

As previously noted, many applications of field test sets are transitory in nature, with potentially long

periods of time in which use is not required. For maximum economy in assembly space, it is most desirable to provide portability to provide rapid changeover from one weapon processing to another. In addition, portability is often mandatory because of the physical location of variables; *i.e.*, a specific weapon may entail final testing at air bases, on supply ships, or in temporary forward areas. For these reasons, most field test set programs for specialized naval ordnance (mines, depth charges, depth bombs) are based on the concept of portability. However, since such portability itself introduces problems of ruggedness, infrequent maintenance and repair capabilities, the designs emphasize reliability over miniaturization. The equipment of Fig. 20 illustrates the degree of portability nominally employed.

Environmental Ruggedness

The four factors which cause better than 90 per cent of the failures in such test equipment include moisture, temperature, shock, and vibration. As can be seen in the unit of Fig. 21, a special design technique has been developed to cope with these problems. Careful attention to mechanical detail has permitted complete moisture proofing at every control and indicator junction passing through the panel. Use of completely sealed panel components has permitted the use of a gasket sealed rear cover or pan to insure complete test set moisture-proofing. A small indicating type humidity monitor is provided on the panel to assure that this integrity has been maintained. Resistance against vibration effects is provided through use of thoroughly tested components, together with a generous "brute force" mounting framework. This framework, illustrated in Fig. 22, provides a wrap-around or box-type mounting for the heavier components in the internal area, and permits mounting of the smaller items on the outer periphery of the chassis. Additional benefits from this type construction are evident from the easy-to-get-to wiring terminations, openness for maintenance, and easy part identifications. Since a sealed unit results, it is apparent that temperature rise from internal heating must be considered. Effective solutions have been the minimization of heat producing items through use of solid state elements, provision of heat sinks and heat transfer through conduction of the metal panels, and generous derating of electrical components. Shock is a major problem in any portable equipment. These shocks come from dropping the unit from loading platforms, trucks, etc., or that received in normal use position such as when used aboard ships. To provide effective protection against this environment, extensive use is made of self contained fiberglass carrying cases which serve as a shock absorber. Since these fiberglass containers are also the only needed shipping containers, the problem of insuring adequate "packaging" is eliminated.

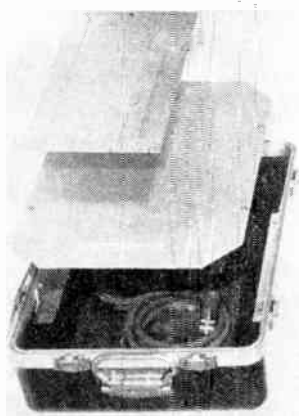
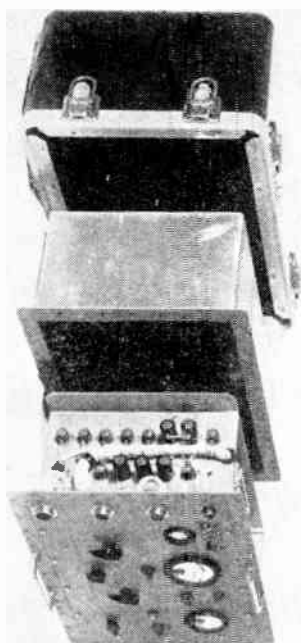


Fig. 21—Exploded view of underwater mine field test set.

Maximum Simplicity in Operation and Minimum Training

Recognizing that highly trained operators will not always be available, especially in time of war, a large amount of the electronic design effort is concentrated on simplification of the operator requirements. Although a minimum of dependence on automatic sequencing or automated operation is sought, it is necessary to minimize the manual effort required for test set operations. Major emphasis is placed on careful test and circuit design to permit a single measurement to convey a maximum of information about the weapon's condition. Closely related to this need for engineering a maximum amount of information from each individual test is the demonstrated gain in reliability from use of easily interpreted "go-no-go" read outs. Over a period of several years naval ordnance programs have sponsored a series of controlled experiments in the area of "Human Engineering" to insure that a maximum

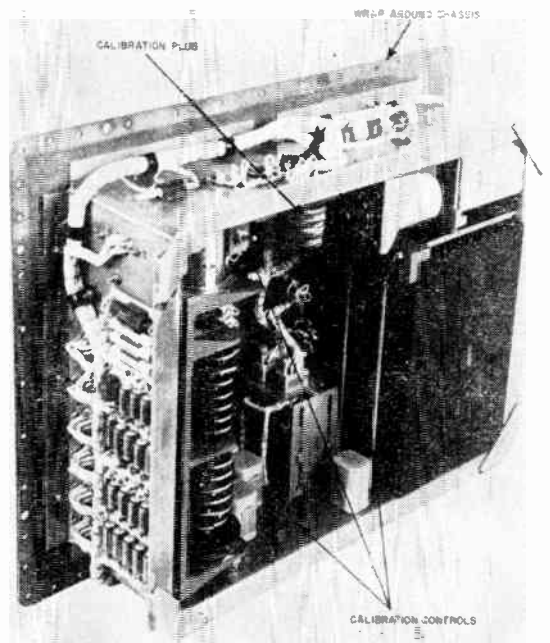


Fig. 22—Test set internal construction.

amount of effectiveness is preserved in the end use of the test equipment. This has resulted in guide line information which is made available to the test set engineering personnel. A sample page from one such reference is reproduced in Fig. 23. A typical front panel can be seen in the photograph of Fig. 24.

Precision and Stability Commensurate with Weapon Needs

As noted earlier, the use of field type test sets in naval ordnance is usually limited to the needs of Items 4-7 of the testing sequence outlined. In general, the accuracy requirements of the equipment are an order of magnitude less severe than the corresponding measurements made in Steps 1-3 of this test sequence. In establishing such limits, it is essential that test schedules be so designated that each succeeding test will not reject assemblies because of the spread of testing specifications, overlap of specifications, or differences in the accuracy of test equipment used at any point in the cycle. This can only be attained by adequate recognition at the start of the weapon development itself of the needs for a logical testing program and degradation of testing limits throughout the logistic supply of the weapon.

Fail-Safe and Self-Testing Circuitry

To insure a satisfactory field test set application, the test set circuitry must be engineered for fail-safe features. Circuits must be designed such that failures in either the test set or the material under test will result in rejection rather than acceptance of the material. This requirement usually results in quite complex test circuitry. As much as 60 per cent of the circuitry of a

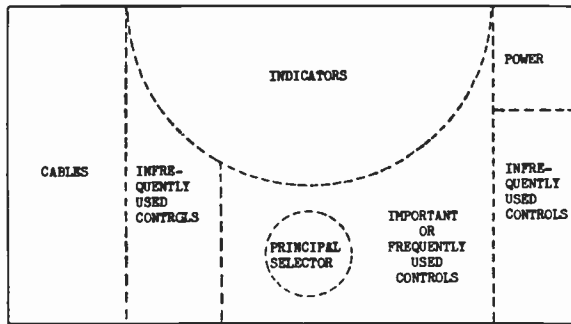


Fig. 23—Control panel design criteria. 1) All controls and indicators have labels *above* them. 2) Numbered or lettered components, such as switches, are numbered or lettered from left to right on the panel. 3) Cable receptacles, power receptacle, and fuses are on the left side of the panel. 4) All cables and connectors are numbered and/or color coded. 5) Cable receptacles are stacked with the larger or more bulky receptacles nearer the bottom of the front panel and the power receptacle and fuse at the top. 6) Clockwise movement of rotary controls results in clockwise movement or increased reading of indicators. 7) Toggle switches move up or to the right for the "ON" position. 8) The test set control panel is generally laid out according to the figure.

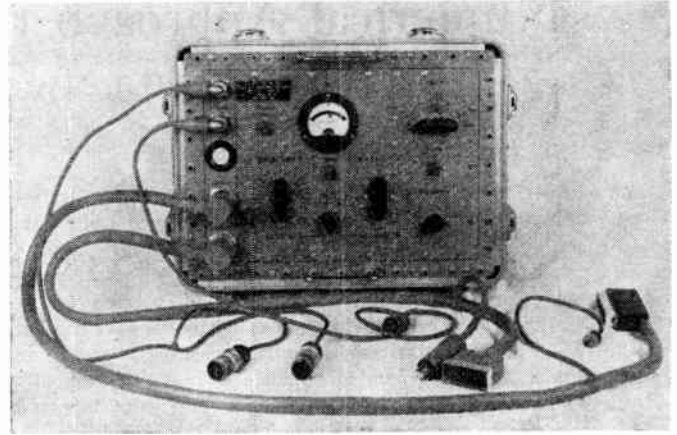


Fig. 24—Test set for underwater mine assembly checkout.

particular unit may be devoted to this feature. Since complete adherence to this goal is sometimes virtually impossible to achieve, this "fail-safe" characteristic is sometimes augmented with self-testing circuits to supply the needed reliability.

Ease of Maintenance and Calibration

In addition to the features already discussed, which also provide for ease of maintenance, most of the field test sets illustrated are provided with an internally accessible uniform connector which is used to provide access to desired calibration or check points in the circuitry. Such provision of a uniform means of access to calibration connections, and the knowledge that a given type of measurement is always made on the same connector pins, considerably simplifies and speeds the calibration and checkout by maintenance personnel.

Through careful attention to design detail and the use of the special techniques outlined, considerable success has been achieved in providing adequate, trouble free field test sets for checking of naval ordnance by non-technical personnel. Such a program is only feasible

where the test program is planned in concert with the design of the weapon itself, and the needs of both are made harmonious. To achieve this end, considerable ingenuity is required of the electronics engineer in providing clever, rather than just complex designs.

ACKNOWLEDGMENT

This paper was compiled and edited by J. J. Redmond, U. S. Naval Ordnance Laboratory, White Oak, Silver Spring, Md. The authors also wish to acknowledge the work of the following individuals on the projects discussed in the following sections.

Section III

Harry W. L. Street, Naval Ordnance Laboratory and Dr. Benjamin Snively, Naval Ordnance Laboratory, on leave to Pennsylvania State College, State College, Pa.

Section IV

H. P. Bell, J. K. Burke, E. H. Dilley, A. Heller, H. S. Nikirk and M. J. Parker, and in particular R. E. Grantham and H. C. Hoffman, Jr., all of the Naval Ordnance Laboratory. The contributions of Aircraft Armaments, Inc., Cockeysville, Md. as described in this section, are acknowledged.

Numerical Approach to Electronic Reliability*

JOSEPH J. NARESKY†, MEMBER, IRE

Summary—This paper describes the progress at Rome Air Development Center (RADC) in converting reliability from a black art to a science. A brief history of and the need for a reliability program for Air Force ground electronic equipment are described. Also described are the results to date and future efforts and expected results.

The RADC reliability program has had as its aim the development of techniques for quantitatively specifying, measuring and predicting the reliability of Air Force ground electronic equipment. Concurrently with these efforts, RADC has been pursuing a program of improved design techniques in such areas as cooling, reliable circuit design, redundancy and preferred circuits.

The following achievements are described: a reliability design handbook; a cooling techniques handbook; reliability specifications for development and production equipments which quantitatively specify the reliability to be achieved and outline the test procedure to determine whether the specified figure has been met; a reliability design procedure which permits an engineer to design electronic equipment to meet a specified reliability figure; and a prediction technique which permits a designer to predict, within limits, the quantitative reliability that an equipment will achieve in the field.

INTRODUCTION

THE designers of ground electronic equipment have been, as a rule, unaffected by some of the problems plaguing airborne and missile electronic equipment designers such as the need for miniaturization, operation at high temperatures and extreme shock and vibration requirements. They have, however, been disproportionately affected by a problem present in many aspects of our modern day world: the trend towards increasing complexity. With the airborne and missile people forced to build smaller and smaller electronic packages, the ground equipment designer is forced to design into his equipment many of the functions formerly performed by airborne electronic devices. This problem has been compounded by the fact that faster, higher flying aircraft and the dawn of the missile age require, in themselves, a quicker reaction capability in ground equipment and, hence, a greater complexity.

Some designers who, it must be admitted, were in the minority began feeling a bit uneasy about the increasing complexity trend in the late 40's and early 50's. In a little more than a decade, the United States had progressed from the B-17, containing a total of perhaps 100 vacuum tubes, to the B-36, containing in the neighborhood of 2700, and had advanced from a simple World War II radar set containing several hundred vacuum tubes to immense radar defense systems containing tens of thousands of vacuum tubes. However, at the time, these designers had little more than their uneasiness to justify their fears for the future, and so they were dismissed as pessimistic harbingers of gloom.

Most designers of military equipment went on their

merry way, adding components and functions to military electronic equipment and creating frighteningly complex creatures almost human in their capability. Slowly at first and then in increasing volume, reports began filtering back to the designers, reports to the effect that their gleaming monsters were just that—monsters to keep operating normally and monsters to maintain. It was found that the Navy was forced to keep nine tubes in the pipeline for every tube in operation, and the Air Force was barely able to get 20 hours of trouble-free operation from the electronic gear on bombers or 80 hours from the ground electronic equipment. Radio equipment was in trouble 14 per cent of the time, and radar equipment, 84 per cent of the time.¹

RELIABILITY COMES OF AGE

The designers began searching for the cause of the difficulty and were finally forced to concede that complexity was the main factor. Meanwhile, the harbingers of gloom who had been so lightly dismissed refused to remain dismissed and were burrowing into the problem. They defined the problem by means of the term "reliability;" and by utilizing the techniques of probability theory and statistical analysis the investigators arrived at a quantitative definition of reliability—"reliability was the probability that a device, equipment, or system would perform its intended function for a specified period of time under given conditions." They also derived the mathematical relationships existing between the number of electronic components and such terms as hazard, failure rate, reliability, and mean time to failure (see Appendix). From the accumulation of failure data on a large number of equipments, the curve of failure rate vs operating time, shown in Fig. 1, was plotted.²

The center portion of Curve A indicates that the failure rate and its reciprocal, mean time to failure, tended to approach a constant value, independent of operating time. Random, catastrophic failures seemed to predominate during this portion of the equipment's life. The initial high mortality rate (first section of Curve A) indicates a debugging period during which marginal parts were eliminated; these were predominantly quality control failures. The tail-end portion of the curve marks wearout of major equipment items. Curve B of Fig. 1 shows the effect of poor quality parts which cannot be

¹ F. J. Gwen, "Reliability in Electronic Equipment," presented before Guided Missile Committee of Aircraft Industries Assoc., White Sands Proving Ground, N. M.; April 20, 1954.

² I. Mirman, "Reliability of Electronic Equipment," RADC, Griffiss AFB, Rome, N. Y. Presented at National Airborne Electronics Conference, Dayton, Ohio; May, 1952.

C. M. Ryerson, "RCA Reliability Program and Long Range Objectives," Engrg. Products Div., RCA, Camden, N. J.; 1954.

* Original manuscript received by the IRE, December 1, 1958.

† Rome Air Dev. Center, Griffiss AFB, Rome, N. Y.

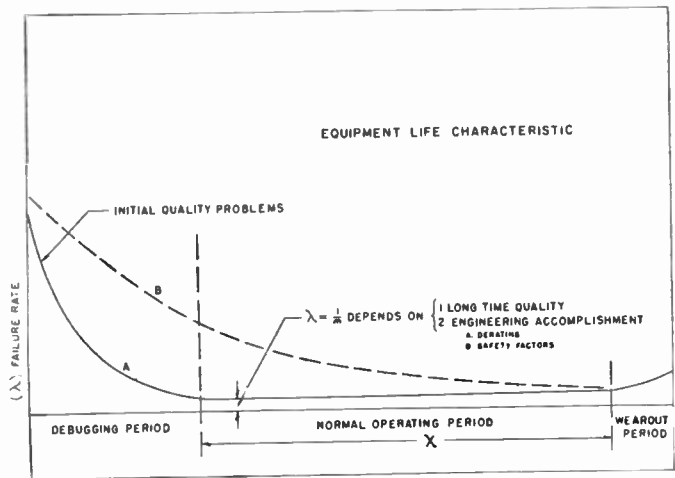


Fig. 1—Equipment life characteristic.

debugged and thus lead to unreliability and poor equipment life characteristics.

Again referring to the Appendix, the reader can see that if the hazard and, hence, its reciprocal, mean time to failure, are constant the reliability equation reduces to

$$R = e^{-\lambda t} = e^{-t/m} \quad (1)$$

where

R = reliability, as defined,

λ = failure rate,

z = hazard = λ (under the conditions given in the Appendix),

m = mean time to failure, and

t = desired time of successful operation.

Thus, a significant relationship was obtained between equipment reliability and mean time to failure, the latter term more readily lending itself to measurement. If one assumes that the curve of Fig. 1 is valid and if equipment-failure data are taken on the constant failure rate (flat) portion of the curve of Fig. 1, mean time to failure may be measured as follows:

$$\text{MTF} = \frac{\text{total number of operating hours}}{\text{total number of failures}} \quad (2)$$

Since this is a statistical process, there are problems of sample size and confidence limits for various operating times and number of failures. Time and space prevent their treatment here; the interested reader should consult the references.

Reliability could be defined quantitatively and measured under certain specified conditions. Slowly, reliability was becoming less of an art and more of a science. It was at this stage that the RADC reliability program was formulated, circa 1954.

RADC PROGRAM OBJECTIVES

From the beginning, the objectives of the program have been 1) to develop and disseminate improved design techniques for increasing the reliability of ground electronic equipment,

2) to perfect a prediction technique which would permit rapid analysis of circuit and/or equipment design and enable quantitative prediction of reliability,

3) to develop techniques for quantitatively specifying the reliability of ground electronic equipment so that these figures could be inserted into equipment specifications, and

4) to develop techniques to measure accurately the reliability of an item of ground electronic equipment and thus verify whether the specified reliability figure has been met.

It is felt that considerable progress has been made toward meeting all of the objectives. The next several sections are devoted to a description of results achieved in the areas outlined.

RELIABILITY DESIGN TECHNIQUES

Reliability Handbook

Any program of reliability improvement must, of necessity, begin by educating the equipment engineers as to what reliability is, the need for it, and how it can be designed into equipment. At RADC, this educational process took the form of a design handbook,³ a first attempt to compile in one publication most of the general information then available on reliability and reliability design practices in such fields as circuit design, components, mechanical design, heat transfer, packaging, human engineering, etc. It was disseminated to design personnel at RADC, and it was recommended that RADC contractors use the handbook as a guide to increase the reliability of ground electronic equipment. A sample page from the handbook is shown in Fig. 2.

The field of reliability has literally exploded during the past three years, so that the handbook requires extensive revision. In 1955, when the handbook was prepared, only general qualitative information on reliability design techniques was available; today, on the other hand, there is a wealth of specific quantitative information available. The need now exists for an actual designer's handbook which will permit a designer to design his equipment to meet and maintain some reliability figure. Such a handbook is being prepared in loose-leaf form with replacements or additions to be published periodically in the future.⁴

Cooling Techniques

Since temperature is one of the more significant stresses causing electronic equipment failures, it was felt that the problem was of large enough magnitude to be treated separately. Electronic designers are not noted for using optimum thermal design techniques in the design of their equipments; only through years of design experience is such knowledge acquired. This process had

³ K. Henney, et al., "Reliability Factors for Ground Electronic Equipment," RADC, Griffiss AFB, Rome, N. Y., TR-57-27, ASTIA Doc. No. AD-114274, McGraw-Hill Book Co., Inc., New York, N. Y.; April, 1955.

⁴ I. Lopatin, "RADC Reliability Notebook," McGraw-Hill Book Co., Inc., New York, N. Y. This should be completed by April, 1959.

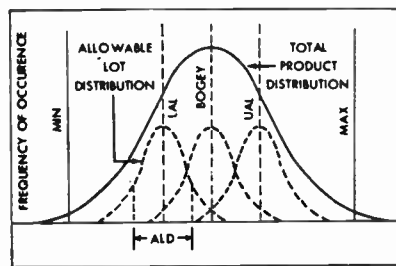


Figure 8-1. Typical distribution of I_b , S_m , μ , I_f , I_{c2} and capacitance

obtained for a tube type from the testing specification sheet.

The normal, or symmetrical, distribution is typical for the characteristics usually considered most important. It is to this type of distribution only that the present type of variables controls are applied and have meaning. The curve shown demonstrates how the total product distribution is related to the various lots distributions and how control of mean and dispersion are accomplished.

The averages or means of the samples from acceptable production lots must be between a predetermined allowable lower limit (LAL) and upper limit (UAL). The dispersion of each lot is controlled by requiring that the average range of a group of samples from the same lot be not more than some predetermined allowable lot distribution (ALD). It should be apparent that such control of each lot will produce a total product distribution of the normal type.

The distribution in Figure 8-2 is typical of such characteristics as control grid current and heater-to-cathode leakage current. These characteristics tend toward zero, with fewer tubes reading higher. Those that measure above the testing specification maximum are out of control for one of a number of possible reasons and are discarded by the tube manufacturer. Progress in reducing the amount of these currents is reflected in lower maximum values in the reliable-type testing specifications.

The skewed-type distribution (Figures 8-3 and 8-4) have a left or right skewness. The first type is typical of plate resistance and audio-frequency noise. The second, or right, skewed type, however, is typical of plate current at cutoff bias, power output, and vibration noise. Some of the characteristics of both skewness types, because of their circuit effects, should be as high as practicable. Hence, for plate resistance and power output, a minimum figure only is specified. For others, which are detrimental to reliable (or even satisfactory) circuit operation, their amount should be reduced as much as possible. Thus, for such characteristics as cutoff

8-10

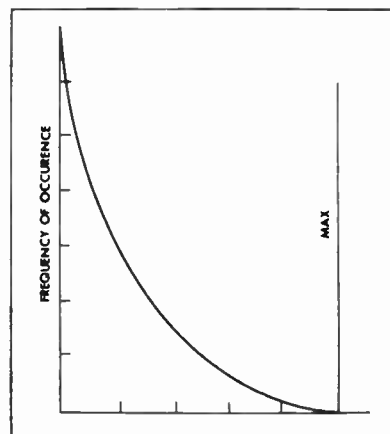


Figure 8-2. Typical distribution of I_{c1} and I_{hk}

plate current, vibration noise, and audio-frequency noise, a maximum is specified.

Almost without exception, characteristic curves included in handbooks and data sheets are average or typical curves. They describe the amount of change in the important characteristics only as test conditions are varied, and a number of curves are given to describe completely the characteristic changes. When correctly done, the curves are based on a tube that has bogey characteristics at the specified test conditions.

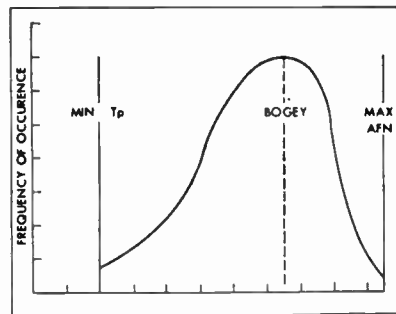


Figure 8-3. Left skewed distribution

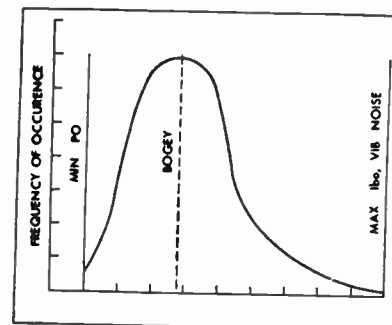


Figure 8-4. Right skewed distribution

These curves, however, do not contain any information about the variation of the other tubes of the same type around the characteristics of the bogey tube. The characteristic curves in Figure 8-5 describe in a generalized form the manner in which the total production of a type varies around the average curves.

This curve shows the spread of the product around the familiar triode plate characteristics. In this instance the spreads around the bias curves do not overlap, and in some types this will be the case. Other triodes will exhibit overlap of the spreads, and the amount will vary with the bias values selected and the tube type. Although specific curves will differ from the relative spreads shown here, these curves demonstrate the pattern of the spread to be expected in all types as the operating voltages are changed. It is characteristic for the absolute amount of spread

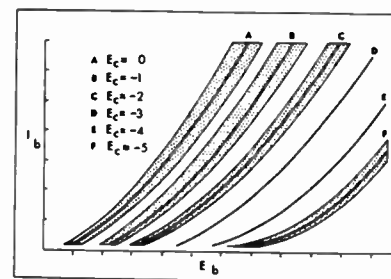


Figure 8-5. Triode plate characteristics - Effects of variation of E_c

to decrease as the amount of current drawn decreases. That is, lower plate voltages and higher bias voltages reduce the absolute differences in plate current between tubes.

Effect of Heater-Voltage Variations. Figure 8-6 shows the effect of a ± 10 percent variation in heater voltage on one of the triode plate characteristics. The increased heater voltage moves the limit curves upward from those for the normal or specified heater voltage. Of course, the average curve is moved upward also. Conversely, the reduced heater voltage causes a lowered position of the $I_b - E_b$ limit curves. The amount of change in the plate current limit curves at ± 10 percent E_f may depart as much as 20 or 25 percent from the values at specified heater voltage.

The amount of plate current change (and other characteristics) varies with the operating plate voltage and the emission-temperature characteristic of the cathode. If a 10 percent change in heater voltage occurs over a flat portion of the emission-temperature characteristic, the resulting change in plate currents will be small, and the upper and lower limits of the spread will be affected equally. If the emission-temperature characteristic is not flat, the change in plate current limits will be greater. The amount of change of the upper and lower limits may not be the same for the ± 10 percent E_f and -10 percent E_f values. In some instances it may be found that a proportionately greater drop will occur for the lower limit of the -10 percent curve. Here, as is the case in applying the other curves to specific tube types, data on the effect of heater voltage variation on the characteristics of a tube should be obtained for the type under consideration.

There have been indications, although the point should be investigated more thoroughly, that E_f variations affect triode characteristics. Everything else being equal, such as design considerations, cathode

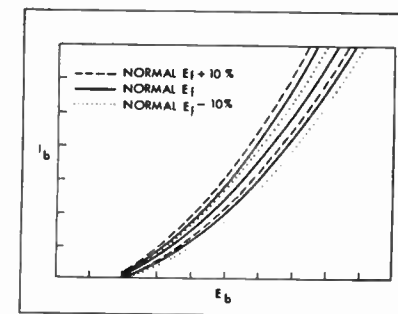


Figure 8-6. Triode plate characteristics - Effects of variation of E_f

Fig. 2—Sample page from RADC reliability handbook.

8-11

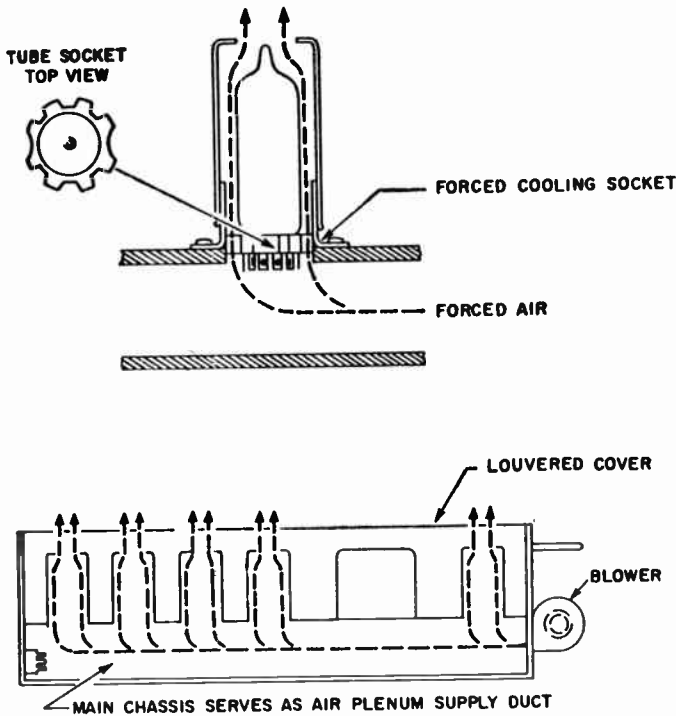
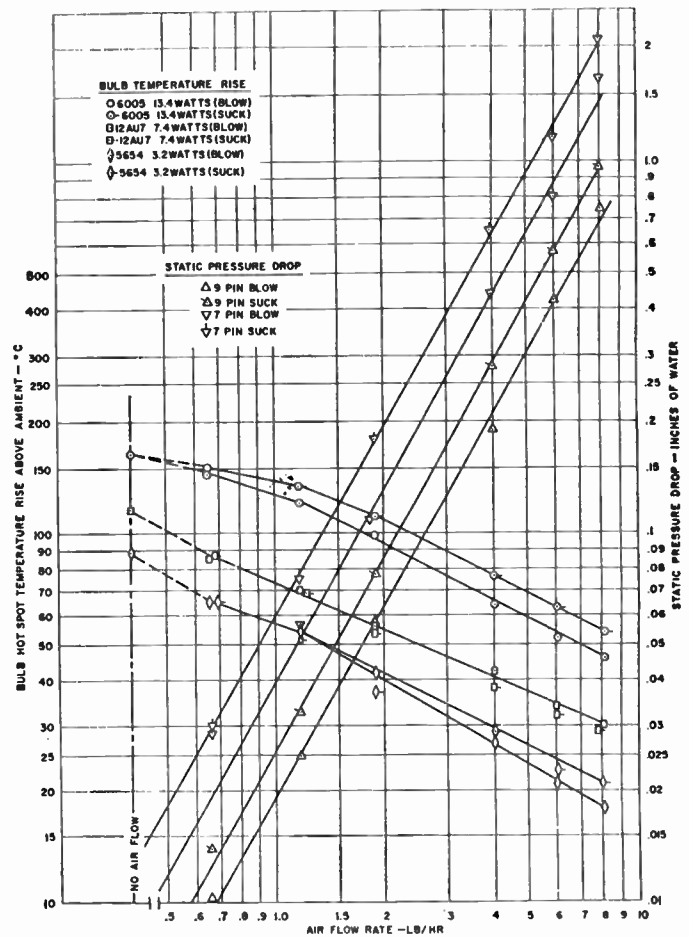


Fig. 3—Sample page from cooling techniques handbook. (Courtesy Collins Radio Corp.)

to be hastened if reliable ground electronic equipment were to result in the near future. Therefore, a contract was let to Cornell Aeronautical Laboratory to prepare a design handbook on methods of cooling Air Force ground electronic equipment. After review at RADC, the final version of this handbook is about to be published.⁵ Just as with the reliability handbook, it is planned to distribute the cooling handbook to engineers and contractors for use as a guide to optimum thermal design. Fig. 3 is a sample page from the cooling handbook. As an example of the increase in reliability which can be achieved through the use of better thermal design, RADC had Cornell Aeronautical Labs perform a thermal analysis of a representative item of ground electronic equipment—in this case, a 200- to 400-mc transmitter receiver. As a result of the analysis, some simple modifications were made to the equipment and life tests begun. Preliminary test data indicate that the mean time to failure of the equipment has been increased from 500 to 1500 hours, a 3 to 1 increase. Such startling results are not expected in all cases. It is the fervent

⁵ J. Welsh, "Handbook of Methods of Cooling Air Force Ground Electronics Equipment," RADC, Griffiss AFB, Rome, N.Y., TR-58-128, ASTIA Doc. No. AD-148907; December, 1958.



hope of the author, however, that the cooling handbook will cause a ground electronic equipment designer to think twice before cutting down on the capacity of the cooling equipment under the mistaken impression that electronic components have the same failure rate at 125°F as they have at 75°F.

Figs. 4 and 5 indicate another area which has virtually been ignored until recently. Fig. 4 depicts the effect of bulb temperature on vacuum tube life (in this case a 6AQ5),⁶ and Fig. 5 indicates the effect of various tube shields on tube bulb temperature.⁷ From Fig. 4 it can be seen that a bulb temperature differential of 41°C is the deciding factor in determining whether 5 or 95 per cent of the tubes survive after 4000 hours of operation. From Fig. 5 it can be readily observed that the JAN type of tube shield, formerly required in most military electronic equipments, serves as an excellent oven to cook the tubes and shorten their lives. From studies per-

⁶ "Study of Effects of Operation and Environment on Electron-Tube Life," General Electric Co., Owensboro, Ky., Rep. on Signal Corps Contract DA-36-039-SC-2524; 1956.

⁷ J. McAdam, "NEL JAN-shield insert for reducing tube temperatures," in "NEL Reliability Design Handbook," U. S. Naval Electronics Lab., San Diego, Calif., sec. 502; October, 1955.

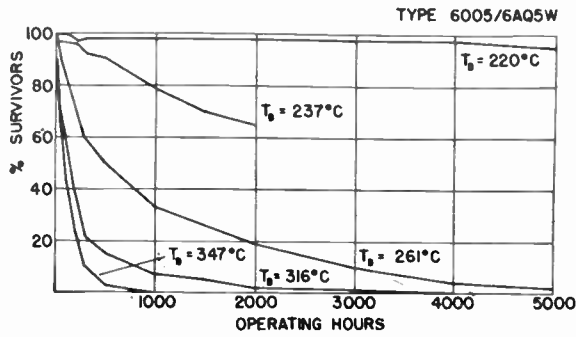


Fig. 4—Effect of envelope temperature on vacuum tube life.

formed at RADC⁸ and other organizations,⁹ it was found that, in some cases, tube life was increased by a factor of 12 simply by lowering tube operating temperature through use of heat dissipating tube shields. As a result of these studies, it is now required that all RADC equipments in development and production be supplied with heat dissipating tube shields rather than the JAN type of shield. Attempts are also being made to replace JAN shields in equipments now in the field.

Redundancy

Since the advent of the reliability problem, many reliability designers have been attempting to determine the optimum method of enhancing equipment reliability. Should it be by improving the reliability of electron tubes and components, or by replacing the tube with transistors, or by using the presently unreliable components in redundant or parallel circuits? It is now agreed by many designers who have looked into the problem that redundancy is the most economical approach.¹⁰ However, although many theoretical analyses of the power of this tool have been made, there has been, to date, little factual data to support the theoretical conclusions. Personnel of RADC have also contributed to this area mainly from a theoretical viewpoint.¹¹ During the next year, however, it is planned to put the theories to test by designing various configurations of redundant circuits and testing them under dif-

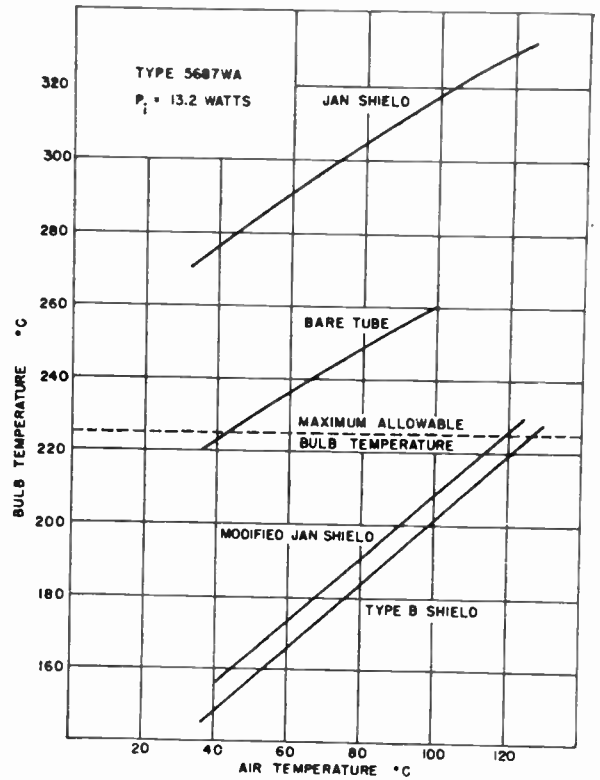


Fig. 5—Effect of tube shields on envelope temperature.

fering conditions in an attempt to verify the theoretical results.

Solid-State Circuits

Realizing that the devices developed by the solid-state physicists show promise of having great inherent reliability, RADC has let a contract to Vitro Corp. to look into this area.¹² The intent of the study is to attempt to predict those electronic functions which it is felt ground electronic equipment will be required to perform in the 1963–1965 era. Once this has been accomplished, a series of solid-state, functional, electronic building blocks of known reliability will be designed and developed. RADC engineers and contractors will be required to use these building blocks as the basic elements of future reliable electronic systems as much as they now use as their basic elements, resistors, capacitors, transformers, etc. It is hoped that these new elements will free the designer from the necessity of designing basic circuits and allow him to concentrate on how these circuits can best be put together to create reliable systems. This program should also greatly alleviate the present logistics problem in which, for example, components must be stocked for as many as twenty different IF strips, each of which has essentially the same functional characteristics but differing circuit design. The other military services are pursuing similar programs in this

⁸ A. Coppola, "Extending Electron-Tube Life with Tube Shields," RADC, Griffiss AFB, Rome, N. Y., TR-58-64, ASTIA Doc. No. AD-148716; July, 1958.

⁹ W. D. Campbell and R. B. Jordan, "Heat Dissipation in Electron Tubes," presented at Second RETMA Symp. on Appl. Relativity, Syracuse, N. Y.; June, 1957.

P. Klass, "New shield gets rid of tube heat fast," *Aviation Week*, vol. 59, pp. 62–66; December 14, 1953.

¹⁰ J. S. Chin, "Optimum design for reliability," *Sperry Eng. Rev.*, vol. 2, pp. 16–21; March, 1958.

C. Shannon and F. S. Moore, "Reliable circuits for unreliable relays," *J. Franklin Inst.*, vol. 262, pp. 1330–1335; September–October, 1956.

¹¹ F. Moskowitz and J. McLean, "Some Reliability Aspects of Systems Design," RADC, Griffiss AFB, Rome, N. Y., RADC-TN-55-4; 1955.

F. Moskowitz, "The Analysis of Redundancy Networks," RADC, Griffiss AFB, Rome, N. Y., RADC-TN-58-42, ASTIA Doc. No. AD-148588; February, 1958.

D. J. Kenneally, "Notes on the Effect of Redundancy as Applied to an RC Coupled Audio Amplifier," RADC, Griffiss AFB, Rome, N. Y., Gen. Engrg. Lab. Rep. No. SG-2; 1955.

¹² "Solid-State Circuits for Earth Electronics in the Space Era," Vitro Corp., Silver Spring, Md., and RADC, Griffiss AFB, Rome, N. Y. Engrg. Rep. on Contract No. AF 30(602)-1906; 1958.

area, and these are being monitored for application to the RADC program.¹³

One of the dangers of such a program is that the urge for standardization could become so strong that it stifled newer, more creative designs. It is anticipated that this can be avoided by constant review of the building blocks, as well as by permission to use other circuits if it can be shown that they offer advantages in performance and at least equivalent reliability. The goal of this development is not so much one of standardization as it is one of the achievement of some acceptable and known minimum reliability in electronic equipment.

RELIABILITY PREDICTION TECHNIQUE

The development of a valid technique for predicting the quantitative reliability to be achieved by an electronic circuit and/or equipment, would serve a three-fold purpose:

- 1) When quantitative reliability requirements become part of equipment specifications, it would permit evaluation of a contractor's performance during design stages to determine whether the completed design will meet the specified reliability figure.
- 2) It would permit design of circuits and equipments of known predictable reliability.
- 3) In those cases where only a limited number of test samples are available or testing time is limited, it could be used to aid in determining equipment reliability.

Basis for Prediction

If one considers an item of electronic equipment consisting of *n* components each of which has some known reliability *R*, the expression for the over-all reliability of the equipment will be given by

$$\begin{aligned}
 R_{\text{over-all}} &= R_1 R_2 \cdots R_n = \prod_{i=1}^n R_i \\
 &= e^{-\lambda_1 t} e^{-\lambda_2 t} \cdots e^{-\lambda_n t} \\
 &= \exp - \sum_{i=1}^n \lambda_i t = e^{-t/m}, \quad (3)
 \end{aligned}$$

where λ_i is the failure rate of the *i*th component. Therefore, the over-all equipment failure rate is simply the sum of the individual component failure rates. The mean time to failure (*m*) is equal to $1/\sum\lambda_i$ and is measured as shown in (2). This is true providing the following assumptions are made:

- 1) All components are said to be in series, so that if one fails the equipment is considered to have failed.
- 2) The failure of one element does not affect the future probability of failure of any other elements.

¹³ "General Usage Assemblies for Navy Electronic Equipment," Vitro Corp., Silver Spring, Md., and Navy Dept., Bureau of Ships, Washington, D. C., Engrg. Reps. on Contract No. NOBsr-72538 (1718); May 26, 1958.

"Microminiature Module Program," RCA, Camden, N. J., and USASRD, Fort Monmouth, N. J., Engrg. Reps. on Contract No. DA36-039 SC-75968.

J. Muncy, "Standardization of Electronic Circuits," Navy Dept., Bur. of Aeronautics, Washington, D. C., NBS Contracts No. NAcr-01435, 01622, and 01835; 1955.

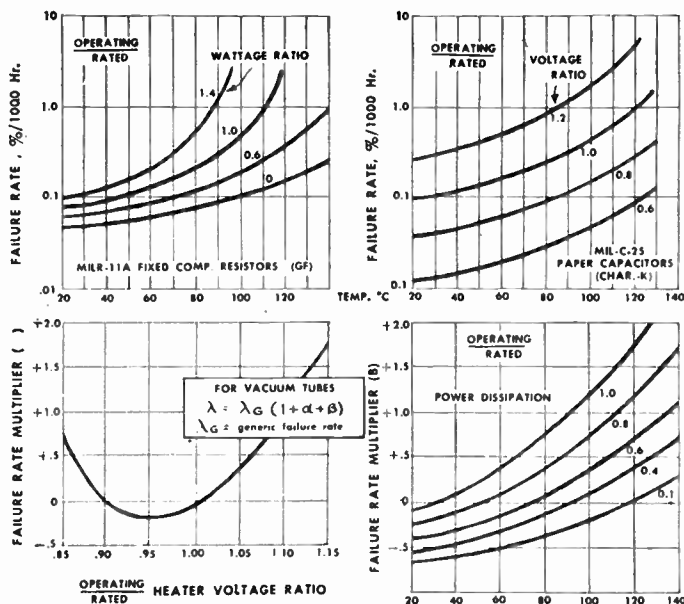


Fig. 6—Failure rate vs stress curves.

- 3) The equipment is in the constant failure rate portion of its life curve.

It would appear, if one accepts these assumptions, that to make a reliability prediction on an item of electronic equipment, one need only know the λ 's or failure rates of each component used. Adding the λ 's would provide the λ_T or equipment failure rate, and $1/\lambda_T$ would provide the mean time to failure, a measurable quantity.

Because there are many types of stresses affecting the failure rate of each component, one would have to search for those stresses having the most significant effect on each component and measure their quantitative effect. If this were done and enough component-failure data gathered, one would arrive at a series of curves for each component, indicating the quantitative effects that the most significant stresses have on the component failure rates. A sample of such a set of curves is shown in Fig. 6 for some types of resistors, capacitors, and vacuum tubes.¹⁴ Armed with a complete set of such curves for each component used in his equipment, a designer could then measure the magnitude of the more significant stresses on each component (temperature and rated operating wattage for resistors and temperature and rated operating voltage for capacitors) and arrive at more realistic and meaningful failure rate figures for each component and hence for the total equipment.

Therefore, the mathematical expression for an equipment failure rate, assuming *n* parts, is given by

$$\lambda_T = \sum_{i=1}^n (\lambda_p + \lambda_m + \lambda_e)_i \quad (4)$$

¹⁴ H. Wuerffel, et al., "Reliability Stress Analysis for Electronic Equipment," Defense Electronic Products, RCA, Camden, N. J., TR-1100; November, 1956.

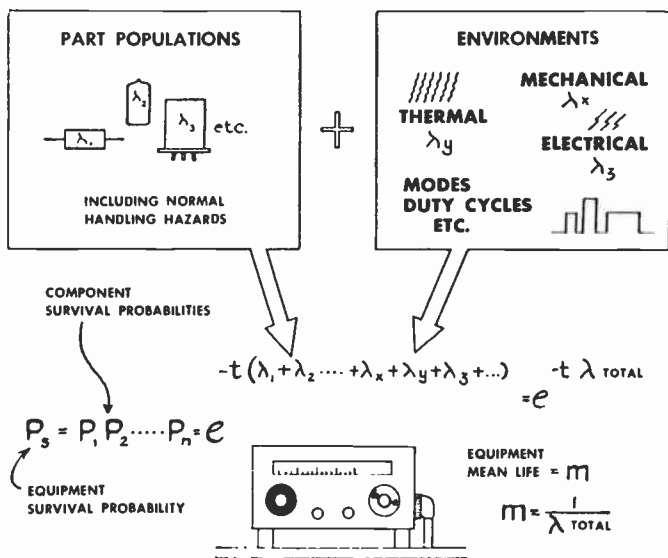


Fig. 7—Basis for failure prediction.

where λ_T is the composite failure rate of the equipment functioning under specified circuit and environmental conditions, λ_p is the part failure rate in terms of a nominal or quiescent failure rate, λ_m is the change in failure rates due to circuit application (electrical stress), and λ_e is the change in failure rates due to environmental (primarily thermal) effects. This is graphically shown in Fig. 7.

It has been shown that if under certain conditions a designer has data on the basic failure rates of each of his components and has a knowledge of the quantitative effects of the more significant stresses (electrical and environmental) to which the components are subjected, he should be able to predict with a fair degree of accuracy the total equipment failure rate and, hence, the mean time to failure of his equipment.

Prediction Program

During 1957, RADC contracted with RCA to apply the previously mentioned prediction technique to several representative items of ground electronic equipment to compare the predicted figure with those obtained from laboratory and field tests of the equipment, and to recommend methods of refining and improving the prediction technique if required.¹⁵

The equipments chosen were the AN/GRC-27, a UHF transmitter-receiver containing 1594 component parts; AN/FPS-3, a radar set operating in the 1300-mc region and containing 3766 component parts; and the AN/GPX-20, an IFF equipment operating in the 1000-mc region and containing 542 component parts. Component parts refer to nonredundant parts since the prediction technique assumes a series dependency relationship of components, and obvious redundancies are culled out before the technique is applied.

From the latest available data on component rates vs stresses, a thorough reliability analysis was performed

on each of the equipments and a predicted reliability figure in terms of mean time to failure was obtained. The analysis included measurement of electrical and environmental stresses on each component in each circuit. The subsequent premises were made and followed during the prediction;

1) The equipment was assumed to have passed through the early failure rate portion of its life.

2) All parts used, particularly the nonstandard parts, had failure rates normally associated with high-quality parts of similar types.

3) Only catastrophic (chance or true random) failures that occurred during normal equipment operation were considered.

4) Special parts with relatively short operation times to wear out were to be replaced in accordance with normal maintenance practices associated with the particular type of equipment and specifically as cited for particular motors, etc. (Such replacements were not included in the reliability calculation.)

5) Portions of the equipment where part misapplications and questionable circuit operation were detected, would be modified in accordance with recommendations included as part of the reliability analysis. Failure-rate values for misapplied or overstressed parts would be used on the modified design.

6) Only electronic, electrical and electromechanical parts would be considered. (Past experience had shown that approximately 90 per cent or more of the catastrophic failures were caused by such parts.¹⁴) Mechanical parts were more likely to exhibit gradual deterioration permitting replacement during normal maintenance periods.

7) The equipment was required to be operated continuously except for breakdowns and scheduled maintenance shutdown periods. In the case of the AN/GRC-27 radio set, it was felt that a more realistic simulation of actual operational conditions would occur with the transmitter on 20 per cent of the time and receiver on 80 per cent of the time.

8) Vacuum tube replacements were assumed to be of the reliable type.

Little satisfaction can be derived from predicting equipment reliability unless failure data accumulated during laboratory and field operation can be obtained to permit correlation studies. These studies provide the required feedback loop necessary for refinement of the prediction technique. RADC supplied the facilities and manpower for life-testing the equipments to obtain laboratory reliability figures, and RCA provided the manpower for performing the field reliability analysis.

Results, Comparisons and By-Products of Study¹⁶

The results of this study to date are summarized in Table I. It can be seen that correlation between predicted, laboratory and field figures is good.

¹⁵ "Reliability Prediction and Measurement Techniques Study," RCA Service Co., and RADC, Griffiss AFB, Rome, N. Y., Engrg. Repts. on Contract No. AF 30(602)-1623; 1957-1958.

¹⁶ J. Naresky, "Reliability Prediction and Test Results on USAF Ground Electronic Equipment," RADC, Griffiss AFB, Rome, N. Y., RADC-TN-58-177, ASTIA Doc. No. AD-148794; July, 1958.

TABLE I

MEAN LIFE OF AIR FORCE GROUND ELECTRONIC EQUIPMENT

Equipment	Complexity (electrical parts)	Prediction	Lab test to July 1, 1958	Field measurement to July, 1958
AN/FPS-3	3776	74 hrs.	83 hrs.	55 hrs.
AN/GPX-20	540	425 hrs.	540 hrs.	337 hrs.
AN/GRC-27	1594	399 hrs.	434 hrs.	390 hrs.

For a more thorough description of the program and discussion of the results, the interested reader is referred to the reports published concerning the work.¹⁷ Efforts are now being concentrated on refining the technique to increase its validity and accuracy. A report describing the technique and its method of application has been disseminated to RADC engineers and contractors for their use.¹⁸

The performance of a thorough reliability prediction and analysis on an item of complex equipment is a time-consuming process. It has been estimated that a rigorous reliability analysis of 25 electron tube circuits consumes one engineering man month.¹⁴ The author felt that a requirement existed for a prediction technique which, although less accurate than the rigorous one, could be performed rapidly and would supply "ballpark" reliability estimates. Such a technique was developed using the failure results of this study and based upon the total number of vacuum tubes an equipment contains.¹⁹ With a knowledge of tube count a designer of an item of ground electronic equipment can estimate field reliability from the empirically derived relationship,

$$\text{mean life (hours)} = \frac{1.8 \times 10^4}{N} \quad (5)$$

where N is the total number of vacuum tubes. The correlation between ballpark and field figures has been close in most cases. Figures for equipments used in this study are

	Ballpark figure (hours)	Field figure (hours)
AN/FPS-3	55	55
AN/GPX-20	392	337
AN/GRC-27	225	390

¹⁷ "A Prediction of AN/GRC-27 Reliability," RCA Service Co. and RADC, Griffiss AFB, Rome, N. Y., RADC-TN-58-18, ASTIA Doc. No. AD-148554; August 26, 1957.

"A Prediction of AN/FPS-3 Reliability," RCA Service Co. and RADC, Griffiss AFB, Rome, N. Y., RADC-TN-58-19, ASTIA Doc. No. AD-148555; October 1, 1957.

"A Prediction of AN/GPX-20 Reliability," RCA Service Co. and RADC, Griffiss AFB, Rome, N. Y., RADC-TN-58-30, ASTIA Doc. No. AD-148562; December 20, 1957.

"The Prediction and Measurement of Air Force Ground Electronic Equipment Reliability," RCA Service Co. and RADC, Griffiss AFB, Rome, N. Y., Final Engrg. Rep. on Contract No. AF 30(602)-1623; August 15, 1958.

¹⁸ "Philosophy and Guidelines for Reliability Prediction of Ground Electronic Equipment," RCA Service Co. and RADC Griffiss, AFB, Rome, N. Y., RADC-TN-58-20, ASTIA Doc. No. AD-148556; October 15, 1957.

¹⁹ J. Naresky, "A Technique for Estimating Ballpark Reliability Figures by Tube Counting," RADC, Griffiss AFB, Rome, N. Y., RADC-TN-58-81, ASTIA Doc. No. AD-148647; March, 1958.

The ballpark figure has been extremely useful in estimating reliability of systems in the planning stages, in estimating reliability of systems in the field on which reliability data are meager, in determining reliability goals for equipment specifications, and in other uses where rapid estimates are required and some inaccuracy can be tolerated.

RELIABILITY SPECIFICATION AND MEASUREMENT

It is a well-known fact that unless and until reliability can be made a quantitative specification requirement capable of measurement, much of the work on design techniques and handbooks is virtually useless. A contractor faced with the two specification statements, one of which requires, "the power output of the transmitter shall be no less than 10 megawatts" and the other of which declares, "the equipment shall be made as reliable as possible, consistent with time and money limitations," will concentrate his efforts on meeting the first requirement because it is specific and capable of measurement.

The military services have for some time been guilty of inserting "I love babies" type of qualitative reliability requirements into equipment specifications and then complaining because the contractor expended little effort on reliability design.

RADC has been striving to develop quantitative reliability figures for ground electronic equipment specifications. As a result of the prediction technique study and other studies such as the AGREE report,²⁰ this has finally been accomplished. Two reliability specifications have been prepared for ground electronic equipment: one for development models, and the other for production models.²¹ The specifications contain quantitative reliability requirements related to equipment complexity and the test procedure for determining whether the requirements have been met. Examples of some of the specifications paragraphs are shown below for development equipments:

3.2 MINIMUM MEAN-TIME-BETWEEN-FAILURES (MTBF)—The mean operating time between independent failures shall be not less than that specified in the detailed equipment exhibit. When not specified therein, the mean operating time between independent failures shall be not less than that given by the following formula:

$$\text{MTBF (in hours)} = 1/F,$$

where

- $F_1 = 30 \times 10^{-6} \times N_t + 15 \times 10^{-6} N_m + 2 \times 10^{-6} \times N_s + 0.5 \times 10^{-6} N_c$
- N_t = Total number of tubes (envelopes) included in the equipment parts complement,
- N_m = Total number of motors and relays included in the equipment parts complements,
- N_s = Total number of semiconductors (transistors, diodes, et cetera) included in the equipment parts complement,
- N_c = Total number of remaining electrical and electrical mechanical parts included in the equipment parts complement.

²⁰ Advisory Group on Reliability of Electronic Equipment (AGREE), "Reliability of Military Electronic Equipment," Office of the Asst. Secy. of Defense (Research and Engineering), Washington, D. C.; June, 1957.

²¹ "Reliability Requirements for Ground Electronic Equipment (Development and Service Test)," RADC, Griffiss AFB, Rome, N. Y., Exhibit No. 2629; October 31, 1958.

"Reliability Requirements for Ground Electronic Equipment (Preproduction and Production)," RADC Griffiss, AFB, Rome, N. Y., Exhibit No. 2623; October 31, 1958.

3.3.3 PRELIMINARY RELIABILITY ESTIMATE—Not later than 30 days after contract award, or at such other time as is approved by the procuring activity, the contractor shall prepare and submit to the procuring activity a statement of the indicated reliability of the equipment. This statement shall include but not necessarily be limited to the following:

a. The estimated or predicted reliability of the total final equipment configuration. This estimate shall be based on the failure rates of 3.2 except for the parts whose failure rates must be derived from special tests. The method of presentation shall include a reliability block diagram which is broken down by functional divisions in the manner of an electrical block diagram and which indicates, for each block, the parts complement and failure rates per category listed in 3.2. The estimate and, if possible, the reliability block diagram shall clearly indicate the number and total failure rate of parts which comprise support circuitry.

b. When the equipment is capable of operation in more than one mode, separate estimates as in "a" shall be accomplished for each mode.

c. When the estimated MTBF for the over-all equipment is less than 100 hours, or when requested by the procuring activity, the statement shall include a diagnosis of optimum methods and locations for the scope of work required and the degree of reliability improvement to be achieved through redundancy.

4.1 RELIABILITY TESTS—The contractor shall conduct reliability tests in such manner as to demonstrate to an acceptable confidence level (90 per cent) that the required reliability has been provided. Unless otherwise specified in the detailed equipment exhibit, the reliability demonstration test shall be accomplished under normal ambient room temperature and humidity conditions. The conditions of the test and the accept-reject criteria shall conform to that outlined in Table I.

TABLE II
FAILURE RATE TESTING (ACCEPT-REJECT CRITERIA)

Multiples of MTBF (See Note 1)	Accumulated Number of Failures (See Note 2)		
	Rejection (See Note 3)	Continuation of Testing (See Note 4)	Acceptance (See Note 5)
3.00	8	2-7	1
3.32	8	2-7	1
3.58	9	3-8	2
4.01	9	3-8	2
4.27	10	4-9	3
4.70	10	4-9	3
4.96	11	5-10	4
5.39	11	5-10	4
5.65	12	6-11	5
6.08	12	6-11	5
6.34	13	7-12	6
6.77	13	7-12	6
7.03	14	8-13	7
7.46	14	8-13	7
7.72	15	9-14	8
8.15	15	9-14	8
8.41	15	10-14	9
9.10	15	11-14	10
9.79	15	12-14	11
10.30	15	—	14

NOTE 1: Column 1 indicates the ratio of total operating time (accumulated by all equipments under test) divided by MTBF. Required MTBF shall be as specified in 3.2.

NOTE 2: Failures listed are total independent failures experienced by all equipments under test.

NOTE 3: Rejection shall occur if the number of failures indicated occur on or before the time indicated in Column 1.

NOTE 4: Testing will be continued if the number of failures fall within the range indicated at the time indicated in Column 1. (See 4.1.1 for exception.)

NOTE 5: Acceptance shall occur if no more than the number of failures indicated occur by the time indicated in Column 1.

It is felt that these specifications represent a significant step toward the goal of achieving increased reliability in ground electronic equipment.

To date RADC has made no attempt to develop new reliability measurement techniques; it has applied those described in the literature.²² One of the first needs for

²² "Techniques for Reliability Measurement and Prediction Based on Field Failure Data," Vitro Labs., Silver Spring, Md., TR No. 80; October 10, 1955.

"Concepts and Tentative Techniques for Reliability Assurance," Aeronautical Radio, Inc., Washington, D. C., Air Force Reliability Assurance Program Progress Rep. No. 1; February 15, 1956.

D. A. Hill and H. D. Voegtlen, "Time-Samples Measure Equipment Performance," Hughes Aircraft Co., Culver City, Calif., Rep. No. R-56-7; October 10, 1956.

measurement was to establish the current state of the art in reliability of ground electronic equipment. How could reliability be improved until it was known how bad or good it was? Also, could some reliability index be established for ground electronic equipment which was independent of equipment complexity? Several studies were performed in the measurement field, and a fairly large amount of field failure data was acquired.²³

Using the failure rate information, it was possible to determine a reliability index for ground electronic equipment, a measure of the reliability state of the art of ground electronic equipment. This index, in terms of average component failure rate per hour, is a convenient measure of the state of the art of reliability design since it is independent of equipment complexity. Once obtained, such an index can be compared with the state of the art in other areas and the magnitude of future work estimated. The reliability index was computed in terms of average failure rate per component:

$$\text{reliability index} = \frac{(\text{total number of component failures})}{(\text{total number of components})(\text{total time of operation})}$$

Computations of the index were made for true and replacement failures for each equipment under study; replacement failures include secondary failures and preventive maintenance replacements. From the figures for each equipment, an average index was computed for ground electronic equipment for true and replacement failure cases. Reliability indexes obtained were

$$RI_1 (\text{true failures}) = 3.7 \times 10^{-6} \frac{\text{failures}}{\text{component} - \text{hour}} = \text{average failure rate/component,}$$

$$RI_2 (\text{replacement failures}) = 11.3 \times 10^{-6} \frac{\text{failures}}{\text{component} - \text{hour}} = \text{average failure rate/component.}$$

RI_1 and RI_2 were then plotted on a curve of total number of components vs mean life and compared with reliability indexes of other equipment (Fig. 8). The curve fell approximately where expected, thus indicating that the present state of the art of ground electronic equipment reliability design is not significantly lagging behind that of other electronic equipment areas. However, from observation of Fig. 8, it is obvious that there is a great deal of room for improvement.

²³ "Establishment of Methods and Procedures of Testing for Reliability in Ground Electronic Equipment," RCA Service Co. and RADC, Griffiss AFB, Rome, N. Y., TN-58-70, ASTIA Doc. No. AD-1486527; December 1, 1957.

"Methods of Field Data Acquisition, Reduction and Analysis for Ground Electronic Equipment Reliability Measurement," RCA Service Co. and RADC, Griffiss AFB, Rome, N. Y., TN-58-183, ASTIA Doc. No. 148801; May 1, 1958.

"Missile Guidance Equipment Reliability Study," RADC, Griffiss AFB, Rome, N. Y., and Fairchild Guided Missiles Div., Wyandanch N. Y., Final Engrg. Rep. on Contract No. AF 30(602)-1553; June, 1957.

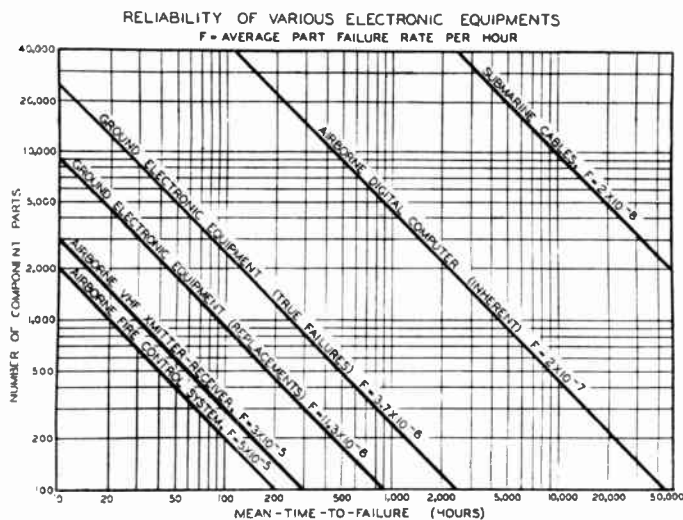


Fig. 8—SOAR curves based upon average failure rate per component.

Another interesting fact revealed by the reliability index data is that the average replacement rate is approximately three times the true average failure rate in the field. As additional data of this type become available, more scientific procedures should result for a determination of spare parts requirements, a method of evaluating maintenance training, and, coupled with mean life figures, a means for optimizing preventive maintenance procedures.

The USAF field failure reporting system, by virtue of sample size, is potentially an excellent source for component and equipment failure rate data. This potential, however, is not now being realized because of the following:

- 1) not all failures are reported,
- 2) not all types of equipment are included in the list of those to be reported,
- 3) no indication of total population is available,
- 4) equipment running time is usually not reported,
- 5) some common failures cannot be listed, and
- 6) the data are slow in arriving and are, in some cases, incorrectly processed.

Studies are under way to correct these deficiencies so that the failure data can be used in reliability analyses. To be effective any reliability program requires a good feedback system; the failure reports could represent that feedback. Not only could it provide the means for verifying reliability prediction figures or measured figures on limited time samples, it could also be used as a tool for rapid elimination of design deficiencies which might have escaped previous detection. The designer would have a timely running account of his equipment's reliability in the field.

Another foreseeable problem in achieving more reliable ground electronic equipment is that of attempting to verify specified reliability on a small number of equipments which can be life tested for only a short time. This is particularly true of development models for which economy dictates that no more than one or two

models be purchased. A valid prediction technique will help alleviate this difficulty; however, it is no panacea. The requirement still exists for developing some type of accelerated reliability test for electronic equipment which will accelerate the failures occurring per unit time without changing the normal long-time failure pattern of the equipment. RADC personnel are looking into this problem to determine if a valid test of this type can be developed for ground electronic equipment. It is too soon to list definite results.

CONCLUSIONS AND FUTURE PLANS

The history and current status of the RADC Reliability Program have been described. Our future reliability program will probably be modified many times to take into account new developments in this rapidly expanding field.

As with any new area of scientific investigation, each new advance has opened to investigators a vista of unsolved problems. The author would now like to pose some of the problems encountered at RADC and probably other organizations working in this field.

Concerning the prediction of reliability, is the product rule equation (3) too much of a simplification? Would the development of a system of weighting failures be more realistic, the weighting factor being proportional to the amount of degradation in system performance caused by the component failure? If one assumes that present design and prediction techniques, requiring a knowledge of component failure rates vs stresses, are valid, how can the required component failure rate data be most economically and rapidly gathered and disseminated to designers? Is an astronomical test program the only solution, or might a basic study in material failure mechanisms provide the answers faster? For example, what is the average failure rate for solid-state components such as transistors?

Reliability, despite its importance, is not the ultimate measure of the operational effectiveness of electronic equipment. There is no point in designing an equipment which fails once a year if, when it fails, as much as a month might be required for repairs. Thus, maintainability, or decreasing equipment maintenance time, must become another goal of military electronic equipment design. As with reliability, techniques must be developed for quantitatively specifying, measuring, and predicting equipment maintainability as well as the development of design techniques which will improve maintainability. This area promises to provide more difficulty than that of reliability since one must take into account the human variable as well as the machine design in any analysis of the system.

Therefore, although great strides have been recently made at RADC and other organizations toward the conversion of reliability from a black art, there is a great deal to be accomplished before it can be termed a science.

If the reader is left with one conclusion, it is hoped that it will be the realization that tools, still crude, have

been developed to specify quantitatively the reliability of ground electronic equipment, to measure it, and to predict it. It is the plan of the author to apply these tools in helping to develop Air Force ground electronic equipment of increased reliability.

APPENDIX MATHEMATICS OF RELIABILITY

If

N = the number of components at the start of a test,

S = the number of survivors,

λ = the failure rate,

R = the reliability,

Z = the hazard, and

F = the number of failures,

then

$$S = N - F.$$

Reliability is obtained directly from the fraction of the initial components which survive at the end of each hour,

$$R = \frac{S}{N}.$$

Failure rate, the rate of decrease in the reliability, can be determined from the reliability as

$$\lambda = \frac{dR}{dt} = -\frac{1}{N} \frac{dF}{dt}.$$

Hazard is the ratio of the number of failures per hour to the number of survivors at that time, or

$$Z = \frac{1}{S} \frac{dF}{dt} = -\frac{1}{R} \frac{dR}{dt}.$$

Integrating this equation gives an important relation between the reliability and the hazard,

$$R = \exp - \int_0^t Z dt.$$

If the hazard is constant, as is the case with the middle portion of the life curves of most electronic equipments,

$$R = e^{-Z_0 t} = e^{-t/m}.$$

This is the standard reliability equation where m is the mean time to failure, and

$$m = \frac{1}{Z}.$$

When F is small (as is usually the case),

$$\lambda(\text{failure rate}) = Z_0 = \frac{1}{m}.$$

Special Electronic Equipment for the Analysis of Statistical Data*

E. R. CARLSON†, C. C. CONGER†, J. C. LAURENCE†, E. H. MEYN†, AND R. A. YOCKE†

Summary—Four instruments useful in the mechanized analysis of noise signals are described. They are a noise data digitizer, a sound field contour plotter, a probability analyzer, and a correlation computer.

A description of analytical and experimental approaches to the extraction of useful information from stored statistical data is included.

INTRODUCTION

THE advent of the jet airplane has intensified interest in the origin of aircraft noise and its elimination or reduction. To solve the problem,

considerable research¹⁻⁴ has been undertaken. As in other noise problems, the phenomena of interest cannot be specified as of a particular value at a particular time. Rather the magnitude of the signal at any time is a random function of time. These random signals have been analyzed in two ways:

¹ M. J. Lighthill, "On sound generated aerodynamically. I—General theory," *Proc. Roy. Soc. (London) A*, vol. 211, pp. 564-587; March 20, 1952. "II—Turbulence as a source of sound," *Proc. Roy. Soc. (London) A*, vol. 222, pp. 1-32; February 23, 1954.

² R. Westley and G. M. Lilley, "An Investigation of the Noise Field from a Small Jet and Methods for Its Reduction," The College of Aeronautics, Cranfield, Rep. No. 53; June, 1952.

³ W. L. Howes, E. E. Callaghan, W. D. Coles, and H. R. Mull, "Near Noise Field of a Jet-Engine Exhaust," NACA Rep. No. 1338; 1957. (Supersedes NACA TN's 3763 and 3764.)

⁴ E. E. Callaghan and W. D. Coles, "Far Noise Field of Air Jets and Jet Engines," NACA Rep. No. 1329; 1957. (Supersedes NACA TN's 3590 and 3591.)

* Original manuscript received by the IRE, January 21, 1959.
† Lewis Res. Center, Natl. Aeronautics and Space Admin. (NASA), Cleveland, Ohio.

1) The average or true root-mean-square has been obtained for the over-all frequency range or in specific bands of frequencies (octaves, $\frac{1}{3}$ octave, etc.).

2) Statistical analyses involving space-time correlations, probability density, and/or spectral density are made.

To save time and to increase the reliability of these statistical analyses, several instruments, which are described in this paper, were designed to process the data obtained from the noise measurements at the Lewis Research Center of the NASA. Microphones and hot-wire anemometers reduced the sound pressure fluctuations and velocity fluctuations to electrical signals which in general were stored on magnetic tape for analysis.

In studies of noise and noise reduction problems, the total acoustic power and the sound level spectrum are needed. To obtain this information from the tape recordings sometimes requires laborious computations using a large number of such recordings. The definition of sound pressure level⁵ is

$$SPL = 20 \log \frac{\sqrt{\bar{p}^2}}{p_{ref}} \quad (1)$$

This requires the measurement of the signals in a series of band-pass filters to obtain a display of the sound pressure level in the form of a spectrum.

The total acoustic power is

$$W = \int_S I(\vec{X}) dS = \int_S \frac{\bar{p}^2(\vec{X})}{\rho c} dS \quad (2a)$$

which for practical measurement techniques can be approximated in the far field by

$$W = \sum_{f=1}^{f=m} \sum_{a=1}^{a=n} \frac{k_f \cdot A_s a \bar{p}^2}{\rho c} \quad (2b)$$

where

- SPL = sound pressure level
- \bar{p}^2 = mean-square acoustic pressure
- ρ = density of the ambient medium
- c = velocity of sound in ambient medium
- a = a station on sphere of radius r
- n = total number of stations on a sphere of radius r
- A_s = area of zone on sphere of measurement
- m = number of bands
- k_f = frequency response and other instrument constants
- p_{ref} = 0.0002 dyne/cm²
- f = frequency pass band
- I = sound intensity at \vec{X}
- S = area of sphere.

The instrument to be described reduces the voltage recordings and certain calibration constants to a punched paper tape. This tape supplies the information to a digital computer which calculates the total acoustic power.

NOISE-DATA DIGITIZER

The noise-data digitizer is composed of a data signal section, a word-forming unit, and a control section. These are shown in Fig. 1. All necessary functional units required to convert continuous electrical variations to digital form and to punch this information on paper tape are contained in the data signal section and the word-forming unit. The control section provides a link between the data signal section and the word-forming unit and also determines the operational sequence of the instrument.

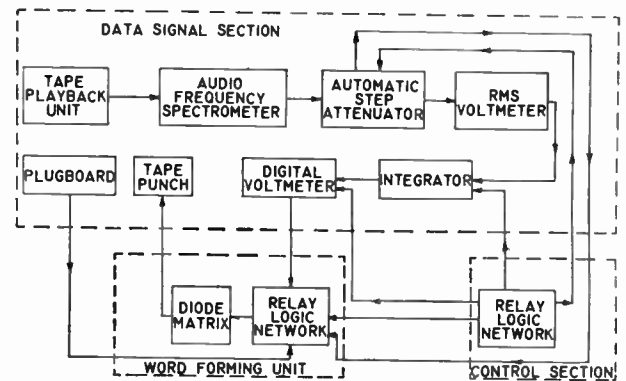


Fig. 1—Block diagram of noise data digitizer.

The electrical variations resulting from the noise measurement are first operated on by an audio frequency spectrometer which uses band-pass filters, automatically switched to separate the signal into prescribed bands. Each filter provides a voltage to be digitized and recorded as a word—a pattern of holes on the paper tape. The number of filters used determines the upper limit for the first summation of (2b). The output voltage of the spectrometer becomes the input to the rms voltmeter after being modified by the automatic step attenuator, which will be discussed later. The rms voltmeter produces a dc output proportional to the square of the input despite the voltage waveform, performing a required mathematical computation as well as contributing to the reduction of the signal to digits. The input voltage is proportional to the sound pressure level actuating the microphone so that the output is proportional to \bar{p}^2 required in (2b). To obtain \bar{p}^2 , however, a time average is required for which an active integration system is used. (See Fig. 2.) The integrator operation is expressed as⁶

⁵ P. M. Morse, "Vibration and Sound," McGraw-Hill Book Co., Inc., New York, N. Y.; 1936.

⁶ W. W. Soroka, "Analog Methods in Computation and Simulation," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 75-81; 1954.

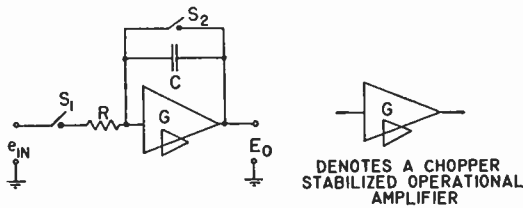


Fig. 2—Integrator.

$$E_0 = \frac{1}{RC} \int_0^T e_{in} dt.$$

The length of time (determined by the control section) that s_1 is held closed sets the upper integration limit T and is longer than the period of the lowest frequency of interest. When s_2 is closed the integrator is reset to zero, reducing drift effects to a minimum and preparing the system for another integration. The output voltage is applied to a standard stepping-switch-type digital voltmeter, in which a level of each decade switch provides contact closures representing the input voltage in decimal form.⁷

After the digital voltmeter has balanced on the integrated voltage, an internally generated readout pulse is sensed by the control section. This then activates the word-forming unit, which presents the information contained in the contact pattern of the digital voltmeter to the diode matrix for conversion to the binary coded decimal form required by the tape punch and initiates a punch operation.⁸

The automatic step-attenuator is a constant input-impedance device composed of a capacitively compensated resistive network in conjunction with a bidirectional stepping switch. It permits the rms voltmeter to operate in its 10-db linear range with a large amplitude variation in input level. The attenuator switch is operated by the control section when the rms voltmeter is not operating in the proper range. With the bidirectional characteristic of the stepping switch the attenuation can be increased or decreased with one switch step.

Two levels of the step attenuator switch provide contact closures describing in decimal form the attenuation in decibels. The plugboard provides means for insertion of other required information, also in the form of contact closures not included in the data. Therefore, the word-forming unit during a punch operation not only obtains information from the digital voltmeter but also from the plugboard and attenuator stepping switch. The function of the word-forming unit is to arrange these various pieces of information in a definite order on the tape.

⁷ T. E. Nawalinski, "Principles and Application of Non-Linear Systems Inc. Stepping Switch Type Digital Voltmeters," Non-Linear Systems Inc., Del Mar, Calif.; 1957.

⁸ R. H. Selzer, "A Preparation Unit for Punched Paper Tape," Jet Propulsion Lab., California Inst. of Technology, Pasadena, Calif., Memo. No. 20-127, pp. 4, 16-17; June, 1956.

The control section is a relay logic network receiving information from various units of the other sections. Many important operations of this section have already been discussed to determine the operational sequence of the instrument.

Each complete cycle of the instrument provides a group of tape words containing digitized noise data from one microphone station in the sound field. The number of these word groups (one for each station) determines the upper limit of the second summation of (2b). The paper tape with information from all stations can then be used by a digital computer to calculate the desired results for the acousticians.

SOUND FIELD CONTOUR PLOTTER

Another parameter of considerable interest is the directivity. It is obtained from contours of constant sound pressure level throughout the jet. Such contours are obtained from surveys of the sound field and cross plotting from (1). An instrument described in this paper does this automatically and plots the result directly on an x - y plotter.

The object of the sound field contour plotter is to reduce automatically the tape recorded data of a sound field in contour form for discrete frequency bands. Fig. 3 is an example of the type of plot produced. Fig. 4 is a photograph showing the airjet and the microphone mounted on the survey boom.

A polar coordinate survey method was adopted by the acousticians. The microphone was moved along a boom oriented at a discrete angle to the centerline of the airjet at a constant velocity. A standard, instrument-quality, direct-record tape recorder was used. The instantaneous intensity as recorded was therefore related to the microphone position as a function of elapsed time from the start of the recording. The data reduction problem now resolves itself essentially into two parts: 1) transferring a polar coordinate system into rectangular coordinates, and 2) measuring the level of the sound data as recorded on the magnetic tape and plotting a system of symbols which will designate discrete sound pressure levels.

Rectangular Coordinate Plotting of a Polar Coordinate System

As already discussed, each survey was made along a radius of fixed length, at constant velocity, and at a discrete angle. The instantaneous positions of the microphone taking the survey are directly related to the elapsed time from the beginning of the recording. Referring to the block diagram (Fig. 5) it is seen that the signal from the tape playback is monitored by the time to radius translator. This consists of a pair of ganged 10-turn potentiometers each supplied by constant voltage and driven by a synchronous motor. The potentiometer drive-motor is started by the signal-sensitive gate. The output voltage of the potentiometers is thus proportional to radius at any instant of time. The output

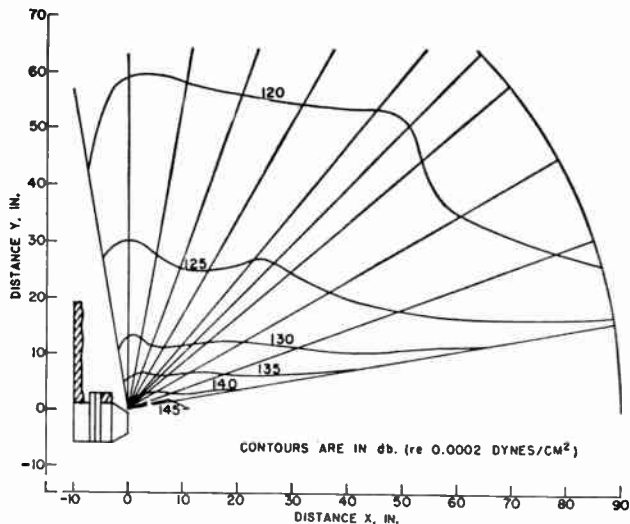


Fig. 3—Sound field contour plot. Over-all sound pressure level for a five-inch diameter nozzle, jet velocity 894 feet/second.

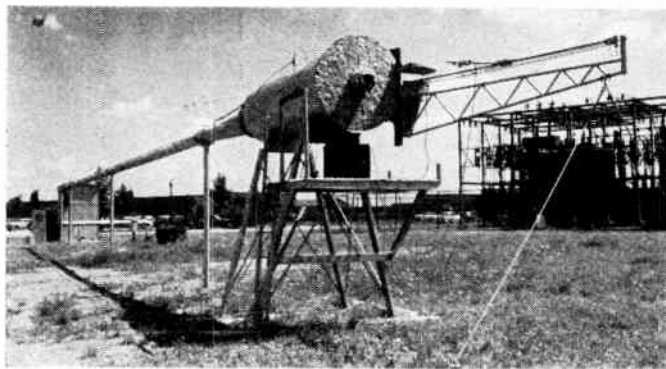


Fig. 4—Sound field probing apparatus.

of each of the motor-driven potentiometers is impressed across a 10-turn potentiometer calibrated from 0–1,000. The setting of each of these corresponds to the sine and cosine of the survey angle \angle as shown in Fig. 3. Thus polar coordinate $r\angle\theta$ is transferred to its equivalent rectangular coordinates $X=r \cos \theta$ and $Y=r \sin \theta$.

Calibration, Measurement, and Plotting

Before the data are recorded, a recording of a calibrated sine wave acoustical source is made. This recording may then also serve as a reference signal for calibrating the contour plotter as long as no change in recording gain is made subsequently. Any gain changes found necessary during recording must be made by a factor of ± 20 db, as will be explained later. The intensity measurements (Fig. 5) are accomplished by feeding the output of the tape recorder whose level is set to recreate the original signal amplitude to the signal sensitive gate which controls the time to radius translator and to the audio frequency spectrometer. The output of the latter is impressed on an rms voltmeter. The step attenuator and rms voltmeter have already been outlined in the first part of this paper.

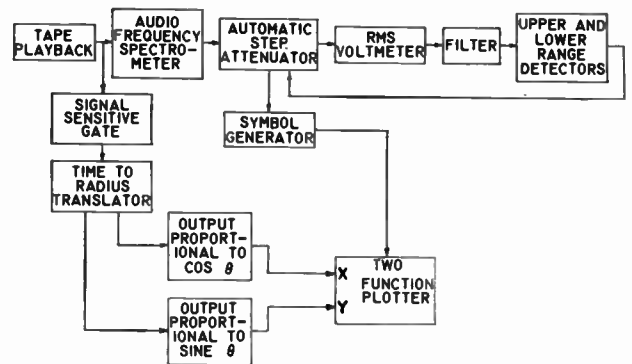


Fig. 5—Block diagram of contour plotter.

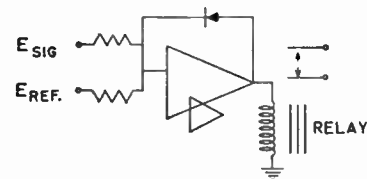


Fig. 6—Level detector.

The rms voltmeter serves a threefold function. It provides a direct means of reading signals in terms of decibels, reads directly in true root-mean-square, and also provides a direct current output proportional to the square of the input signal. For any given range of the rms voltmeter, an input signal which produces 0-db reading provides an output equal to 20 mv, and for 10-db input, an output equal to 200 mv. The contour plotting function is accomplished by utilizing these two voltage levels corresponding to input voltages which are 10 db apart.

As noted on the block diagram, the output of the rms voltmeter is fed to the upper and lower range detectors. The basic function of these detectors is to provide two discrete outputs: one to indicate the level is below 0 db, and the other, the level above 10 db for the particular meter range being used. Fig. 6 shows the basic circuit used for the level detectors. The output of the operational amplifier is equal to the sum of the input voltages times the ratio of feedback resistance to the input resistance. With equal input resistors,

$$E_{out} = (E_{sig} + E_{ref}) \frac{R_{diode}}{R_{input}}$$

The diode used has an impedance ratio of about $10^6/1$, and a Zener voltage of about 50 volts. The Zener voltage serves to prevent saturation of the amplifier. The use of the diode characteristics provides an amplifier, whose gain and output voltage are a function of the polarity of the signals summed at its input. These range between approximately $1/100$ –10,000, with corresponding output voltages of $+\frac{1}{2}$ v minimum to -50 v maximum. With the diode connected as shown, the E_{ref} set at -0.2 v (corresponding to 10 db on the rms voltmeter) and an E_{sig} of $+0.205$ v, the sum would be

+0.05 v, and the output would be negative and in the Zener direction. The gain of the amplifier would thus be approximately 10,000 and the output voltage, 50 volts. This is sufficient to pull up the relay. Any signal exceeding 0.2 v by 0.005 v or more will energize the relay. Signals less than the reference voltage will produce an output limited to $+\frac{1}{2}$ v, by the low forward impedance of the diode, and will not energize the relay. The upper range detector is therefore extremely sensitive to signals above a set reference level.

The lower range detector circuit is the same except that the diode in the feedback loop of the operational amplifier is reversed. The associated relay will therefore energize only when the signal goes *below* the reference voltage level. The reference voltage for this detector is set to 0.02 v corresponding to 0 db on the rms voltmeter. Thus the combination of these two range detectors determines where the incoming signal is in relation to a 10-db wide band, that is below, within, or above.

A selectable set of filters is located between the rms voltmeter and the upper and lower range detector. The purpose of the filters is to average the dc signal voltage provided by the rms voltmeter. This filtering is necessary to prevent instability of the system due to the manner in which the noise data vary.

The symbol generator noted on the diagram has the function of providing a system of symbols denoting levels. Four symbols (L, J, Γ, T) are generated. They are a point to which has been added first a small *X* deflection and then a *Y* deflection. Two levels of contacts on the step attenuator are supplied with ± 4 v. One of these levels determines the polarity of the pulse to be applied to the *x* axis, the other, the polarity of the pulse to the *y* axis. When a point is to be plotted, a system of relays drops the pen of the plotter onto the paper and runs an auxiliary 10-position stepping switch through its 10-position sequence. *X* and *Y* pulses are added to the *X* and *Y* inputs of the plotter, respectively. The symbol plotting sequence takes about 1 second. The symbols which are plotted repeat every 4 symbols.

To summarize, the entire sequence of operations when a particular sound field survey is plotted is as follows. The *xy* plotter sensitivities are first set to produce the desired calibration for the known radius length. These sine and cosine potentiometers are then set for the functions of the survey angle to be plotted. The tape playback, which has been calibrated as described above, is started just ahead of the recorded signal. At the instant the recorded signal appears, the synchronous motor-driven *x-y* potentiometers begin to turn. The upper and lower range detectors will immediately provide the necessary signals to drive the automatic step attenuator to a level such that its output lies within the chosen 10-db range of the rms voltmeter.

A relay logic circuit has been provided which will prevent any point from being plotted if the automatic step attenuator has to step more than once to bring the sig-

nal within 0-10 db range. This is necessary to prevent false points from being plotted when, for instance, the input level to the tape recorder was changed by 20 db to maintain proper recording level. If there has been a gradual change, the automatic step attenuator will step only once. The 10-position stepping switch runs through its sequence dropping the plotter pen. This pen has been moving continuously to correspond to the original microphone position, and a symbol is plotted denoting the new level that the automatic step attenuator has reached. Each time the level changes by 5 db an appropriate symbol will be plotted. It is only necessary that the operator identify one symbol for each survey with its proper db reading. The succeeding symbols will be in sequential order and easily identified if the order of the symbols is known. Each of the contour maps will be made for a particular frequency band selected on the audio frequency spectrometer. The operator, after completing all of the radial plots, has only to connect the appropriate symbols for each constant-pressure contour to complete the map.

PROBABILITY ANALYZER

Also of interest to the acoustician is what kind of noise is being measured. The spectrum level will answer part of this question, but another important concept is that of probability density. The numerical measurements of a noise source may vary widely when the same experiment is repeated several times. However, the relative likelihood of the measurements being the same can be obtained.

Let e represent a typical measured value of the voltage and think of each e as defining a point at a particular distance from a point of reference on a straight line. If this line is divided into equal increments, Δe , and if the number of e 's in each such interval is counted, the probability density function can be approximated by

$$p(e) = \lim_{\substack{\Delta e \rightarrow 0 \\ N \rightarrow \infty}} \frac{\text{number of values in } \Delta e \text{ at } e}{\text{total number of values, } N} \frac{1}{\Delta e} \quad (3)$$

and the probability distribution function by

$$P(e) = \int_{-\infty}^E p(e) de. \quad (4)$$

The design and operation of an instrument to make these calculations will now be described. The basic elements of the analyzer are shown in Fig. 7. To illustrate the basic principle of the instrument, let us rewrite the expression for probability density function as

$$p(e) = \lim_{\substack{\Delta e \rightarrow 0 \\ t \rightarrow \infty}} \frac{(\text{time the voltage is in range } \Delta e \text{ at } e)}{\text{total time}} \frac{1}{\Delta e} .$$

In this case time duration takes the place of a count of points.

Each time interval during which the instantaneous value of the input, e , exceeds $|e_r|$, $+\frac{1}{2} v$ appears at (2). Also, each time interval that e exceeds $|(e_r + \Delta e_r)|$, $-\frac{1}{2} v$ appears at (3). Conversely, each interval during which e is less than $|e_r|$, $-\frac{1}{2} v$ appears at (2), and each time interval during which e is less than $|(e_r + \Delta e_r)|$, $+\frac{1}{2} v$ appears at (3). Hence, during each period of time that the input wave instantaneous voltage lies within Δe_r at e_r , $+\frac{1}{2} v$ at (2) is added to $+\frac{1}{2} v$ at (3) resulting in an electrical output voltage at (4). If the input wave is not within Δe_r at e_r , zero volts appears at (4). If Δe_r is small and the sampling time is long, the average output is proportional to $p(e)$ at amplitude e_r .

The switch S_1 is shown in the position for analyzing probability density functions. The other switch position inserts $-e'$, so that $|(e' + e_r + \Delta e_r)|$ exceeds the maximum value of the input wave e . In this condition the output of the analyzer is proportional to probability distribution function.

In practice the reference voltage e_r is automatically varied from the maximum negative to the maximum positive value of the input wave. The voltage e_r is applied to one channel of a two function plotter while the output of the analyzer serves as the input to the other plotter channel. Thus a plot results of $p(e)$ or $P(e)$ for all values of e_r .

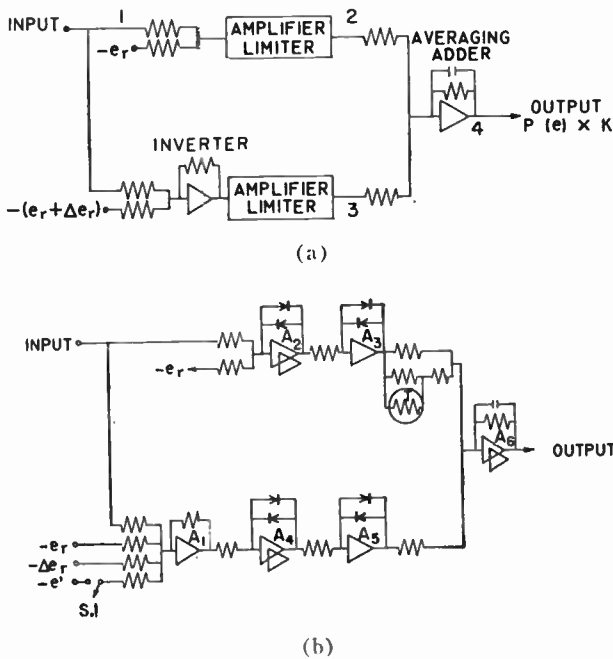


Fig. 7—Probability density analyzer. (a) Modified block diagram. (b) Schematic diagram.

The probability distribution function as defined in (4) is the probability that a value is less than some specified value. If $|(e_r + \Delta e_r)|$ (Fig. 7) were made greater than any value of the input waves, the output would be proportional to $P(e)$ at amplitude e_r .

Fig. 8 is a more detailed schematic diagram. Operational amplifiers are used with A_2 , A_4 , and A_6 chopper stabilized for minimum dc drift. All diodes are silicon. Amplifiers A_2 , A_3 , A_4 , and A_5 function as high-gain amplifiers at low-output voltage levels. The diodes in the feedback path limit each amplifier's maximum output voltage to the diodes' forward conduction voltage, about $\frac{1}{2}$ volt. Diodes associated with A_3 and A_5 were selected for matched conductivity. In addition, to make temperature effects negligible, a thermistor plus trimming resistors in one input channel of A_6 are used. The diodes associated with A_3 and A_5 plus the resistors and thermistor associated with the input of A_6 are potted to minimize the magnitude of temperature gradients.

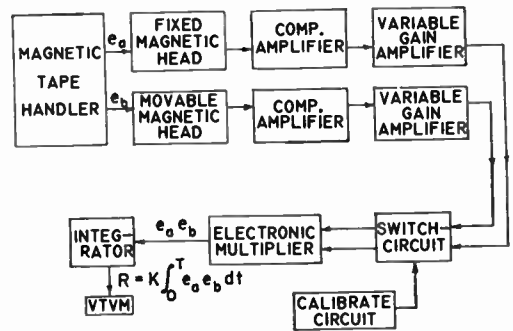


Fig. 8—Block diagram of correlation computer.

A CORRELATION COMPUTER

An alternative to the spectral analysis method is the autocorrelation technique. Let $w(t)$ and $w(t + \tau)$ be two samples of a stationary process, the statistical properties of which depend only on τ . If a large number of measurements on the noise wave are made at t and $t + \tau$, the autocorrelation function is⁹

$$R_w(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T w(t)w(t + \tau)dt. \quad (5)$$

The autocorrelation function can also be obtained by the Fourier transform of the power spectral density function

$$R_w(\tau) = \int_{-\infty}^{\infty} W_w(\omega)e^{i\omega\tau}d\omega \quad (6)$$

and conversely,

$$W_w(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_w(\tau)e^{-i\omega\tau}d\tau. \quad (7)$$

Experimentally, the relation shown in (6) is much easier and more accurately obtained than the transform of (7). To solve (6) an instrument was designed and built.³

⁹ W. R. Bennett, "Methods of solving noise problems," Proc. IRE, vol. 44, pp. 609-638; May, 1956.

In analog form (5) may be written as

$$E_0 = \lim_{T \rightarrow \infty} \frac{K}{T} \int_0^T e_a(t) e_b(t) dt$$

where E_0 is output voltage, T is real time, and e_a and e_b are the instantaneous values of the input voltages whose correlation is desired. If e_a and e_b are from different sources, then the value of E_0 obtained is proportional to the cross correlation of e_a and e_b . If, however,

$$e_b(t) = e_a(t + \tau)$$

where τ is an increment of time considered small with respect to T , then E_0 is proportional to the autocorrelation of e_a and e_b . The device to be described is capable of computing either auto or cross correlations.

By referring to the block diagram (Fig. 8), the operation of the circuit may be described as follows. The two signals whose correlation is to be measured are recorded on a dual channel magnetic tape recorder.

The magnetic tape is played back on a nearly standard tape handler. The exception from standard is that a special playback head arrangement is used. This arrangement will be described later.

In order to reproduce as nearly as possible the signal recorded it is necessary to compensate for the distortions of the signal introduced by the magnetic tape and recording-playback system. This compensation is accomplished by means of a passive filter in the compensating amplifier.

From the compensating amplifier the signals are sent to the variable gain amplifier. The adjustable gain is used to set the scaling coefficient K of (1). The switching circuit (Fig. 8) allows connection of either the data or the calibration signals to the multiplier. The electronic multiplier, which is commercially manufactured, produces an output voltage proportional to the instantaneous product of the input voltages.

To obtain the time average indicated in (8), that is,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t) dt,$$

if low frequencies are involved, it is desirable to use an active rather than a passive filter. Here an analog integrator is used as shown in Fig. 2. The operation has been described previously. The integrator permits a steady readable output indication from the unsteady dc voltage usually obtained from the computer.

When used for autocorrelation measurements, the same data are recorded on each channel of tape. The

computer is adjusted using the variable gain amplifier so that the value of the correlation for each value of τ in (3) may be read directly from the voltmeter.

The calibration circuit (Fig. 8) provides ac and dc voltages. These are switched into the electronic multiplier prior to its use for calibration and adjustment.

The compensating amplifier design is one of the most critical in the computer. Desirable features are: low noise, low phase shift over the usable band, and high gain. The introduction of the head compensating filter in the early stages of the amplifier prevents saturation in the later stages. The filter is passive, adjustable, and of plug-in construction.

The instrument has been provided with facilities for automatic plotting of the autocorrelation function. It is possible to compute the autocorrelation coefficient R from

$$R = \frac{1}{\{[e_a(t)]^2 [e_a(t + \tau)]^2\}^{1/2}} \frac{1}{T} \int_0^T e_a(t) e_a(t + \tau) dt \quad (9)$$

for various values of τ by allowing the drive motor to vary head gap separation x according to

$$\tau = \frac{S_t}{x}$$

where S_t is the tape speed. The head position is measured by a linear, infinite resolution potentiometer attached to the movable head mechanism. When automatic plotting is used, the active filter is replaced by a low-pass passive filter so that R may be obtained continuously. The output of the potentiometer is applied to one input of an x - y recorder. The other input is obtained from the computer output.

Time delays of the order of 50 msec are required for acoustic correlation measurements. To achieve the required accuracy of time delay a high degree of precision is required. Tolerances of ± 0.00025 inch were achieved in the final construction of the head mechanism.

CONCLUSION

1) A system of instrumentation has been described which treats statistically the type of data of interest to acousticians studying aircraft noise.

2) The instrumentation reduces the labor of processing this type of data by analog and analog-digital methods which are described.

3) The instrumentation has had extensive use and has proved to be a practical engineering solution to the data reduction problem.

Hybrid Velocity Data for the Velocity Measuring System of the Supersonic Naval Ordnance Research Track*

R. J. STIRTON† AND B. GLATT†, ASSOCIATE MEMBER, IRE

Summary—A system is described for measuring the velocity of test sleds on the Supersonic Naval Ordnance Research Track (SNORT) with an rms error of less than 0.1 ft/second over a velocity range of 200 to 2000 ft/second and a bandwidth of 50 cps. Instrumentation consists of two data sources: sled-position vs time is measured with a magnetic-pickup track-coil system, and sled acceleration vs time is measured with a sled-borne accelerometer and PDM telemetry system. The tape-recorded data are converted to digital form and entered into the IBM-704 computer by automatic assessment equipment. The two different sets of data are combined by a near-optimum digital filtering technique to provide a set of hybrid wide-bandwidth data.

INTRODUCTION

THE supersonic Naval ordnance research track offers possibilities of testing ordnance equipment over a large range of velocities and accelerations without loss or destruction of the equipment. Detailed performance of the ordnance equipment during the run down the track may be obtained with extensive recoverable sled-borne instruments. Precise velocity of the instrument-carrying sled is often required. The specifications for a velocity measuring system (vms) to meet these requirements were 0.1-ft/second rms error with 0- to 50-cps bandwidth in the range of 200 to 2000 ft/second. The bandwidth was rather arbitrarily set at 50 cps, but it appeared to be a good compromise between practicality and ideal requirements.

In themselves, cameras, position markers along the track and the like are not capable of yielding data of the required precision under field conditions. However, it is possible to obtain precision velocity information by the proper combination of two types of measurements (hybrid data). The vms developed for SNORT uses the longitudinal acceleration data from a sled-borne accelerometer and time-position data from some 215 magnetic-pickup coils spaced at 100-ft intervals along the track.

A magnet on the sled generates an electrical pulse as it passes over each coil. These pulses recorded with a time base provide accurate position-time data at discrete points along the track. The accelerometer provides longitudinal acceleration data for the sled and is sampled at rates of 800 to 900 times per second. The problem is now to combine these data to provide a detailed and precise velocity history.

COMPUTATIONAL SCHEME

The following outline of the hybrid-data scheme is based largely on the work compiled by the staff of the test department,¹ here referenced for more detail. Since most of the work was initially reported in internal memoranda and by private communications, sufficient detail is given here to suggest the basic theory and computational methods. Details of the electronic equipment associated with transmitting and recording the data and translating it into a form suitable for input to the IBM-704 computer will be discussed immediately following this section. We shall start here with the assumption that track-coil position data and the acceleration data are at hand in a digital form.

Differentiation of position data will in principle give velocity. However, noise due to vibration, uncertainties in the track-coil survey and the location of the magnetic centers of the magnet, and aliasing error introduce noise in the velocity obtained by differentiation. If the spectral density of the position noise is substantially constant up to some limiting frequency, *i.e.*, frequency-limited white noise, differentiation multiplies the rms noise spectral density by the angular frequency. Similarly, if the noise of the accelerometer data is frequency-limited white noise, integration to velocity divides the rms noise spectral density by the angular frequency. This situation is illustrated in Fig. 1. Now, if the two types of data are combined with suitable weighting in frequency (a crossover filter), it is possible to minimize the error by taking advantage of the nature of the spectral densities of the errors.

A method for such combination was proposed by Beutler and Rauch.¹ They assume that the measured acceleration $a^*(t)$ is of the form

$$a^*(t) = K_0 + (1 + K_1)a(t) + R_a(t), \quad (1)$$

where $a(t)$ is the true acceleration, K_0 and K_1 ($\ll 1$) are constants to be determined experimentally, and $R_a(t)$ is band limited white Gaussian noise with spectral density S_{R_a} for $0 \leq \omega < \omega_c$ where ω_c is the cutoff frequency. Two successive integrations of (1) give the indicated velocity $v^*(t)$ and position $x^*(t)$ in terms of the true velocity and position $v(t)$ and $x(t)$. Thus,

$$v^*(t) = K_0 t + (1 + K_1)v(t) + R_v(t) \quad (2)$$

* Original manuscript received by the IRE, December 3, 1958.

† U. S. Naval Ordnance Test Station, China Lake, Calif.

¹ Staff of the Test Dept., "A Precision Velocity-Measurement System for SNORT," U. S. Naval Ord. Test Station, China Lake, Calif., NOTS 1449 NAVORD 5247 1; April, 1956.

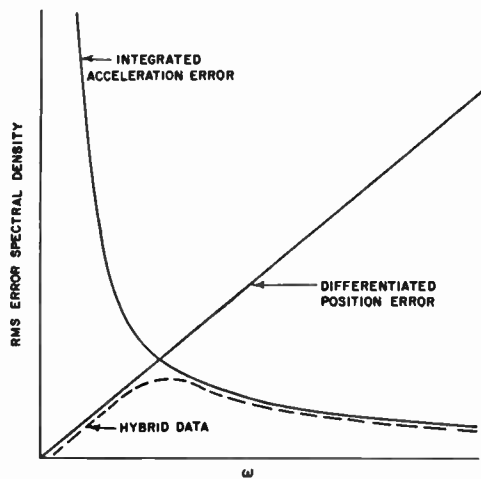


Fig. 1—Plot of spectral density of rms velocity error for two different primary measurements: position and acceleration. Curves assume flat acceleration and position error spectra.

and

$$x^*(t) = \frac{1}{2}K_0t^2 + (1 + K_1)x(t) + R_x(t), \tag{3}$$

where

$$R_v(t) = \int_0^t R_a(t)dt \quad \text{and} \quad R_x(t) = \int_0^t R_v(t)dt.$$

The spectral densities of $R_v(t)$ and $R_x(t)$ are

$$S_{R_v}(\omega) = \frac{S_{R_a}}{\omega} \tag{4}$$

and

$$S_{R_x}(\omega) = \frac{S_{R_a}}{\omega^2}. \tag{5}$$

On the basis of $R_a(t)$ being an uncorrelated discrete sequence, Beutler and Rauch showed that the (ensemble) rms values of $R_v(t)$ and $R_x(t)$ are

$$\sigma_{R_v}(t) = S_{R_a}(\pi t)^{1/2} = \sigma_{R_a} \left(\frac{\pi t}{\omega_c} \right)^{1/2} \tag{6}$$

and

$$\sigma_{R_x}(t) = S_{R_a} \frac{\pi t}{(2\omega_c)^{1/2}} = \sigma_{R_a} \frac{\pi t}{\sqrt{2}\omega_c} \tag{7}$$

for $t \gg \pi/\omega_c$.

The positions $x^*(t)$ are now corrected as follows: K_0 and K_1 are determined from three widely spaced track-coil points (x_0, t_0) , (x_1, t_1) , and (x_2, t_2) . The first corrected velocity $\bar{v}(t)$ is then computed from

$$\bar{v}(t) = \frac{v^*(t) - K_0t}{1 + K_1}. \tag{8}$$

If $x_0=0, t_0=0$ (i.e., take the initial sled position and time as the origin) (3) gives

$$K_0 = 2 \frac{(x_2^*/t_2^2)(x_1/t_1^2) - (x_1^*/t_1^2)(x_2/t_2^2)}{x_1/t_1^2 - x_2/t_2^2}, \tag{9}$$

$$K_1 = \frac{x_1^*/t_1^2 - x_2^*/t_2^2}{x_1/t_1^2 - x_2/t_2^2} - 1, \tag{10}$$

where the noise $R_v(t)$ has been temporarily ignored.

The error due to the noise $R_v(t)$ is corrected by comparing the distance between two successive track coils $\Delta x = x_{n+1} - x_n$ with the distance computed from

$$\int_{t_n}^{t_{n+1}} \bar{v}(t)dt$$

over the time interval $\Delta t = t_{n+1} - t_n$ corresponding to Δx . The second corrected velocity $\hat{v}(t)$ is thus

$$\hat{v}(t) = \bar{v}(t) + \frac{1}{\Delta t} \left[\Delta x - \int_{t_n}^{t_{n+1}} \bar{v}(t)dt \right]. \tag{11}$$

The correction

$$\Delta \bar{v} = \frac{1}{\Delta t} \left[\Delta x - \int_{t_n}^{t_{n+1}} \bar{v}(t)dt \right]$$

is computed for each track-coil interval and used for that interval.

Beutler and Rauch made an error analysis of this process of correction. They showed that if $x^*(t_1)$ and $x^*(t_2)$ are essentially uncorrelated, the rms errors of K_0 and K_1 are

$$\sigma_{K_0} = \frac{2(x_1\epsilon_{x_2}^* + x_2\epsilon_{x_1}^*)}{t_1^2t_2^2 |x_1/t_1^2 - x_2/t_2^2|} \tag{12}$$

and

$$\sigma_{K_1} \leq \frac{\epsilon_{x_1}^*/t_1^2 - \epsilon_{x_2}^*/t_2^2}{|x_1/t_1^2 - x_2/t_2^2|}, \tag{13}$$

where $\epsilon_x^{*2} = \epsilon_x^2 + \sigma_{R_x}^2(t)$; that is, the positional error of x^* consists of the rms error of the track-coil position ϵ_x and the error due to noise in the accelerometer system. They also showed that the rms error in the corrected velocity is

$$\bar{\epsilon}_v^2 = \frac{\pi}{\omega_c} \sigma_{R_a}^2 \Delta t + \frac{2\epsilon_x^2}{(\Delta t)^2} + \frac{1}{2} \left(\frac{\pi}{\omega_c} \right)^2 \sigma_{R_a}^2 \tag{14}$$

for the interval Δt .

With the following set of values

$$\begin{aligned} \sigma_{R_a} &= 3.2 \text{ ft/second}^2 \\ \omega_c &= 314 \text{ rad/second} \\ \epsilon_x &= 0.01 \text{ ft} \\ x_1 &= 10^4 \text{ ft}, \quad x_2 = 2 \times 10^4 \text{ ft} \\ t_1 &= 12 \text{ seconds}, \quad t_2 = 24 \text{ seconds}, \end{aligned} \tag{15}$$

and referring to (12) and (13), they compute that

$$\begin{aligned} \sigma_{K_0} &\leq 7.6 \times 10^{-3} \text{ ft/second}^2, \\ \sigma_{K_1} &\leq 8.1 \times 10^{-5}, \end{aligned}$$

and that with these values it requires 5.1 seconds and 1.9 seconds to develop a 0.1-ft/second error in velocity

due to errors in K_0 and K_1 , respectively, for a 20-g change in acceleration. By correcting the velocity computed from (8) at 1.0-second intervals or less, errors due to K_0 and K_1 can be kept within the required 0.1 ft/second. This correction is just that given by (11). With a coil spacing of 100 feet and velocities in the range 200–2000 ft/second, corrections occur at least every 0.5 second.

Now referring to the error equation (14) with $\Delta t = 0.5$ second and the other constants from (15), the error is

$$\sqrt{\bar{\epsilon}_v^2} = 0.08 \text{ ft/second.}$$

This, then, is the basic computational scheme used to compute precision velocity.

The error analysis discussed briefly here has been further investigated by Mallinckrodt.^{1,2} He points out that there is an implicit assumption of the independence of successive velocity errors in the analysis leading to the error equation (7). In general successive velocity errors are not independent; considering this, Mallinckrodt's analysis heads to a corrected error equation:

$$\sigma_{R_x}(t) = \sqrt{\frac{\pi}{3\omega_c} \sigma_{R_a} t^{3/2}} \quad (16)$$

for $t \gg (\pi/\omega_c)$, to replace (7). There is $t^{3/2}$ rather than t dependence of $\sigma_{R_x}(t)$ in this case.

Now, in the original analysis leading to the error equation (14), the maximum value of $\sigma_{R_v}(t)$ was introduced with an included dc component. The dc component is taken into account by the correction process, and thus only the random ac component $\sigma_{R_{v,r}}$ need be used in the error equation. Further, Mallinckrodt suggested a time average over the track-coil interval of $\sigma_{R_{v,r}}$ be used. He then arrived at an error equation to replace (14).

$$\epsilon_v^2 = \frac{\pi}{\omega_c} \sigma_{R_a}^2 \frac{\Delta t}{2} + \frac{2\epsilon_x^2}{(\Delta t)^2} \quad (17)$$

This gives a somewhat smaller error than (14).

From this equation an optimum reset time may be computed, that is, the interval Δt minimizing ϵ_v . This is

$$(\Delta t)_0 = \left(\frac{8\omega_c \epsilon_x^2}{\pi \sigma_{R_a}^2} \right)^{1/3} \quad (18)$$

When the velocity is high, the time interval between successive coils will be small and it has been suggested¹ that at high velocities Δt be taken over several coil intervals to more nearly match $(\Delta t)_0$. Curves of error given by Hendrix, Hilker, and Gates¹ are shown in Fig. 2 with Δt taken for single track-coil intervals. It is seen that the error grows for high velocity in this case. Mallinckrodt¹ computed the error for Δt corresponding to

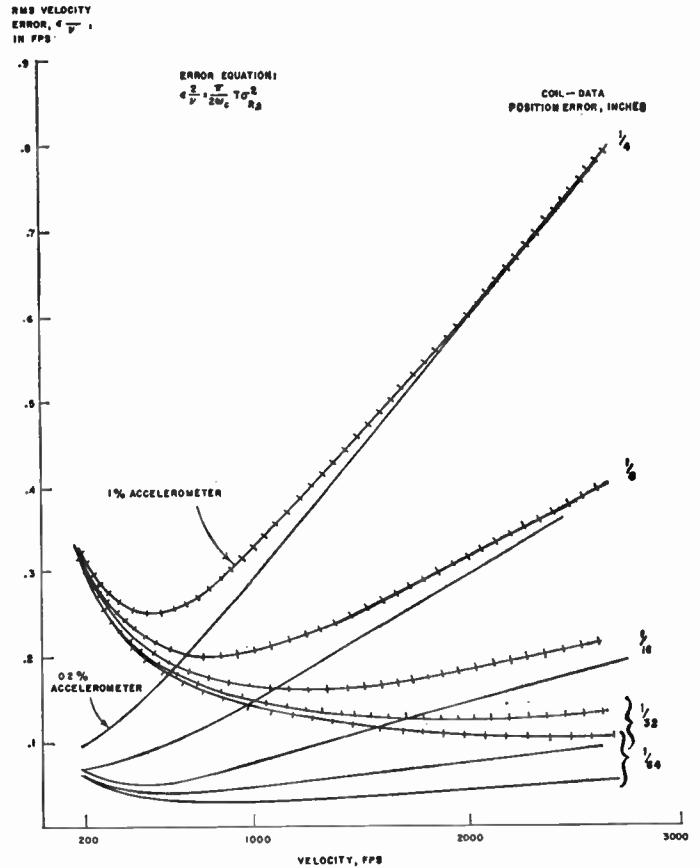


Fig. 2—Plots of rms velocity error vs velocity for the Beutler-Rauch method.

skipping n coils. Curves are shown in Fig. 3. He also showed curves for the optimum reset time $(\Delta t)_0$. By skipping coils, it is possible to keep the error within the required limit.

This process of combining the acceleration and position may be thought of in the frequency domain as using a self-complementary crossover filter. Mallinckrodt treated hybrid data from this point of view in some detail.^{1,2} His conclusions are: 1) the original Beutler-Rauch method when compared to the continuous optimum gives an error three or four times that of the continuous optimum. 2) The optimum infinite-lag crossover filter is

$$y(\omega) = \frac{1}{1 + \tau^4 \omega^4} \quad (19)$$

where

$$\tau = \frac{\Delta T \epsilon_x^2 \omega_c}{\sqrt{\sigma_{R_a} \pi}} \quad (20)$$

(Note that infinite lag filters are permissible in this case since all computing is done after the entire set of data is available.³) 3) The impulsive response of the Beutler-Rauch filter is approximately a triangular wave whose

¹ A. J. Mallinckrodt, "A Hybrid Data System for SNORT," Interstate Electronics Div., Interstate Eng. Corp., Anaheim, Calif., IE-1001-F1; August 1, 1956.

³ O. N. Strand, U. S. Naval Ord. Test Station, China Lake, Calif., unpublished NOTS memo; November 12, 1956.

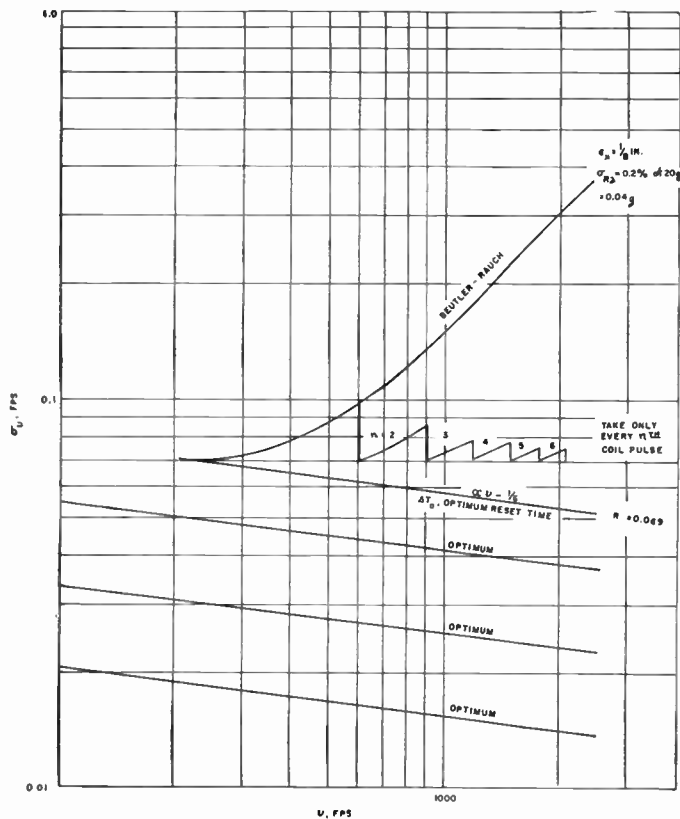


Fig. 3—Plots of rms velocity error vs velocity as a comparison of Beutler-Rauch and optimum methods

transfer function is then of the form $y(\omega) = \sin^2 x/x^2$. Comparison of these two transfer functions is shown in Fig. 4, using the optimum reset time $(\Delta t)_0$. It is seen that the Beutler-Rauch near-optimum filter is not as sharp as the theoretical optimum.

The error analysis discussed here is based on an accelerometer noise spectrum cutting off at 50 cps. It was thought at first that a presampling filter would be required, but the accelerometer and the digitalizing network serve adequately for this filter.

The accelerometers used for vms have transfer functions of the form

$$F(\omega) = \left[1 - \left(\frac{\omega}{\omega_0} \right)^2 + 2j\xi \frac{\omega}{\omega_0} \right]^{-1}$$

where ω_0 is the resonant frequency of the accelerometer and ξ is the damping factor. The computations of velocity are done on a digital computer (IBM 701 and 704), and a digital compensating filter could be introduced during the velocity computation. But, since this involves significant amounts of computing time, compromise filters were considered. If $Y^*(\omega)$ is the compensating filter, then the rms in velocity due to lack of compensation is

$$\epsilon_v^2 = \int_0^\omega \frac{S_{Ra}}{\omega^2} |1 - Y^*(\omega)F(\omega)|^2 d\omega$$

It was found through some work of Strand³ that a simple time delay filter $Y^*(\omega) = e^{kj(\omega/\omega_0)}$ will keep this error

well below the required 0.1 ft/second. In this work, k was adjusted to minimize the error, which was then on the order of 0.02 to 0.03 ft/second.

The basic computational scheme used for the SNORT hybrid data system follows these steps: 1) Compute the over-all scale factor $1 + K_1$ and bias or offset K_0 from widely spaced positions along the track. This step provides the first corrected velocity $\bar{v}(t)$ from the integrated acceleration, $v^*(t)$ (2), (3), (8), (9), and (10). 2) Correct $\bar{v}(t)$ for noise by comparison of the difference of position computed from $\int \bar{v}(t) dt$ with that of the track coil, (11). These computations are all performed on a digital computer (currently the IBM 704). The digital acceleration data are compensated for the transfer function characteristics of the accelerometer in the computer by a simple time delay before the above calculations are made. Some consideration of curve fitting of the track coil and accelerometer data was made by Zilmer and Fay¹ and it appeared that some improvement of the over-all error was possible but it has not currently been incorporated because the low-velocity error is increased by curve fitting.

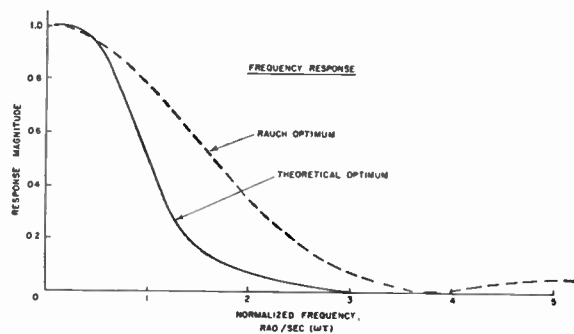


Fig. 4—Plots of impulse and frequency responses, comparing Rauch's near-optimum filter with the theoretical optimum.

There are objections to assuming that the noise is band-limited white Gaussian noise and that it is a stationary process. Leipnik⁴ has investigated this subject in some detail, but the error analysis becomes extremely cumbersome unless some simplifying assumptions such as the process is stationary over short intervals at least, are made.

TRACK-COIL SYSTEM

The track-coil system consists of an E-shaped permanent magnet mounted on the test sled, U-shaped pickup coils installed every 100 feet along the track, and ten twisted-pair wire transmission lines connecting the coils to the terminal equipment in a test-control building. The transmission lines are modified by the addition of shunting resistors across the lines,⁵ and the lines are terminated in their characteristic impedance at the receiving end.

⁴ R. B. Leipnik, U. S. Ord. Test Station, China Lake, Calif.; unpublished NOTS manuscript.

⁵ B. W. Pike, "The Sled Position-and-Time Measuring (Track-Coil) System at SNORT," U. S. Naval Ord. Test Station, China Lake, Calif., NOTS 1202; September 1, 1955.

When the sled-borne E magnet passes over a track coil, the changing magnetic field induces a current pulse in the winding, resulting in a voltage pulse in the transmission line to which the coil is connected. The transmitted pulses are received in the SNORT test-control building, where they are combined in a mixer⁶ and amplified.

Fig. 5 illustrates how the voltage of the pulse varies as the magnet passes over a coil. The reference point on the voltage pulse is defined as that point corresponding to the instant the electrical center of the E magnet is over the electrical center of the coil. The reference point is an accurately detectable node.

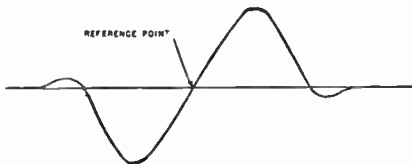


Fig. 5—Waveform of voltage induced in a U coil by excitation with an E magnet.

The task of the precision pulse converter is to pick out the reference point from the other two nodes on the E-magnet waveform and present this information in suitable form for recording on tape.

The heart of the precision pulse converter is the transducer used to detect the three null points on the induced waveform. This device is the National Bureau of Standards error-voltage detector.⁷ The NBS detector has a zero-drift stability approaching 1-mv dc per day.

The NBS null detector consists of a Wheatstone bridge in which two arms are inductors and the other two arms are germanium diodes. The two bridge inductors are excited by a third inductor which is connected to a 30-mc RF source as illustrated in Fig. 6.

The track-coil waveform applied at the terminals E_x and E_y varies the resistive balance of the bridge by differentially biasing the two diodes R1 and R2. Since the track-coil waveform is effectively chopped at a 30-mc rate by the bridge RF excitation, the result is a series of instantaneous imbalances of the bridge. When the diode resistances become instantaneously equal due to the zero-biasing effect of the control signal at crossover, a null or minimum RF bridge output is obtained.

The output of the bridge is amplified and detected. The resulting signal is three sharply defined null spikes. The spike corresponding to the reference point is selected from the other two and used to generate a square pulse which is tape recorded along with a 50-kc time base and acceleration data obtained from the sled-borne accelerometer.

⁶ P. M. Sherk, "Instruction Manual for the SNORT Track Coil Pulse Shaper," U. S. Naval Ord. Test Station, China Lake, Calif., NOTS Tech. Memo. 1785; November 1, 1954.

⁷ B. Chance, et al., eds., "Waveforms," M.I.T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., pp. 45-47; 1949.

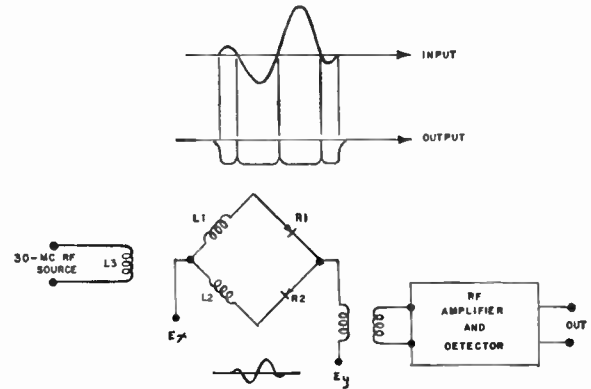


Fig. 6—Circuit diagram of null detector in precision pulse converter.

Sources of Track-Coil System Errors

It can be shown that for small values of Δt the proportional error in velocity is given approximately by

$$\frac{\Delta v}{v} = \frac{\Delta x}{x} + \frac{\Delta t}{t}$$

where Δv , Δx , and Δt are the errors in velocity, track-coil spacing, and time determination, respectively. Therefore, at any given velocity an error in the time determination between any two consecutive track coils can be related to an equivalent distance error, and conversely a distance error can be related to an equivalent time error. The errors associated with the track-coil system are:

- 1) the uncertainty of the exact location of the electrical centers of the track coils;
- 2) the uncertainty of the distance between the electrical centers;
- 3) the time error in detecting the zero crossing of the track-coil signal;
- 4) the uncertainty of the determination of the time between pulse reference points;
- 5) the time error associated with distortion of the generated pulse due to different magnet velocities and anomalies in individual track coils;
- 6) the time error caused by differences in pulse propagation velocity because of different line propagation characteristics from line to line, and different spectrum content of individual pulses, and
- 7) the error caused by the flexure of the magnet support and angular motion of the sled.

Discussion of Track-Coil System Errors

If the U-core track coils were perfectly symmetrical electrically, their electrical centers would coincide with their geometrical centers and the distance between them could then be determined by an accurate survey. Unfortunately, there is no guarantee that the discrepancy between the electrical and geometrical center is negli-

gibly small. For this reason, it was first necessary to locate the electrical center of each coil to a known tolerance and make a permanent legible mark related to this center to use as a survey point.

An E-shaped ac electromagnet of $n-s-n$ polarity when passed over a U-shaped coil induces a voltage with the waveform shown in Fig. 7.

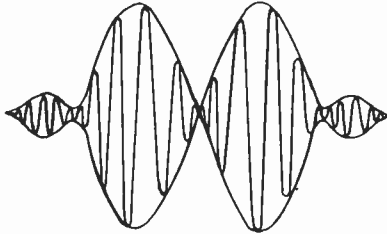


Fig. 7—Waveform of voltage induced in track coil by E electromagnet.

The center of the waveform where the phase reversal takes place corresponds to the electrical center of the track coil. By means of such a magnet attached to a jig, together with a jeweler's saw for etching a line on the track-coil cover, a mark was produced that was reproducible to ± 0.003 inch.

The distances between track-coil center marks were surveyed to an accuracy that varied from the absolute distance by not more than ± 0.001 foot.

The error of the pulse converter in measuring the null is inversely proportional to the velocity of the magnet. The greater the velocity, the faster the phase change in the input waveform and the sharper the null indication from the RF amplifier. At a velocity of 200 ft/second for example, the approximate maximum time error corresponds to a displacement error of ± 0.010 inch. At velocities in excess of 500 ft/second, the error is less than ± 0.001 inch.

The time between crossings of successive track coils is determined and recorded in digital form by equipment called the tape data assessor. This equipment will be described later. The least count of the time base used in this time determination is one μsec . There is also a time indeterminacy in detecting the occurrence of the recorded track-coil pulse as played back on other tape equipment. However, the track-coil position can be determined to within two least counts of the time base used in the tape data assessor. Tests have demonstrated that the rms error in the time determination is approximately $0.5 \mu\text{sec}$.

A large number of oscillographs was made by recording the track-coil pulses with varying magnet velocities at the site of several coils. At no instance could a change in the general pulse shape or a change in the symmetry of the pulses be detected. In this manner, a confidence was established that the reference null of the track-coil wave shape was not velocity dependent and that any anomalies in individual coils were of a minor nature.

To determine the characteristics of the track-coil

transmission lines and to determine the effect of the lines on a transmitted pulse, a study⁸ was made. An analysis of the spectrum of the track-coil pulse showed that for velocities between 200 and 2000 ft/second the significant energy was in a frequency band for which the phase shift was a linear function of frequency and the attenuation factor was constant except at very low frequencies. However, as pointed out,⁸ the effect on the cross-over point contributed by the low-frequency components is approximately 1/25 that contributed by the high-frequency components.

The parameters of several 1000-ft sections of the transmission lines were measured and were found to be within 1 per cent of each other. From these measurements and the analysis above, it was concluded that there is no appreciable error arising from propagation over the track-coil transmission lines.

It was found, however, that the most significant error source in the track-coil system is the uncertainty of the E-magnet location, relative to other parts of the test sled. The sleds ride on skids attached to the tracks; the wearing of the metal in the skids as the sled proceeds down the track will often result in a $\frac{1}{8}$ -inch clearance between the skid and the rails. The test sled is not a truly rigid structure, and under the extreme environmental conditions of high-speed testing there is motion between its various parts.

Sled designers estimate that there is as much as a 0.01-ft relative motion between the position of the E magnet and the center of gravity of the sled, during a run. Table I summarizes the values of the errors in the track-coil system.

TABLE I
ERRORS IN TRACK-COIL SYSTEM

Location of electrical centers of track coils	0.003 in.
Distance between electrical centers at 100 ft.	0.012 in.
Position of pulse reference point at 200 fps (worst condition)	0.010 in.
Time between pulse reference points	2 μsec *.
Approximate motion between magnet and sled center of G	0.125 in.
Over all position error (rms)	0.125 in.

* The time error of 2 μsec is equivalent to 0.048 inch at 2000 ft/second and decays linearly to 0.0048 inch at 200 ft/second.

Obviously, the limiting factor in the precision of the track-coil system is the accuracy of the location of the magnet with respect to another part of the test sled. For more exacting data, it is necessary that greater effort be made to support and restrain the magnet relative to a desired location on the sled.

PDM TELEMETRY SYSTEM

Accelerometer

The quality of the acceleration data depends on the

* "Final Engineering Report on the Study of SNORT Supersonic Research (Track Coil) Transmission Line," U. S. Naval Ord. Test Station, China Lake, Calif. Submitted to Commanding Officer, U. S. NOTS, under Contract N 123 (60530)11346A and prepared by RCA Service Co., Inc., Govt. Service Dept.

accuracy of the acceleration transducer (accelerometer) and the accuracy of the transmission link to the IBM computer. The accelerometers chosen for the vm system have an accuracy of 0.1 per cent full scale; a ± 20 -G range; natural frequency of over 100 cps; and have less than a 1 per cent effect from acceleration in the two axes perpendicular to the sensitive axis. It is demonstrated by Sollenberger¹ that under the fixed-sled restraints and a reasonably assumed frequency spectrum that a cross-axis effect of less than 1 per cent would contribute a negligible component to the sensitive axis. It is also demonstrated that the accelerometer must be aligned to within $\frac{1}{2}^\circ$ to maintain the same order of accuracy.

Telemetry Package for Sled

The telemeter package consists essentially of a free-running blocking oscillator with a nominal repetition rate of 900 cps, a one-shot multivibrator keyer to sample and convert the voltage amplitudes of the accelerometer signal to corresponding pulse durations, a crystal-controlled FM transmitter with an RF power amplifier and necessary power supplies (Fig. 8).

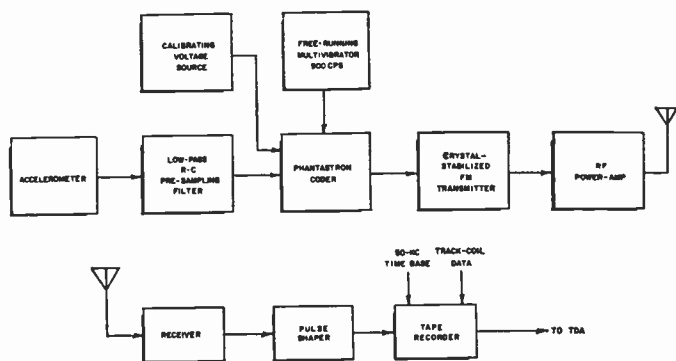


Fig. 8—Block diagram of PDM telemetry package and receiving-recording equipment for SNORT vm system.

In order to avoid excessive aliasing errors arising from the sampling process, a presampling filter is inserted between the accelerometer and the keyer. The concept of aliasing errors is developed by Mallinckrodt and Bosch,⁹ in the introduction of the chapter on sampled data.

The keyer is linear to about 0.2 per cent over its full range of 200–800 μ sec and is subject to long-term drift. Therefore, a 17-point calibration is provided immediately before and after a sled test. This calibration also takes care of drift at the receiver.

To insure low noise in the system, all of the circuits in the telemeter package are transistorized except for the RF portion, the power delivered to the antenna is approximately 10 watts, and the sections of the system that are critical are uniquely isolated from vibration and shock (Fig. 9).

⁹ A. J. Mallinckrodt and E. I. Bosch, "Data Smoothing Techniques," Ralph M. Parsons Co., Pasadena, Calif., Rep. 1026-P2 for Edwards AFB; March 21, 1955.

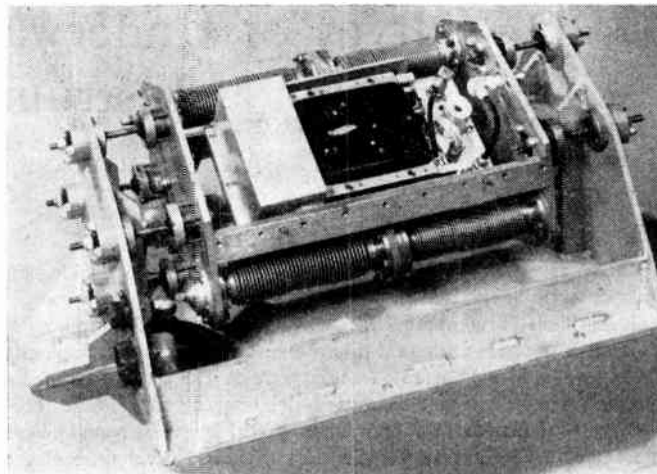


Fig. 9—Telemetry package.

The PDM signal is received by a crystal-stabilized receiver located at the SNORT test-control building. The output of the receiver is direct-coupled to preserve an accurate dc level. The output waveform is a square wave with rise and fall times of 10 μ sec. The 50 per cent voltage level of this wave is detected, and a square wave with rise and fall times less than one μ sec is generated and is ac coupled to a tape recorder and recorded on separate tracks along with the track-coil pulses and a standard 50-kc time base.

Table II presents the errors in acceleration data. Considering 300 μ sec as full scale for the positive signals of the PDM keyer, 0.6 μ sec represents 0.2 per cent.

TABLE II
ERRORS IN ACCELERATION DATA

Linearity and resolution of accelerometer (0.1 per cent)	0.3 μ sec
Time duration of PDM pulse	0.5 μ sec
Over-all Acceleration Error (rms)	0.6 μ sec

TAPE DATA ASSESSOR

The tape data assessor (TDA) is a unit designed to digitize the PDM telemeter signal and the time between track-coil pulses. The basic clock signal of the TDA is one mc. The zero crossing of the 50-kc sinewave signal that has been recorded on the analog tape is detected. The 50-kc signal is the basic time and the one-mc clock signal is used as a fine count between two events, such as a track-coil pulse and the next 50-kc pulse. Allowing a difference of 5 per cent between the recording speed and the playback speed, and for flutter and wow in the tape recorders, the error introduced is the least count, or one μ sec of the TDA. The output of the TDA is recorded in a digital code compatible for input into the IBM-704 computer. On one tape is information containing the time between successive track-coils; on the other is the time duration of the PDM acceleration data, the PDM pulse repetition rate, and the time to nearest track coil, so that acceleration data and track-coil data can be coordinated.

Survey of Underwater Missile Tracking Instrumentation*

D. T. BARRY†, MEMBER, IRE, AND J. M. FORMWALT†

Summary—The development of underwater missile tracking instrumentation is reviewed, and current underwater tracking ranges are described. Three approaches have been taken to the problem of tracking a torpedo: internal instrumentation, external fixed-site tracking ranges, and portable tracking arrays. The principal tracking method employed is that of a fixed-site tracking range. The technique is to plant an array of hydrophones, which receive sound pulses generated in the moving vehicle. The relative arrival time of the pulses, known array geometry, and propagation velocity are then employed to compute the coordinates, speed, and acceleration of the underwater missile.

I. INTRODUCTION

THE requirement for underwater missile tracking instrumentation first began to arise about 1864, when the English inventor Whitehead designed the first free-running mobile torpedo with gyroscopic control in azimuth and hydrostatic control in depth. With such a missile available, the need for a determination of its placement accuracy became apparent. For many years this determination could be made only by firing the weapon at a target and attempting to estimate miss distance by visual observation. This method eventually led to firing torpedoes to pass under a barge on which observers were stationed every few feet; when the torpedo passed underneath, the closest observer signaled its position, and estimated its depth from the time necessary for bubbles from the exhaust gases of the steam propulsion system to reach the surface. This depth estimate was later compared with the depth record from an internal Depth and Roll Recorder. After the development of "wakeless" torpedoes, a spotlight directed upward was used in the torpedo to indicate its position, and firing was often done at night to facilitate visual tracking. Another method which was employed to a considerable extent was to fire the torpedo to pass through one or more light nets of wire or string supported by floats, and measure the location of the resulting holes in the nets.

Still another approach was to tie a cloth bag of dry powdered fluorescein or rhodamine dye to the torpedo, thus causing it to leave a trail which might be photographed from the air; this generally did not produce a very satisfactory track, but did serve to mark the spot at which the torpedo surfaced at the end of the exercise run. Liquid dye dispensed from a tank under air pressure, and carried upward by air bubbles, was more

successful in marking the track, but was useful only for comparatively shallow runs.

A somewhat more elaborate system (and probably the earliest involving electronics) utilized a radio network linking several transit stations, at each of which the bearing of the beginning of the wake was taken each time a signal was received by radio. The resulting series of cross bearings permitted plotting the course of the weapon on a chart.

Each of these relatively crude methods of tracking underwater missiles has served a useful purpose, but as the capability and complexity of the missiles has increased, improved tracking methods have become necessary. For example, it is now necessary to run torpedoes at much greater depths than those permitting optical tracking or the use of nets, and plotting of complicated trajectories is required to evaluate properly the performance of acoustic homing systems. To meet the requirement for increased information regarding performance of the weapon, internal data-recording instrumentation is frequently provided for the exercise versions of underwater missiles. This approach has the advantage that the same recorder can also be used to record variables other than those immediately concerned with the tracking problem; its disadvantages include the need for spatial or other references in the missile, and the possibility of total loss of information if the missile is not recovered. The latter disadvantage could be overcome by ejection of a buoyant data capsule, if the means of ejection were made sufficiently reliable, and versatile enough to cover all possibilities—conditions that are probably no more difficult to meet for the entire vehicle than for the data capsule alone. Another obvious means of overcoming this disadvantage would be telemetering of the data, if telemetering at a sufficiently high information rate over long salt-water paths were less difficult. Both of the disadvantages cited for internal recording of the tracking data can be overcome by the use of an external tracking installation, and several external tracking facilities based on various operating principles have been built and operated successfully. Presently anticipated performance of planned underwater missiles will require increased capabilities of the tracking installations, however, and the design of a tracking system to meet these anticipated requirements is no small problem, as will be apparent from a discussion of some of the existing tracking instrumentation systems.

* Original manuscript received by the IRE, December 24, 1958.

† U. S. Naval Underwater Ordnance Station, Newport, R. I.

II. INTERNAL TRACKING INSTRUMENTATION

The earliest form of internal tracking instrumentation in general use was the Depth and Roll Recorder, which incorporated a bellows or diaphragm linked to a stylus which scratched a record representing torpedo running depth on a coated paper which was moved by clockwork to provide a time base. Another stylus linked to a pendulum provided a record of the angle of roll of the torpedo. The simplicity and reliability of this instrument permitted its use in large numbers of practice runs, resulting in the accumulation of a large amount of information about torpedo running depth. Unfortunately, much of the data were in error because it was not generally recognized that the placement of the recorder along the length of the torpedo was rather critical if the device was to record actual running depth; if the location chosen were not the correct one, some portion of the dynamic pressure resulting from the torpedo's motion would be added to the depth pressure. A later improvement of this type of instrument, developed by the Naval Ordnance Laboratory, White Oak, Md., was the Depth, Roll, and Speed Recorder. In this instrument more care was used in establishing a reliable static depth pressure, and a differential bellows was added to measure torpedo velocity by subtracting the depth pressure from the total pressure at the nose of the torpedo.

Recording depth and roll, or depth, roll, and speed provides useful information but is far short of the goal of tracking the weapon. A number of attempts have been made to provide sufficient internal instrumentation to permit determination of the path of the torpedo in three dimensions, by analysis subsequent to the run.

In 1945 and early 1946, one of the authors was responsible for the development and testing of an internal recorder intended to provide information concerning the dynamic performance of torpedoes. In this instrument, both digital and analog transducers were used, and the information was recorded by means of an electrically driven 16-mm movie camera. A pair of coupled, pneumatically driven gyroscopes were used to determine course, roll, and pitch, with magnetically clamped pendulous indicators of roll and pitch as well. Bellows assemblies were employed as depth and velocity transducers, with revolution counters to record drive shaft revolutions and digital pick-offs to measure the positions of the control surfaces. The instrumentation was largely mechanical, and required precision workmanship in its construction and careful maintenance to insure accurate results. Data reduction and analysis were tedious and time consuming, but the recorder was used successfully in a number of runs and provided much needed information regarding the dynamic performance of torpedoes, particularly of high-speed, nonhoming types.

The need for reading and plotting the data from a number of frames of the film, in order to estimate the behavior of any recorded variable, was a distinct dis-

advantage of the recorder just described. About this time, however, galvanometer type oscillographs of small size, high sensitivity, and adequate frequency response for most requirements began to become available and to be used widely in internal torpedo instrumentation. For those applications requiring response to frequencies above the range of galvanometers, cathode-ray tubes have also been used extensively, and magnetic tape recorders to a limited extent. With the graphic recorders, the signal from a free or rate gyro, an accelerometer, a depth pressure transducer, and a differential pressure transducer for velocity measurement could be plotted continuously as a function of time, permitting crude estimation of dynamic performance by merely scanning the photographic record. For more refined analysis semi-automatic readout and plotting of the data were possible, using equipment which required an operator only to maintain an index point in coincidence with the record trace as the oscillograph record was pulled through the machine; positioning the index point would also vary an electrical output, which could then be used in any manner desired. As an example, a signal representing instantaneous torpedo heading would be used to orient a plotting-board "bug," which was motor-driven at a speed proportional both to the rate at which the record was pulled through the reading device and to the signal representing instantaneous torpedo velocity; the end product of this device would be a two-dimensional plot, to scale, of the course run by the torpedo.

In addition to the measurements relying on gyros and on pressure measurements to track the underwater missile, those weapons which had homing systems could make use of information derived from the homing system for tracking purposes. With a passive acoustic homing system, only an approximate target direction could be determined after target acquisition; but with an active acoustic system, the range-to-target would be available as well as a more accurate indication of target bearing. This information could be compared with the information derived from dead-reckoning plots obtained by the firing vessel or target vessel.

Since each type of vehicle has its own set of operating characteristics, no single internal instrumentation package has been designed to record the dynamic behavior of all of them. A few attempts have been made to provide a universal torpedo recorder, but the results of these attempts have not been widely accepted. Rather, an "exercise section" or in some cases several different exercise sections have been developed for each type of torpedo, and special recorder packages in considerable variety have been built in small quantities to fit the requirements of specific research and development programs. Fig. 1 shows two views of a typical special-purpose recorder; of the tracking portion of the instrumentation, only a gyro is visible, mounted above the

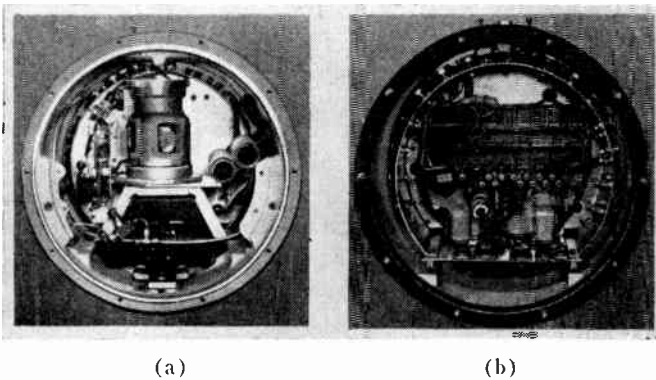


Fig. 1—A typical special-purpose internal instrumentation package. In (a), a gyroscope for recording the instantaneous course of the torpedo is located above the inverter; (b) is the rear view of the same package.

inverter which supplies 400-cycle power; the tubing visible in the same picture is a part of the depth and velocity measuring system, but the transducers are not shown.

III. EXTERNAL TRACKING

Tracking Range, Torpedo (TRT)

In addition to instrumentation systems carried within the missile to provide a record of its trajectory, a number of instrumented ranges have been developed. With the advent of the electric torpedo and the pressures of World War II, it became evident in September, 1942, that a daylight method of ranging electric torpedoes (which were then ranged at night and used a vehicle-installed searchlight to provide a visible path) was necessary in order to increase the number of torpedoes which could be proved and also to improve the accuracy of the speed, running depth, and deflection data for both the steam and electric torpedoes then being tracked visually. At that time the Naval Ordnance Laboratory undertook an investigation of the feasibility of tracking torpedoes magnetically or acoustically¹ and the range at the U. S. Naval Torpedo Station, Newport, R. I., was selected as the site for the development.

Magnetic tracking was attempted first, the concept being to lay flat loops on the bottom of Narragansett Bay. Preliminary attempts to lay such loops soon showed that the positioning of the cables on the bottom with the required accuracy would be extremely difficult and expensive to install and maintain, and magnetic measurement of speed and deflection was abandoned. However, one temporary loop (300×50 feet) was laid at about 800 yards from the firing pier to measure running depth. Firing of magnetized torpedoes over the loop and comparison with internal depth and net penetration measurements indicated that excessive errors were introduced from variation in the separation of the two sides of the loop and from variation in the dipole separation of the magnetized torpedoes.

¹ L. S. Edson, "Acoustic Torpedo Range, U. S. Naval Torpedo Station, Newport, R. I.," NOLR 898; April 15, 1946.

In April, 1943, the magnetic method was abandoned and efforts were concentrated on an acoustic method using hydrophones to measure speed, deflection from a straight course along the initial torpedo heading, and running depth at one point in the trajectory. By the middle of September, a temporary range of seven hydrophone stations had been constructed, placed in operation, proven more accurate than the magnetic system, and sanctioned as the official proving method for daylight ranging of electric torpedoes. In December, proving of steam as well as electric torpedoes was ordered. The initial acoustic tracking range, torpedo (TRT), consisted of a series of 12 hydrophone arrays planted on the bottom and cabled back to a selector switch, amplifier, filter, demodulator, and a recorder. Fig. 2 shows a general arrangement of the initial installation. Two directional (10° beam width) line hydrophones were mounted on a single support at each station. The plane of maximum sensitivity of the speed hydrophone was vertical and perpendicular to the center line of the range, or the line of fire; the plane of maximum sensitivity of the deflection hydrophone was also vertical, but at 45° to the line of fire (Fig. 3). Accurately located hydrophone stations permitted computation of average torpedo speed from the measured time interval between transits of speed planes of successive stations. Knowing the speed from either the preceding or succeeding 1000 yards, deflections from the range centerline could be obtained from the torpedo transit time in crossing the two beams at a single station. The running depth of the torpedo was determined at the 1000-yard point by two hydrophones mounted so that their planes of maximum acoustic sensitivity were at right angles to each other in the vertical plane (Fig. 3) and depth was computed in the same manner as deflection. In order to determine whether the torpedo was traveling on the left or the right side of the center line of the range, a two-channel amplifier was employed; one channel for speed hydrophone signals, and one for deflection hydrophone signals. The outputs of both channels were connected in balanced opposition and fed to a pen recorder. With no signal the pen is centered, with a speed signal the pen would deflect to the right, and with a deflection signal it would kick left. Thus, if the deflection hydrophone beam (Fig. 3) were crossed first, the torpedo would be to the left of the range centerline and the pen would swing left first providing a polarity indication from which the operator could tell at a glance on which side of a hydrophone station a torpedo had passed. Note that the zero station provided only an indication of the time of starting, and employed a nondirectional hydrophone.

The signal on which the TRT system measurements are based is the noise generated by cavitation of the propellers and other surfaces of the torpedo, in a band of frequencies in the high audio range. No signal generator is required to be carried in the vehicle, although at times noise generators have been added to increase the detectability of very quiet weapons.

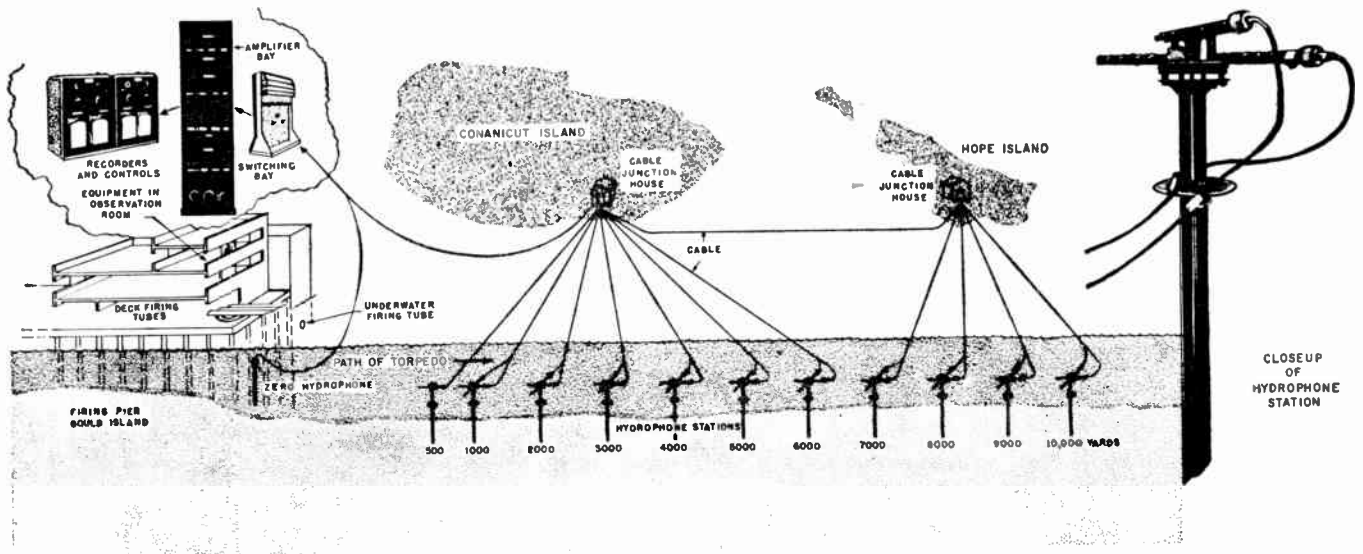


Fig. 2—General arrangement of the Newport Tracking Range, Torpedo (TRT). The hydrophones shown schemetically at each range mark are mounted on piles driven into the floor of the Bay.

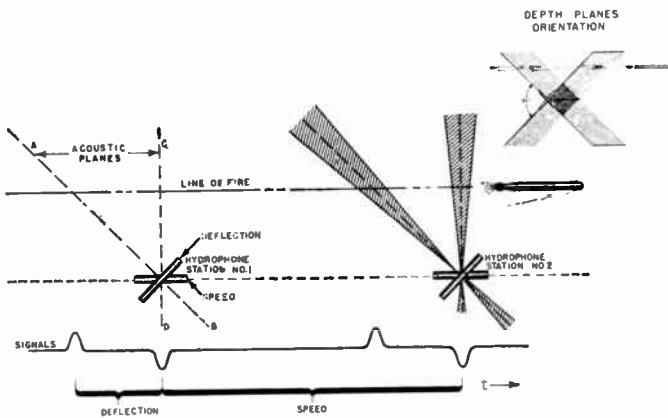


Fig. 3—TRT acoustic plane orientation. As the torpedo crosses each hydrophone beam a signal is generated. Time differences between planes and hydrophone stations can then be employed to compute the speed and deflection of the torpedo. Depth is obtained by orienting the hydrophones so that their planes are vertical.

The success of this first installation led to similar installations at the Naval Torpedo Station, Keyport, Wash.² and the Naval Torpedo Testing Range, Piney Point, Md.³ The Newport and Keyport ranges are still in active operation although improvements have been incorporated, notably the addition of preamplifiers at the hydrophone installations which permitted tracking of quieter torpedoes at greater deflections from the range centerline.

The TRT system, while a major step forward in underwater missile tracking, had limitations which were apparent as weapon systems became more sophisticated in the post World War II period. Lack of initial transient data, the long time interval between data

points, and the limited operating depths were the most serious limitations. Accordingly, development of four new tracking installations was initiated. Each of these was tailored to the problems of the particular activity developing it and therefore differed in either concept or implementation.

The Morris Dam Sound Range

The first system developed was the Morris Dam installation of the Naval Ordnance Test Station, Pasadena, Calif.⁴ It was quite small, only 750 feet long by 150 feet wide by 75 feet deep, and was constructed to measure the initial trajectory of a weapon immediately after water entry from a simulated aircraft launching. Fig. 4 shows the general layout of the sound range.

The tracking is accomplished internal to the array by installing a pulse generator within the vehicle to be tracked and measuring the differences in times of arrival of the sound pulse at the various hydrophones. The locus of any point on the trajectory from which sound is emitted is then the intersection of four hyperbola of revolution.

The array consists of twenty-four hydrophones spaced 150 feet apart horizontally and 75 feet vertically, in six groups of four hydrophones each. The hydrophones (omnidirectional) are cabled to junction boxes and then back to shore. Each hydrophone signal is then independently amplified and the output recorded on a 16-channel oscillograph. A 16-position selector switch is provided to allow monitoring of any of the hydrophones. Recording is usually done on a suitably chosen array of 12 hydrophones.

The sound pulse is generated by installing a special handhole cover in the vehicle which contains six ex-

² E. M. Lindsay, "Acoustic Torpedo Range, U. S. Naval Torpedo Station, Keyport, Washington," NOLR 925; March 1, 1945.

³ E. M. Lindsay and J. H. Treadwell, "Acoustic Torpedo Range, Naval Torpedo Testing Range, Piney Point, Md." NOLR 979; August 20, 1945.

⁴ U. E. Younger, ed., "The Sound Range for Underwater Trajectory Measurement at Morris Dam Torpedo Range." NOTS Tech. Memo. No. 502-13; September 14, 1949.

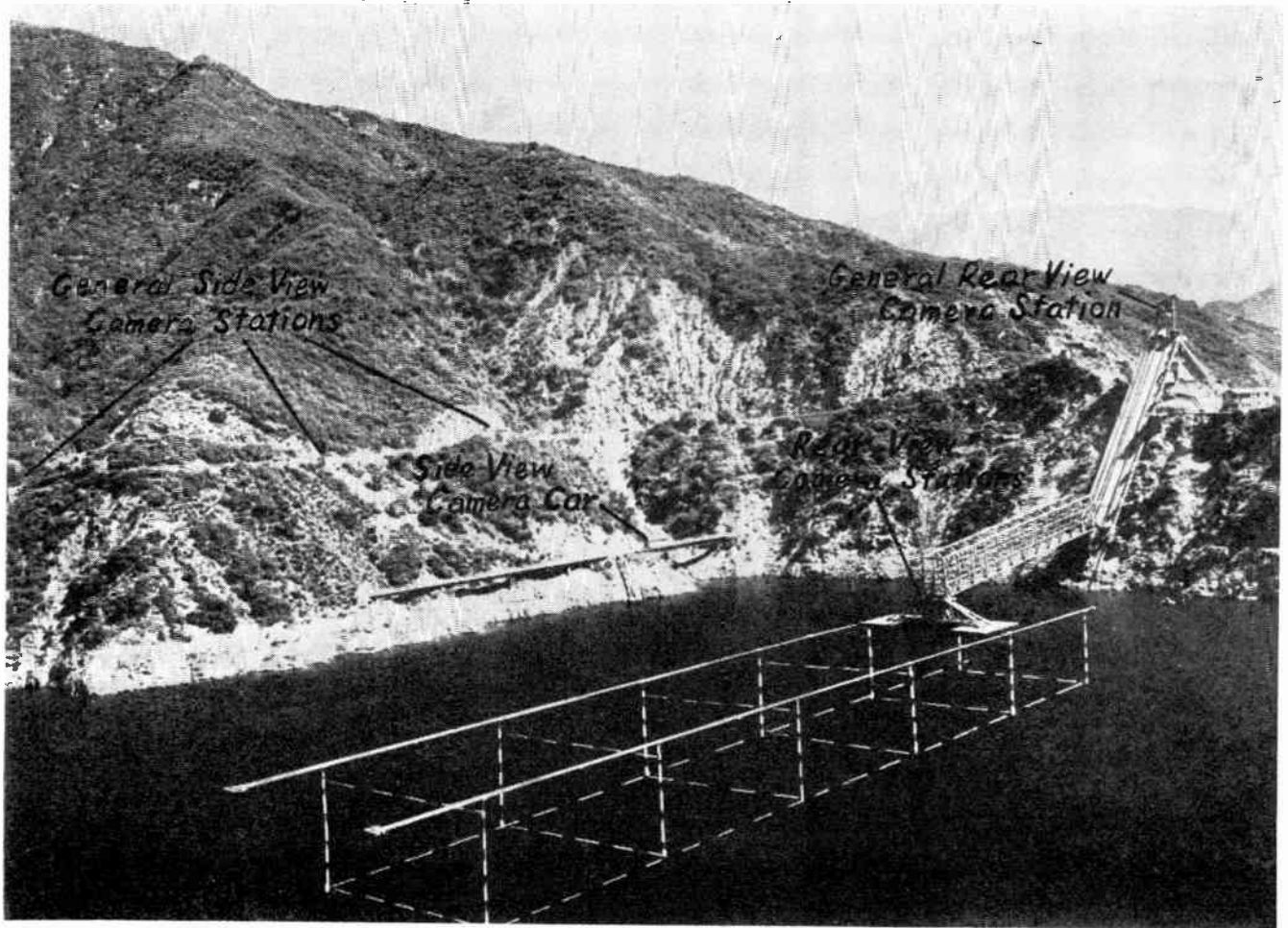


Fig. 4—Morris Dam range installation. The underwater grid provides the general arrangement of the tracking hydrophones. The device at the right of the photograph is the launcher.

plusive squibs detonated one at a time by a spring-operated oil-damped timing switch and a 1.5-volt flashlight battery mounted in the timing-switch housing on another standard handhole cover.

The time of arrival of the pulse on each channel is then manually scaled from the oscillograph tape, using the arrival of the first pulse as the reference. These times (reduced to relative distances in feet) are then inserted in a mechanical model of the sound range (Fig. 5) on a trial and error basis by an operator moving a model sound source through the range until correspondence is obtained between fixed pointers and pointers which are fastened to scaleboard weights. Sound source coordinates are then read from carriage scales to obtain the torpedo position.

The Piney Point Hyperbolic Torpedo Tracking System

The Naval Ordnance Laboratory was faced with a somewhat different problem than the Naval Ordnance Test Station. The staff was interested in the trajectory of the final attack phase of a homing torpedo and therefore built a submerged pontoon-mounted installation at

Piney Point, Md. Its first effort was a mathematical analysis of the geometry of hyperbolic and spherical coordinate systems including the effect of various array parameters on system accuracy.⁵ Since the initial objectives were to provide tracking information of torpedoes over the range of 250 feet to within 15 feet of the target with maximum attainable accuracy,⁶ a four-hydrophone coplanar array and hyperbolic coordinates offered superior over-all advantages.

The system consists of an ultrasonic beacon mounted in the torpedo, four hydrophones mounted as shown in Fig. 6, a signal processing system, a recorder employing chemically treated, electrically sensitive paper, and a set of curves for converting time differences to torpedo position.

The target array consists of a metal pontoon 12 feet in diameter and 50 feet in length, with three hydro-

⁵ D. J. Bordelon, "Analysis of Three Dimensional Hyperbolic Tracking Systems," NOL NavOrd Rep. No. 2718; December 10, 1952.

⁶ C. E. Goodell, "Piney Point Hyperbolic Torpedo Tracking System," NOL NavOrd Rep. No. 2728; March 10, 1953.

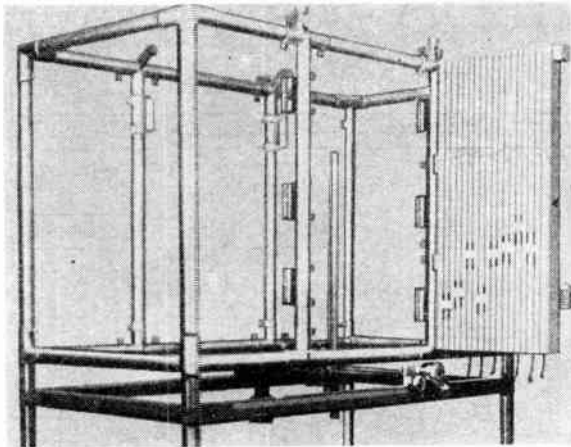


Fig. 5—The Morris Dam sound range computer converts time differences to torpedo coordinates by manually positioning a probe attached to wires until it satisfies all the measured values. The position coordinate of that point is then read from scale weights.

hydrophone (No. 2) pulse and each of the other three pulses to a digital recording system. A 50-kc oscillator feeds a gated amplifier in each of the time difference evaluators (TDE). The amplifier is gated for a period equal to the time difference between the two incoming signals. Thus, the output of each TDE is a group of 50-kc pulses corresponding in number to the time difference. Each TDE is followed by a set of eight cascaded Eccles-Jordan flip-flops, functioning as a binary counter to count the 50-kc pulses and record them in binary notation. The count is then manually converted to time differences and the position of the vehicle obtained from a set of graphs. Special features in the signal processing system include a tuned cable from the hydrophone to the band-pass amplifier, AVC action to provide a logarithmic amplifier voltage response, thyatron gates in the TDE, a sequence identification (Eccles-Jordan) circuit to establish which pulse arrived first in time for each hydrophone pair, and a one-shot multivibrator, which serves as a reset delay in the event that only one pulse is received in a given interval.

A later version⁸ of this system is mounted on a target submarine, and a number of modifications are incorporated: The pulse repetition rate of the vehicle-borne transmitter is reduced to 10 per second and its reliability is improved by projector and packaging modifications. A total of twenty-four hydrophones (three sets of eight each mounted on the top and port and starboard sides of the submarine) are employed in the array to eliminate signal obscuration by the submarine hull.

Noise gates added to the hydrophone amplifiers turn them off when ambient noise is comparable to, or greater than, the signal level.

Data reduction is performed automatically through the addition of a velocimeter, a digital computer, a Clary printer, and two plotters.

The Newport Hyperbolic Coordinate Torpedo Tracking System

The Naval Underwater Ordnance Station, Newport, (successor to the Naval Torpedo Station, Newport), in deciding to improve their range complex, were interested in a somewhat different problem than either NOTS or NOL. The problem here was to measure the trajectory after actual launch (rather than simulated as in the NOTS facility) by either a submarine, surface vessel, aircraft, or launching facilities on a firing pier. Further, at Newport the portion of the trajectory of interest was the over-all course of the powered torpedo. This information is essential to the design, not only of the weapon but also its launcher and fire control equipment. The system consists of a pulsed sound source mounted in the

⁸ J. E. McGolrick, P. A. Trout, K. J. Meese, and D. E. Gerster, "A Three Dimensional Hyperbolic Tracking System," NOL NavOrd Rep. No. 4297; June 1, 1956.

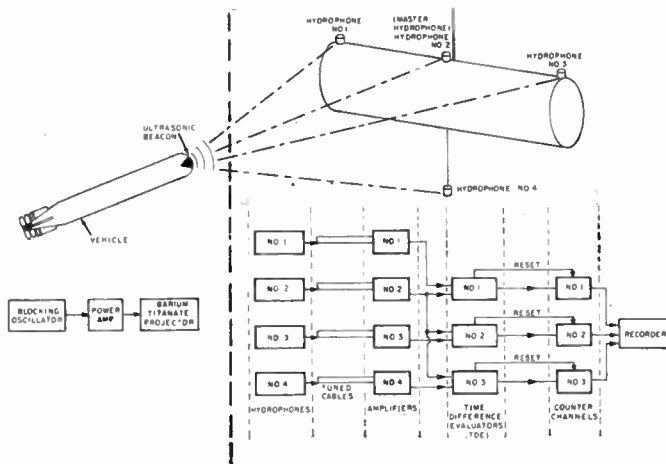


Fig. 6—Piney Point hyperbolic coordinate system. The recorded binary digits output of this system are converted to missile coordinates manually with special graphs.

phones mounted in a line along the top of the pontoon and one suspended below the pontoon.

The ultrasonic beacon carried in the vehicle⁷ has a basic frequency of 250 kc, approximately a one-millisecond pulse width, and a pulse repetition rate of approximately 20 per second. The unit has a peak electrical power of about 20 w and a useful life of approximately one hour. The transmitter, when assembled, is a 7½-inch long cylinder with a 1¾-inch O.D. except for a 3⅜-inch flange on the front end where the projector is housed.

The method of processing this signal after receipt at the array hydrophones and tuned amplification (transducer to detector, 100-db voltage gain) consists of converting the time difference of arrival between the master

⁷ "The Sonic Transmitter XS-2B," NOL NavOrd Rep. No. 4119; August 1, 1955.

torpedo, eighteen bottom-mounted hydrophones located in 500-yard and 1000-yard squares as shown in Fig. 7, a signal processing system, a 14-channel oscillograph recorder, a 16-channel magnetic-tape recorder, a tape playback to punched-card conversion system, an analog computer with oscilloscope output, and a digital computer. The signal processing system, the data reduction to punched cards equipment, and the analog computer were developed to specification by Anderson-Nichols and Company.⁹ The digital computer is a card programmed general purpose IBM 602A calculating punch.

This system, because of the variation in launcher position and location and the extended trajectory of interest, is the largest (in baseline and included water area) system currently in existence, but because of water depth limitations is constrained to two dimensions. Initially, a prototype version of the system (consisting of a detonating squib sound source in the torpedo, the 1000-yard baseline hydrophone arrays, amplifiers, and a six-channel oscillograph recorder) was constructed to prove the system feasible and to serve as an interim facility. As with the other hyperbolic coordinate systems, time difference of arrival of the sound pulses was used to determine the torpedo coordinates. The time differences were manually scaled from the oscillograph records, punched on cards, and the trajectory points including speed and acceleration were computed with the digital computer. In the production version considerable refinement was added and a spark gun sound source developed for use in the vehicles with power available.

Notable improvements include a system for pulsing each hydrophone as a projector prior to a run and receiving on the other hydrophones to determine sound velocity, magnetic tape recording, a medium accuracy analog computer to permit the range operator to follow the torpedo during the run (thus facilitating recovery), and an analog-to-digital converter to provide a punched card output. In addition, 500-yard baseline arrays were added in a deeper water region back from the firing pier to provide an air drop and submerged submarine launching area.

The Dabob Bay 3-D Tracking System

At the Naval Torpedo Station, Keyport, Wash., the problem was to determine the track of homing torpedoes. The Applied Physics Laboratory, University of Washington, was given the task of constructing a three-dimensional tracking range¹⁰ for this purpose, in an area of 600-foot depth water in Dabob Bay where NTS proved and exercised homing torpedoes. The concept employed by APL is somewhat different from the other

⁹ "Hycor Range Manual, Naval Underwater Ordnance Station Contract No. 298s-9674," Anderson-Nichols and Co., Boston, Mass.; April 21, 1955.

¹⁰ "An Introduction to the Three-Dimensional Underwater Tracking Range," Appl. Phys. Lab., Univ. of Wash., APL/UW/TE/58-3; January 31, 1958.

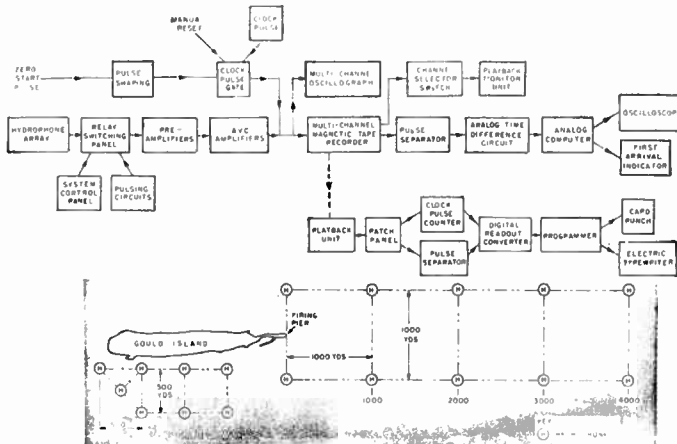


Fig. 7—Newport hyperbolic coordinate system. A two-dimensional tracking range with both digital and analog output.

installations in that interrogation and transponding, a short baseline array, and a spherical coordinate system were chosen rather than pulses, long baselines, and hyperbolic coordinates.

The system (Fig. 8) consists of nine arrays, each with four receiving hydrophones and two projectors, a 330-kc and 190-kc transmitter for interrogating both a torpedo and its target, a 250-kc torpedo transponder for responding to the interrogation, signal processing equipment, a computer to convert elapsed time from interrogation to returned signal to coordinates, a timer to provide initial time and elapsed time for each interrogation-respond-compute cycle, a digital readout and printer, and four strip chart recorders. The latest improvement to this system is the addition of an accurate synchronous clock which is started simultaneously in the torpedo and at the range computer, permitting elimination of the interrogation portion of the tracking cycle.

San Clemente Portable Range

The previously discussed tracking systems (with the exception of the submarine-mounted NOL system) have all been designed for continuous use in sheltered waters. A recent development, conducted by the Minneapolis-Honeywell Company¹¹ for the Naval Ordnance Test Station, is designed as a portable range for use in deep water off San Clemente Island to measure the underwater trajectory of shore-launched, air-to-water missiles.

The system (Fig. 9) consists of an array of four small boats each with a suspended hydrophone and a radio-telemetry transmitter, a squib sound source in the missile, and a shore station with the telemetry receivers feeding a multichannel magnetic-tape recorder. The small boats are towed into the desired position and set adrift. During the test they are located by triangulation with known position cinetheodolites on shore. Upon water entry, the missile fires a series of squibs and the

¹¹ "Range Instrumentation Activity, Honeywell Military Products Group," Minn.-Honeywell Regulator Co., Ordnance Div., Seattle Dev. Lab., Tech. Doc. No. 5358 S; April, 1958.

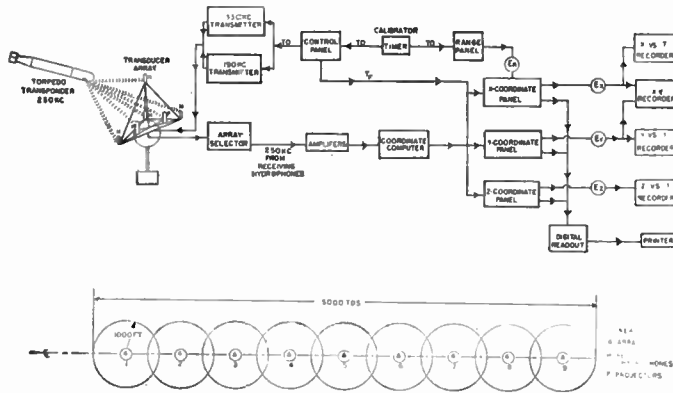


Fig. 8—Dabob Bay 3-D tracking range. A three-dimensional tracking range employing a spherical coordinate system.

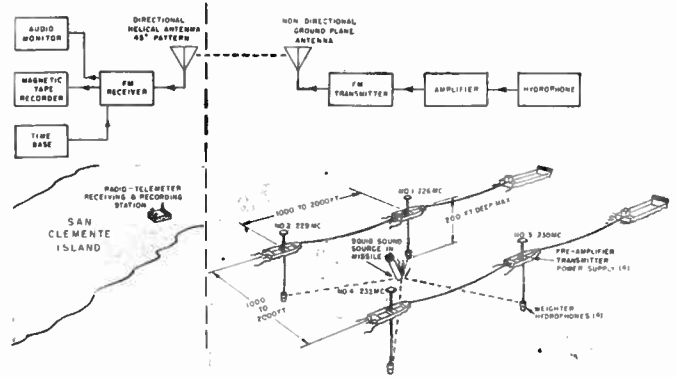


Fig. 9—San Clemente portable range. A range installation designed for missile re-entry trajectory in deep water.

sound is received by the hydrophones and transmitted back to shore for magnetic-tape storage and later processing.

Target-Centered Systems

In addition to the systems designed primarily for missile-tracking in the development and test phase, a number of systems for shipboard use in detecting and tracking attacking torpedoes in order to avoid them have been developed. These systems, which in general operate independent of vehicle-mounted signal sources, are outside the scope of this article for security reasons.

Development Efforts

Having examined our existing external tracking systems it might be instructive to speculate on the course of future development effort. Clearly there are two major directions which might be followed: 1) improved portable systems, and 2) improved fixed site ranges. It is the conviction of the authors that both will play a necessary part in weapons system development. The

portable system approach will be required for open ocean, arctic, and special weapons testing. The fixed site range will be necessary for the major trajectory measurement effort because of its higher inherent accuracy, the more carefully controlled test conditions possible, the potential precision knowledge of the medium at a continuously utilized site, and the inherently lower cost per test of a general purpose facility.

The growth of the weapons system concept provides a basis for predicting another characteristic of future systems. That is, the ability to track simultaneously not only the missile but also its launching platform and target. Since these vessels will require room to maneuver, it is probable that future range systems will encompass a much larger volume of water and will be much more closely integrated with in-air tracking systems. Electronically, the missile-borne projectors and signal processing equipment will take advantage of improvements in low-level signal recognition techniques in order to reduce the amount of underwater equipment and still provide coverage over a larger water volume.

Design of Telemetering Networks for Flood-Control Reservoir Operation*

FRANCIS P. HANES, ASSOCIATE MEMBER, IRE†

Summary—Automatic equipment is described to fulfill a requirement for optimized operation of reservoirs and other flood-control structures by allowing rapid and unattended reporting of river levels and precipitation amounts from remote points. The emphasis is on design of telemetering networks and system engineering for VHF-radio transmission of the data. A recent development in precipitation telemetering is also described.

INTRODUCTION

RESERVOIRS for flood control are located in natural geographical drainage basins at strategic points for storage of rainfall runoff in order to prevent, or materially reduce, downstream flooding. Effective operation is based on storage and release of flood-producing runoff so as always to maintain sufficient available storage capacity while not producing downstream flooding by excessive releases from the reservoir.

The collection and analysis of hydrologic data are basic to the formulation and implementation of any operating plan for a flood-control reservoir. Current reports of the stages of main and tributary streams above and below the dam, as well as amounts of rainfall occurring throughout the drainage basin, are of utmost importance. In some cases these data can be obtained by means of standard, commercially available equipment operating on simple wire-line circuits and utilizing remote-registering systems of the selsyn-position-motor type or other similar types of equipment. Frequently, however, lack of reliable communication facilities results in the need for radio-gauging installations to furnish the required information.

REQUIRED CHARACTERISTICS

In general, the required characteristics of a radio-gauging network of the type under consideration are reliability, accuracy, and simplicity. Advantage may be taken of the lack of a requirement for speed, since there is usually no significant change in the measured variable over a period of a minute or so. Normally, only a single type of information is required from any one gauging station; if more than one is required (*e.g.*, waterstage, precipitation, and perhaps temperature), they can be transmitted in sequence, so that simplicity of the communication system can be maintained.

* Original manuscript received by the IRE, December 18, 1958; revised manuscript received, February 18, 1959.

† U. S. Army Engineer Waterways Exper. Station, Vicksburg, Miss.

The fundamental quantities involved are precipitation amounts (*i.e.*, rainfall or water equivalent of sleet or snow) and stages (water level above some gauge datum) of rivers and streams. The collection of these hydrologic data is logically made at the reservoir location. Thus a dam-operating room may be used as a collection center for a particular drainage-basin hydrologic network. Separate voice-communication facilities are provided by means of telephone and/or radio between these outlying operating rooms and a District office. Here the data are processed to formulate operating schedules for the entire reservoir network within a District.

During normal conditions, a reading from each station every four to six hours will suffice to monitor changes in river stage and amount of precipitation, if any, that have taken place since the last set of readings. However, on fast streams in mountainous terrain, emergency conditions may obtain in a period of a few hours during a heavy rain. It is primarily for this reason that an interrogation-type network is provided in order to allow readings to be obtained every few minutes, if needed. The network to be described later was designed to meet the needs of a reservoir located in mountainous terrain in a drainage basin characterized by flash-flooding conditions.

TYPES OF SYSTEMS

Telemetering systems may be simple or complex; *i.e.*, they may be designed to use a simple time-programmed, keyed-carrier mode of operation or the more complex interrogation plus automatic-programmed-type operation. An interrogation-type system has the obvious advantage of making the information available at any time, whereas the time-programmed system is necessarily a compromise.

Since simplicity and accuracy are prime requisites, the coding, or conversion of the measured variable to a form suitable for transmission, is invariably accomplished with some form of digital device. A commercial version of such coding device that has been found satisfactory for this purpose is the Shand & Jurs Telepulse transmitter. This device (Fig. 1) is basically a shaft analog-to-digital converter. The shaft rotation is produced by a float-and-counterweight assembly; the conversion to pulse-count information is made by the telepulse transmitter. The digital-pulse output consists of two trains of accumulative pulse-count information.

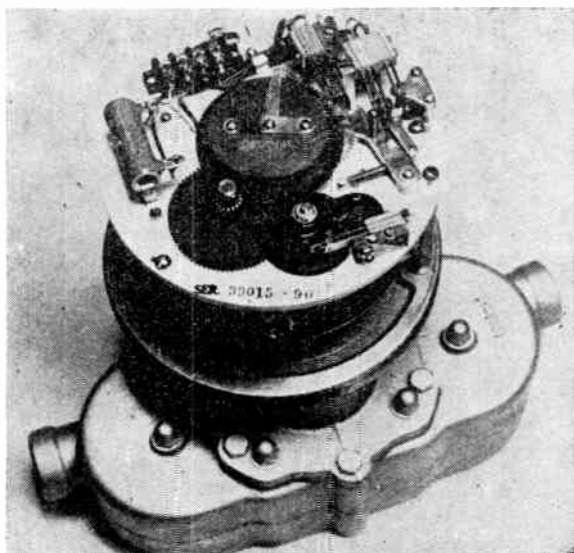


Fig. 1—Shand and Jurs river-stage teletyping device.

When used with the indicator (Fig. 2) the two groups of pulse-count information are channeled respectively to dials used to indicate feet (on one dial) and hundredths/foot (on the other dial). The range of the transmitter is 0- to 60-foot stage, with an accuracy and least count of 0.01 foot.

Several devices for telemetering precipitation information have been developed by the Corps of Engineers, U. S. Army. One such device utilizes a Shand & Jurs water-stage transmitter in conjunction with a standard weighing-recording rain gauge. A differential transformer is used as a motion-pickoff element in the rain gauge, and a servomechanism is used to convert the differential transformer output to shaft rotation that has sufficient torque to drive the water-stage transmitter. Another telemetering device is being developed for low-drain applications and will be described later.

The normal mode of operation for a typical interrogation-type network consisting of six to eight gauging stations makes use of audio-tone interrogation and keying techniques. Speed of operation of the coding devices is normally of the order of 8 pulses per second, thus necessitating the use of frequency-shift-tone equipment (e.g., Motorola voice-frequency-carrier equipment) for reliable transmission of the digital pulse-count information. In operation, the radio transmitter is modulated with the center-frequency tone, while the pulse information is conveyed by dry-contact keying to produce a 30-c frequency shift representing each pulse. This part of the system is similar to typical space-mark teletype transmission. At the receiving end, a matched frequency-shift receiver is utilized to restore the pulse information to dry-contact closures reproducing the pulse output from the telemetering devices. Interrogation of one of several stations in a given network is done with either modified-amplitude-tone equipment or

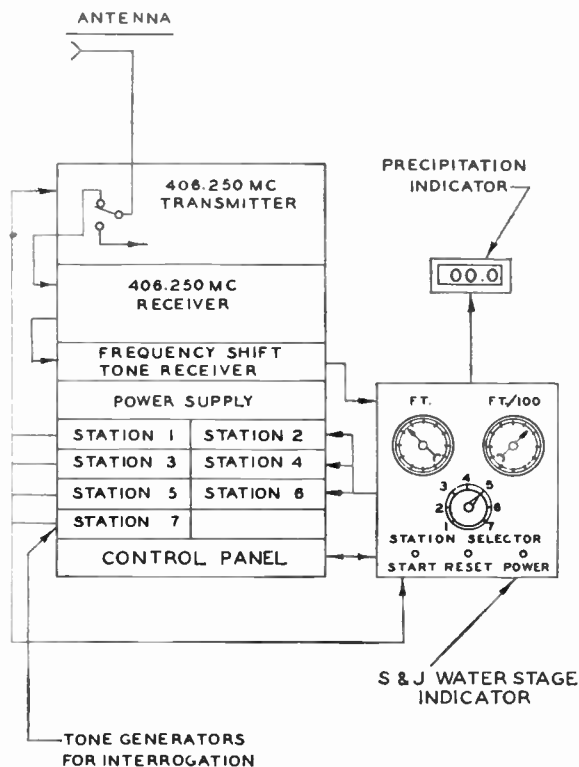


Fig. 2—Central station equipment.

multi-tone selective-signaling equipment such as the Motorola "Quik-Call." In general, accuracy requirements specify measurements to within 0.01-foot stage and 0.1-inch precipitation.

FREQUENCIES

Frequencies available to government and non-government agencies for transmission of hydrologic and meteorologic data are shown in Table I. Assignments of frequencies to government agencies for this purpose are made by the Interdepartmental Radio Advisory Committee (IRAC) through coordination with the FCC.

TABLE I
FREQUENCIES AVAILABLE FOR TELEMETERING
HYDROLOGIC DATA

Megacycles					
169.425	170.225	171.025	171.825	406.050*	412.450*
169.475	170.275	171.075	171.875	406.150*	412.550*
169.525	170.325	171.125	171.925	406.250*	412.650*
169.575	170.375	171.175	171.975	406.350*	412.750*

* Primarily for relay use.

Frequency assignments to non-government agencies are coordinated through the U. S. Weather Bureau. Gauging-station transmitter power of 50 watts is usually adequate and is normally used as a basis for design except for low-drain, battery-operated stations used in special

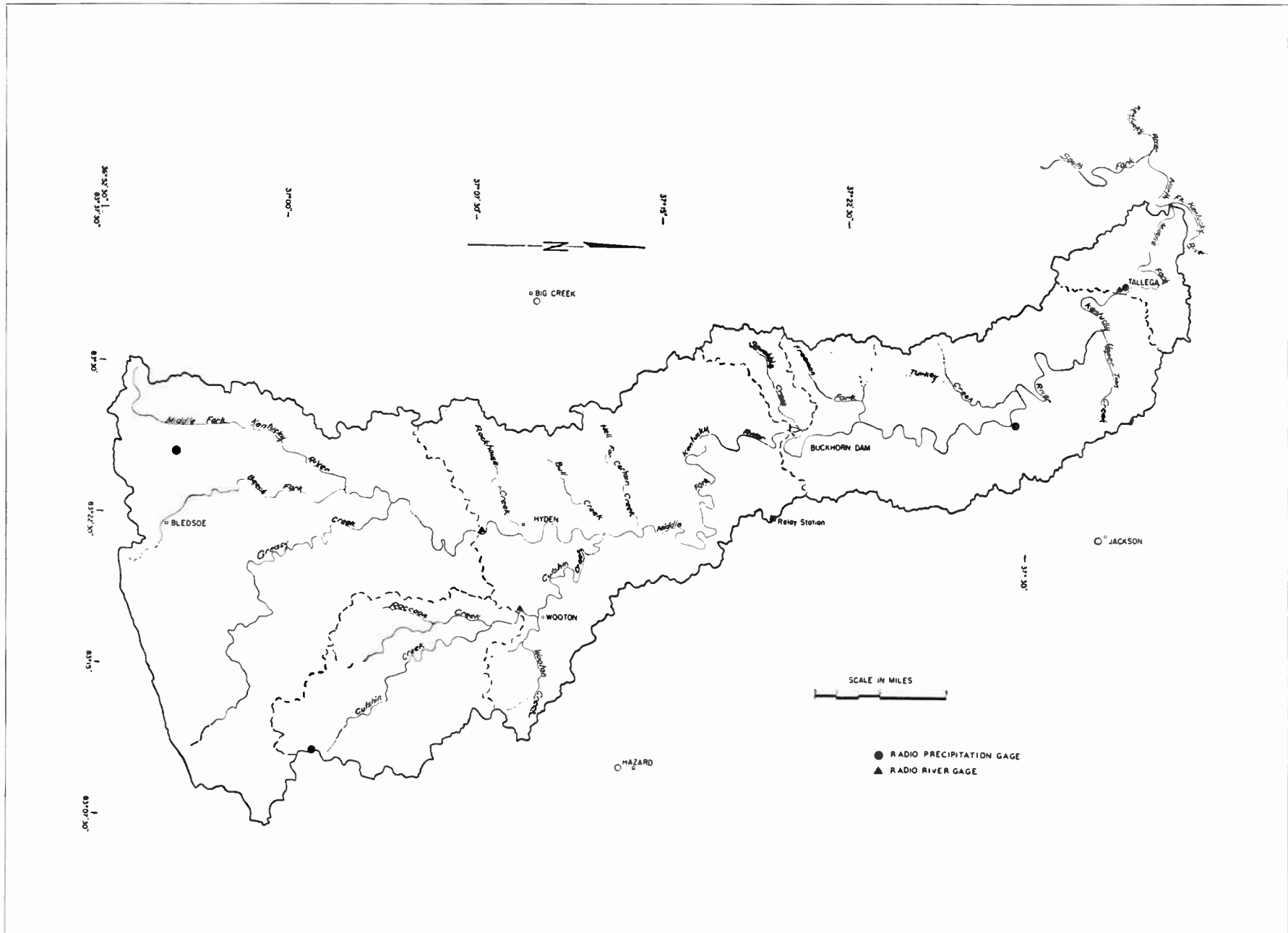


Fig. 3—Map of drainage basin (Middle Fork Kentucky River).

cases. Commercially available high-gain antennas for this service include the corner-reflector and yagi designs.

NETWORK DESIGN

A typical radio telemetering network is that designed for Buckhorn Reservoir in eastern Kentucky, now being constructed by the U. S. Army Engineer District, Louisville. The reservoir is located on the middle fork of of the Kentucky River. A map of the 559-square-mile drainage basin for the reservoir area is shown in Fig. 3, which also shows locations of the three stage and four precipitation stations, as well as a relay station located on a ridge near the dam. Because of the shape of the drainage basin, it was possible to use a bidirectional antenna at the relay station. Signal strengths from the gaging stations to the relay station (169.575 mc) and between the relay station and the dam (406.250 mc) were calculated according to Bullington.¹ Since the paths from the river-gauging stations are non-line-of-sight, measurements of these paths were made for confirmation of the computed signal strengths. Table II shows a comparison of the computed and measured values.

TABLE II

COMPARISON OF CALCULATED AND MEASURED RADIO SIGNALS. BUCKHORN RESERVOIR TELEMETERING NETWORK. MEASUREMENTS MADE TO RELAY STATION, ANTENNA ELEVATION 1620, 50-W POWER

Station	Antenna Elevation*	Received Signal, Microvolts	
		Calculated	Measured
Hyden†	908	6	4-5
Wooten†	920	7	5-6
Tallega‡	710	15	15
Dam‡	920	22	25

* All elevations are in feet above mean sea level.
 † Corner-reflector antenna at gauging stations, bidirectional antenna at relay station.
 ‡ Corner-reflector antenna at both gauging station and relay station.

Equipment for a typical river-gauging station is shown by the block diagram in Fig. 4. Central-station equipment at the dam-operating tower is shown in Fig. 2. The system will operate "on-call," using combinations of audio tones to activate a particular station in the group. Information from the gauging stations will be returned by means of frequency-shift tone equipment. Voice communication will be available on the two-way FM-radio link (see Fig. 5), thus allowing communication with mobile units in emergencies. Precipitation keyers will be either the Hanes² keyer, one model of which is shown in Fig. 6, or the device described below.

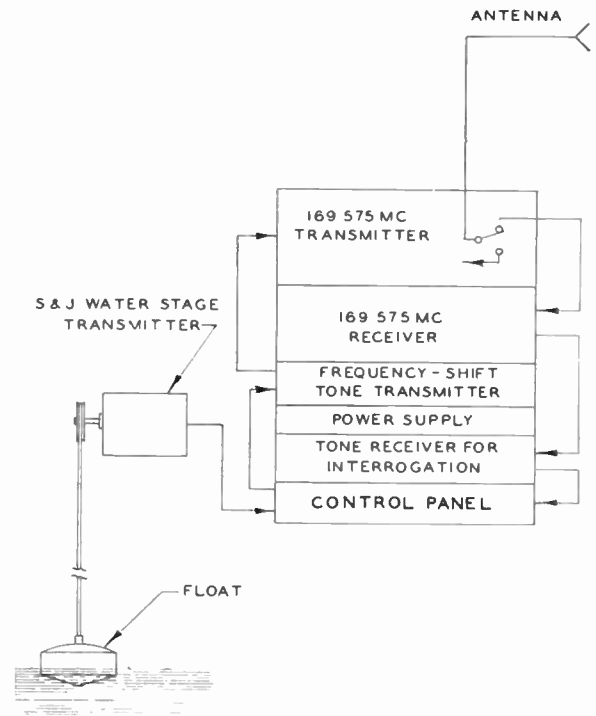


Fig. 4—Typical river-gauging station equipment.

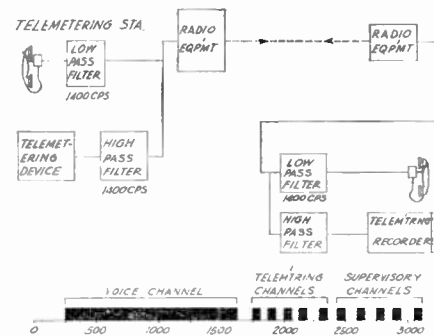


Fig. 5—Use of audio pass band for telemetering and voice communications.

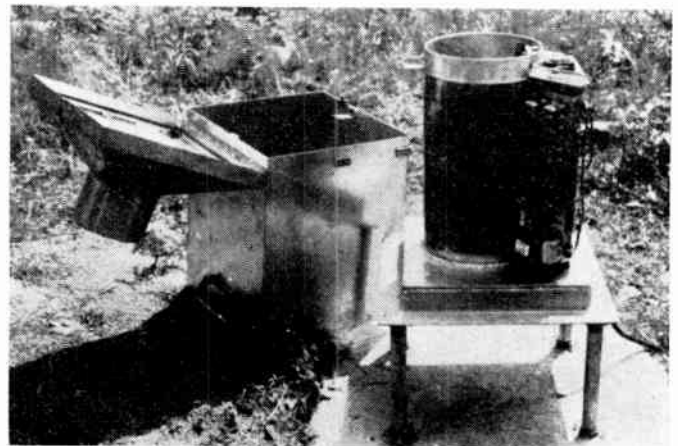


Fig. 6—Hanes precipitation telemetering device installed in telemetering network, Bluestone Dam, W. Va.

¹ K. Bullington, "Radio propagation at frequencies above 30 megacycles," Proc. IRE, vol. 35, pp. 1122-1136; October, 1947.

² F. P. Hanes, "Telemetering hydrologic data," J. Waterways and Harbors Div., A.S.C.E., vol. 83, No. WW3, paper 1365; September, 1957.

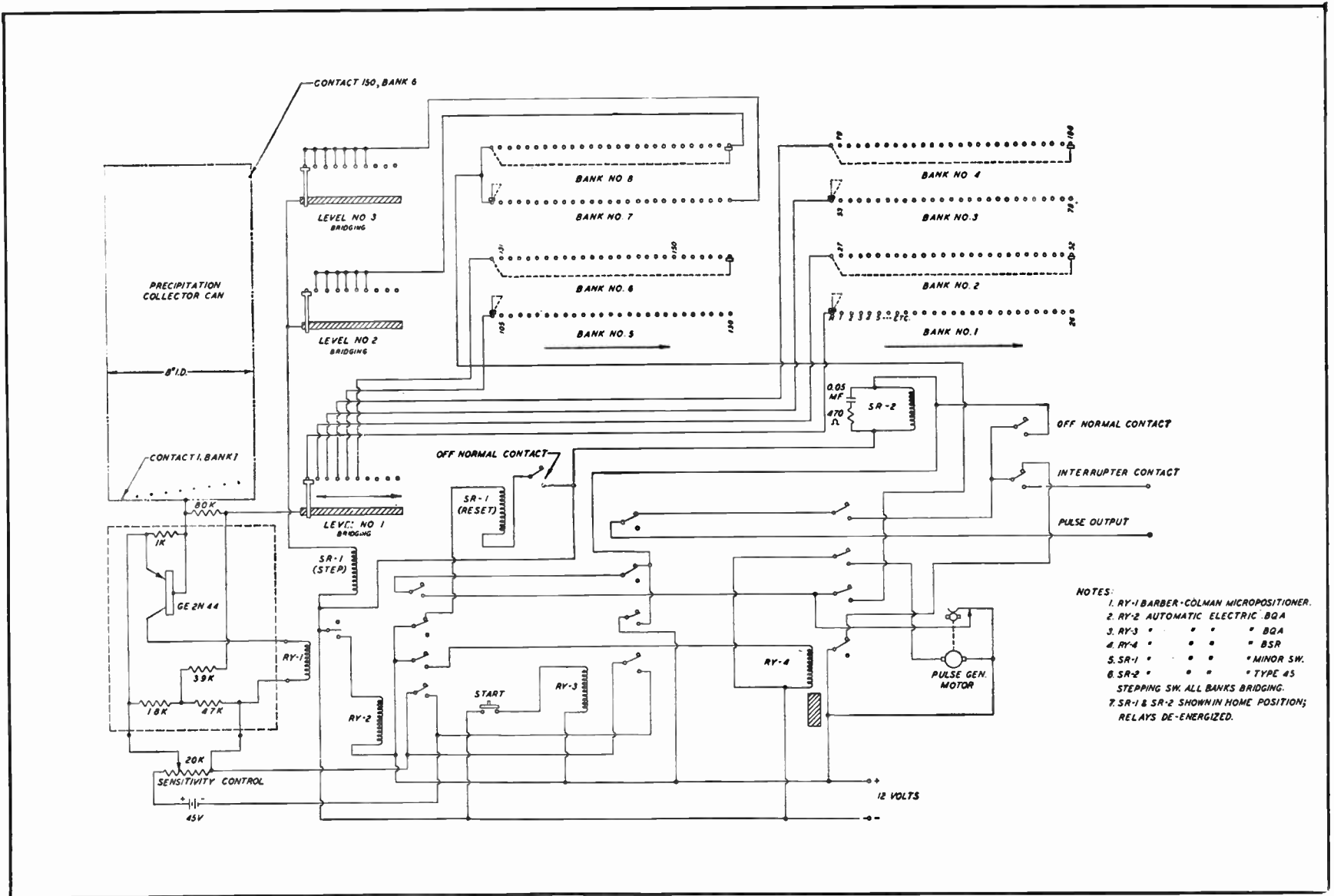


Fig. 7—Electrical circuit diagram of low-drain precipitation telemetering device.

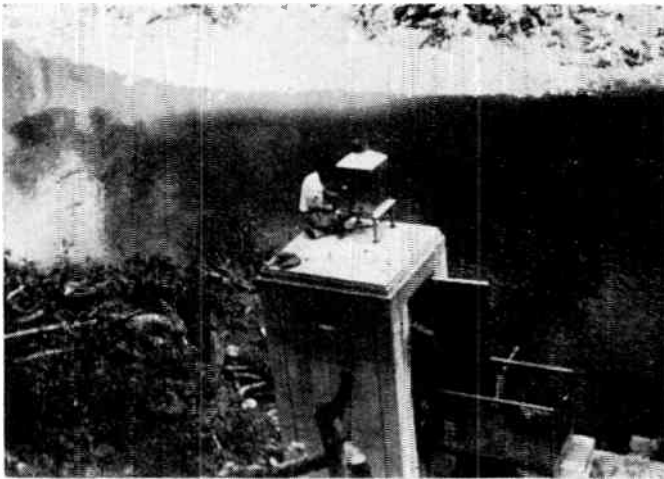


Fig. 8—River-stage station, Bluestone Dam, W. Va.

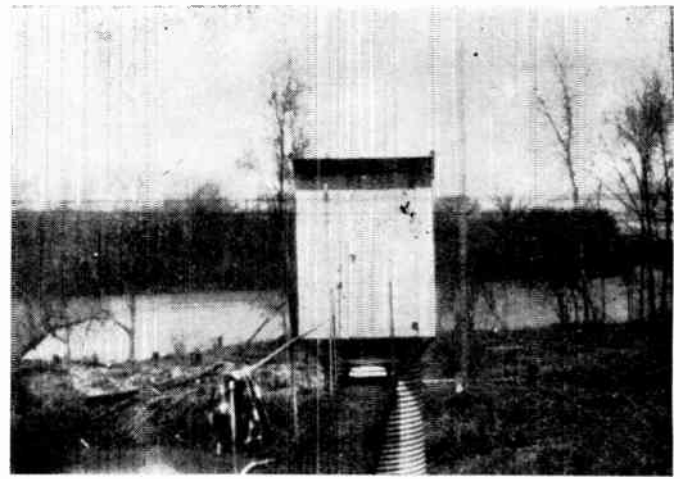


Fig. 9—River-stage telemetering station, Wabash River near Vincennes, Ind.

LOW-DRAIN PRECIPITATION KEYS

In order to permit the telemetering of precipitation data from remote locations by means of battery power, a keyer was developed to measure precipitation in terms of digital pulse-count information. The device consists of a can fitted with insulated contacts spaced at intervals of 0.1 inch in a spiral pattern around the can. A total of 150 contacts gives a capacity of 15 inches of precipitation. For winter use, the can will be fitted with a bottom section containing an antifreeze solution. A transistorized electronic relay is used to scan the contacts by means of a stepping switch until the circuit is lost through the water path. During the scanning cycle, pulses are gated out so that the number of such pulses represents the number of submerged contacts. Upon loss of the circuit through the water path, the pulse gate is closed and the scanning-stepping switch is caused to self-step to the home position. The pulse rate is approximately 8 per second and the count is accumulated on a simple electromechanical counter. A 12-volt nickel-

cadmium battery of 18-ah capacity is expected to operate the device for 300 days, assuming four operations per day. The drain per operation is 3.5 mah. A diagram of the electrical circuit for this low-drain precipitation telemetering device is shown in Fig. 7.

CONCLUDING REMARKS

In general, hydrologic-data telemetering networks are designed in accordance with specific requirements as to quantity and location of gauging stations as well as accuracy of data and mode of operation. At the present time, development work is concerned primarily with low-drain battery operation utilizing wind and/or solar generators to extend unattended operating time. Two gauging-station installations are shown in Figs. 8 and 9. Fig. 8 shows a river-stage gauging station to which has been added a precipitation telemetering device. Fig. 9 shows a river-stage telemetering station where the radio and telemetering equipment are housed in a structure on top of a pipe stilling-well.

An Electronic Control for Coronary Arteriography*

M. EICHER† AND H. C. URSCHEL, JR.†

Summary—An instrument has been developed for the control of diagnostic coronary arteriography. The method requires the automatic injection of a small amount of radio-opaque dye into the arteries at a preselected interval of the cardiac cycle, and the initiation of an X-ray exposure at the time of optimum coronary filling.

The sequence of events is initiated by either the peak of the electrocardiogram R-wave or the aortic systolic pressure wave. Electronic variable time-delay circuits independently open and close a solenoid valve determining the time of injection and the amount of dye injected. A third delay circuit actuates the X-ray exposure.

CORONARY arteriography is accomplished by injecting a radio-opaque dye through a catheter into the region of the coronary ostia and making X-ray exposures when the coronary arteries are filled with this contrast medium. In order for dye to be present in the coronaries at the proper phase of the cardiac cycle, it has been necessary to inject large quantities of the contrast medium which is undesirable due to its potential toxicity. Furthermore, in order to make the X-ray exposure at the moment of maximum coronary filling, it has often been necessary to employ a series of rapid-fire, multiple X rays which involve increased radiation hazard and special technique.

Therefore, the desire of the surgical investigator¹ performing arteriography has been to inject a minimum amount of dye at the optimum phase of the cardiac cycle and to have a single X-ray exposure follow automatically at a precise moment. In order to provide a dependable system for the investigators to perform these required functions, a reliable and flexible electronic instrument was developed which allowed a thorough study of the time intervals involved.

The coronary arterial dye injector is basically a solenoid valve controlled by two electronic time-delay circuits with a third delay provided for triggering the X-ray machine. Following the block diagram, Fig. 1, the initiating signal is obtained from either the R-wave peak of the electrocardiogram or the sharp rise of the aortic systolic pressure wave detected by a strain gauge through a catheter inserted in the abdominal aorta. After sufficient amplification, this signal, when keyed by the operator, enters a thyatron pulse shaper which triggers three variable time-delay circuits. The first delay determines the interval following the initiating signal at which the solenoid valve will open. This permits the flow of dye into the coronaries through a cardiac catheter inserted in the thoracic aorta. The second de-

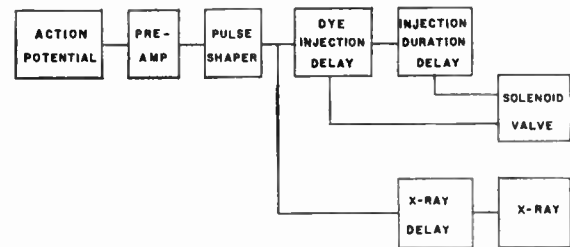


Fig. 1—Block diagram indicating functions of the electronic control.

lay, added to the first, determines the interval for which the valve will remain open, and controls the quantity of dye injected. At the conclusion of the third variable time delay, the X ray is triggered, making the exposure.

This entire sequence of events occurs automatically when the operator presses a switch at any time during the heart cycle. The following signal peak then initiates the arteriography sequence.

The flexibility of the control unit permits injections to be made at any preselected time, independent of heart rate, and over a wide range of signal amplitudes. Dye can be injected as frequently as desired, and the delay intervals repeat with an acceptable accuracy.

Signals from the three delay circuits are fed to the timer stylus of a Sanborn Twin-Viso electrocardiograph in order that the interval of injection and time of X-ray exposure be shown concurrently with the ECG.

The dye supply reservoir and valve system are shown with the control unit in Fig. 2. To inject a quantity of viscous dye through the catheter in a brief interval, pressure from a cylinder of carbon dioxide is applied to the surface of the dye contained in the 40-cc reservoir. The solenoid valve is connected to the lower end of this reservoir controlling the flow into the catheter. A calibrated scale gives an accurate measure of the volume of dye injected.

The volume is determined by the diameter and length of the catheter, the viscosity of the dye, and the arterial pressure. The pressure controlled by the regulator valve on the cylinder, combined with the variable duration control, enables the operator to deliver any desired volume. The upper pressure limit of the system is 100 psi but 5 psi was adequate for the dye, catheter, and durations used in most experiments.

In the design of the electronic control unit for coronary arteriography, the specifications required reliability, flexibility in range, and simplicity of operation for the surgeon using it. The series of events was to be initiated by a remote switch, to be completed and recorded within as short an interval as 0.3 second, and to be independent of heart rate fluctuations.

* Original manuscript received by the IRE, November 27, 1958. The opinions expressed in this article are those of the authors, and not necessarily those of the Navy.

† Naval Medical Res. Inst., Bethesda, Md.

† H. C. Urschel, Jr. and E. J. Roth, "Electronically controlled coronary arteriography," *Ann. Surgery*, accepted for publication, November, 1958.

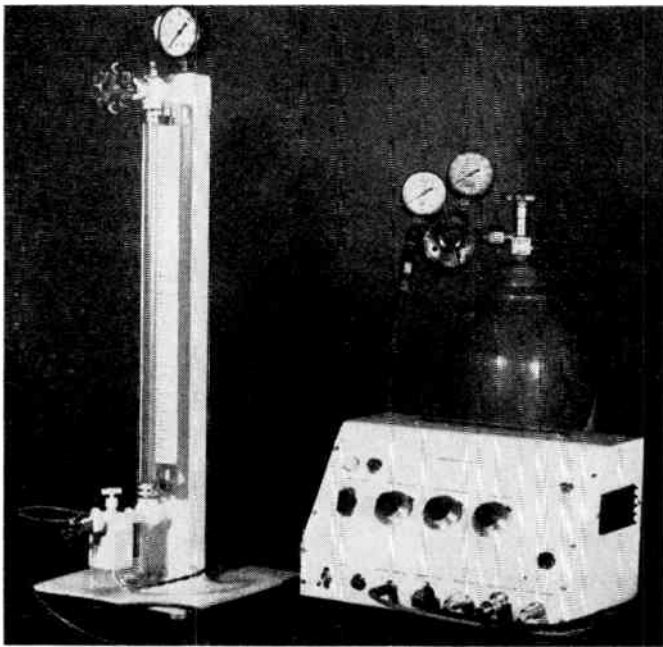


Fig. 2—Control unit with dye reservoir, solenoid valve, and cardiac catheter.

Since the electrocardiogram and aortic pressure were being recorded on the Sanborn Twin-Viso, the pre-amplifiers in these channels were used to provide the action potential signal of approximately 30 volts magnitude.

With reference to the circuit diagram, Fig. 3, this action potential (usually the rapid-rising peak of the aortic pressure wave) signal is applied to the grid of a 2D21 thyratron. This tube is normally held nonconducting by contacts on relay R1 which open the cathode circuit. When the operator depresses the remote switch (or the panel switch), R1 operates as a gate to allow the next action pulse to fire the thyratron. The other contacts on R1 pass ac power on for other circuit functions.

The first thyratron circuit acts as a pulse shaper and peak wave discriminator. Only signals raised above the bias of the grid by the adjustable gain of the pre-amplifier will fire the tube causing it to conduct and energize relay R2. The function of the R2 contacts is to momentarily break the plate supply voltage to the three time-delay circuits.

This type of time delay² does not require a well-regulated power supply. The thyratron is conducting in its stable condition but with a -105 volt grid bias. When the plate circuit is momentarily interrupted, conduction stops and the plate relay contacts switch $+105$ volts through an RC circuit to the grid. As the grid swings through its firing potential after the selected delay conduction in the stable condition is restored. The interval that the thyratron is nonconducting is the delay selected to open the solenoid valve after the action potential signal.

² S. Wald, "Precision interval timer," *Electronics*, vol. 21, pp. 88-89; December, 1948.

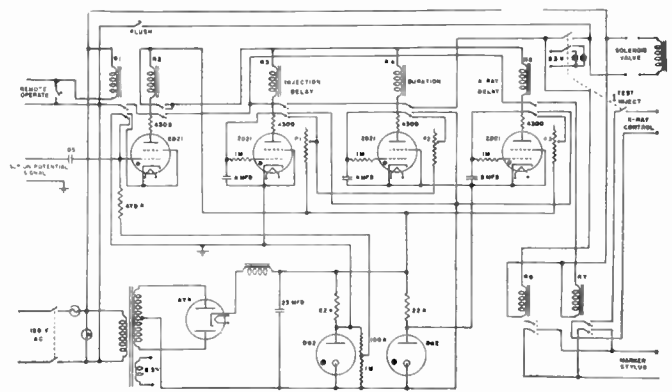


Fig. 3—Circuit diagram for electronic control for coronary arteriography.

In like manner, the action signal initiates the delay of the second delay thyratron, but its plate relay contacts cut off the solenoid valve power, closing the valve, following the delay interval selected. Thus, the "duration of injection" interval is the difference between the delays selected on the two circuits by controls P1 and P2. Because of the series connection of the two resistors this interval is necessarily greater than zero.

The solenoid valve receives its power through the series-connected contacts of R1, R2, R3, and R4. The contacts of R6 close with the valve operation and serve to activate the marker stylus on the Twin-Viso recorder to record the moment and duration of the injection with respect to the electrocardiogram.

The X-ray exposure is initiated by the third time-delay circuit. The period of delay starts with the firing of the pulse-shaping thyratron and is determined by the setting of P3. The firing of the time-delay thyratron actuates relay R5, which actuates R7, whose contacts close the X-ray machine firing circuit. The other contacts on R7 are used to operate the marker stylus on the electrocardiograph, recording the moment of the exposure. The X-ray time delay, being independent of the other two, permits the taking of arteriograms at a selected phase of the cardiac cycle while varying the moment and duration of injection. This flexibility is necessary for exploratory work in determining the optimum timing of the several operational functions.

A "test-inject" switch opens the solenoid valve circuit to permit the setting and recording of the time delays prior to actual injections. The momentary "flush" switch bypasses all other controls to permit holding the solenoid valve open for the purpose of flushing air out of the system. This precaution is extremely important to prevent any bubbles from entering the aorta through the catheter.

Fig. 4 shows a typical recording from the Twin-Viso electrocardiograph. The upper channel record is the electrocardiogram and the lower is the aortic pressure wave. At the bottom is the stylus marker trace indicating the sequential timing of the events. Sometime during the cycle preceding A, the surgeon depressed the



Fig. 4—Reproduction of Twin-Viso recording showing ECG tracing, aortic pressure wave, and stylus marking to indicate series of operations. A-B, injection delay; B-C, duration of injection; A-D, X-ray delay.

remote switch. At point A, the rising peak of the pressure wave fired the first thyatron initiating the sequence. At B, the solenoid valve opened, starting injection of the dye. At C, the valve closed, showing the duration of injection to have been 0.08 second. The X-ray exposure was made at D which was 0.43 second after the initial action potential signal.

Over 500 dye injections and X-ray exposures have been made on dogs with this equipment, exploring the possible variables. The results clearly indicated the ideal moments of injection and X-ray exposure for this type of experiment.

The medical aspects which motivated the development of this control for coronary arteriography and the experimental results are presented separately.³ After thorough testing on dogs, it is believed that the method is now ready for clinical application in diagnosis of human heart disease.

The method is also applicable for dye injection with X-ray demonstration of other vessels in the body, and for injection of gas instead of dye.

ACKNOWLEDGMENT

Early work was conducted in collaboration with Dr. E. J. Roth, who contributed to the theory and application of this project as well as to the interpretation and analysis of the results.

The authors wish to acknowledge the assistance of M. L. Hartley, T. J. Connor, O. M. Wiseman, and J. C. Hertsch of the Instrumentation Laboratory, Naval Medical Research Institute.

³ Urschel and Roth, *op. cit.*

The Impact of Electronics on Sanitary Engineering*

C. D. GEILKER† AND H. P. KRAMER‡

Summary—Although a field not usually thought of as depending upon electronics, sanitary engineering is feeling the impact of growing electronics instrumentation. This paper discusses specific developments in the areas of air pollution, radiological health, milk and food sanitation, and water supply and water pollution.

SANITARY engineering is a field which most people do not associate with electronics. Yet within the past decade electronics has been an increasing influence both in the research and in applied fields at the Robert A. Taft Sanitary Engineering Center. The activities of the Center, which is the research arm of the U. S. Public Health Service in the sanitary engineering field, are aimed not only at developing new and improved techniques of safeguarding the public health, but also at narrowing the gap between such develop-

ments and their widespread application at the state and local level. The Center's four research programs deal with water supply and water pollution, milk and food sanitation, air pollution, and radiological health. There is also an active training program which incorporates the results of the research groups into specialized short-term courses which are attended by professional personnel from state and local health departments, industry, universities, and the military services.

A decade and a half ago the concern of the sanitary engineer for the public health involved mainly chemical and bacteriological pollution of sources of water, milk, and food, and electronics made no great contribution. But with growing diversity of areas of concern extending, for example, into air pollution and radioactivity, and with the advent of automation and servomechanisms on a laboratory scale, electronics instrumentation began to appear in every program of Center activity. In some programs, such as water supply and water pollution, and milk and food, it made a gradual entry and gained increasing acceptance as it demonstrated superi-

* Original manuscript received by the IRE, December 4, 1958.

† Div. of Radiol. Health, Robert A. Taft Sanitary Eng. Center, Public Health Service, U. S. Dept. of Health, Education, and Welfare, Cincinnati, Ohio.

‡ Robert A. Taft Sanitary Eng. Center, Public Health Service, U. S. Dept. of Health, Education, and Welfare, Cincinnati, Ohio.

ority to conventional techniques in such operations as chemical analysis and temperature measurement and control. Other programs—radiological health, for example—were highly electronic from their beginning. At present the Center's investment in electronics represents almost \$400,000 with the distribution by programs as shown in Fig. 1.

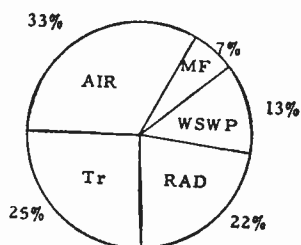


Fig. 1—Electronics investment, by programs—percentage of total investment.

AIR POLLUTION

The youngest of the Center programs has, as might be expected, the greatest emphasis upon electronics instrumentation. The air pollution program came into being with the passage of the National Air Pollution Act (Public Law 159, 84th Congress) in 1955. As a result, the Sanitary Engineering Center became the federal center for engineering and research in the field of air pollution. The air pollution program is the only Center program which has its own instrumentation section, whose objective it is to develop, modify, and evaluate new instrumentation. This section was created in 1957, with an electrical engineer as its chief. It now has three engineers, two technicians, and its own electronics laboratory.

The interest in electronics in the air pollution program is predominantly in highly automated sampling and analysis devices for field use. Automatic particulate tape samplers, odor evaluation apparatus, and multiple gas analyzers have been developed. Laboratory instrumentation includes infrared absorption analyzers, gas chromatographs, and mass spectrometers. These are applied to problems as diverse as smog analysis, auto exhaust studies, and incinerator evaluation.

RADIOLOGICAL HEALTH

No field of sanitary engineering is more completely dependent upon electronics instrumentation than is radiological health. Center work in this field began within four years of the first atomic bomb, and today constitutes one of the largest programs. New techniques are developed, and large numbers of samples from various phases of the environment—water samples, soil samples, air samples, milk samples—are assayed for their radioactivity content. Because the levels encountered are in general very low, the instrumentation is pushed to its limit in sensitivity. Determinations of gross alpha and beta radioactivity are made with gas flow internal proportional counters. These are about 25

per cent transistorized. For analytical determination of specific radioactive contaminants there is an increasing tendency to by-pass as much chemistry as possible and rely upon gamma-ray spectrometers. Characteristic gamma rays are emitted by most radionuclides, and identification may be made from the plots of gamma intensity vs gamma energy produced by the spectrometers. These instruments range in complexity from simple single-channel devices which scan the energy spectrum in ten minutes and display the results on a strip-chart recorder, through 20-channel spectrometers of the multistage voltage discriminator type, to 100-channel kicksorters which involve analog-to-digital conversion, magnetic core memory storage, high speed readout, and other computer-type equipment.

MILK AND FOOD SANITATION

The invasion of the milk and food chemistry fields by electronics has been more gradual. There have generally been conventional alternatives to electronic techniques here, so that change occurred only when electronics clearly offered either greatly increased efficiency and productivity or else made possible a degree of exactness previously impossible. An outstanding example of the first case is the analytical technique of spectrophotometry, in which a material is identified and/or quantitatively measured by means of its characteristic light absorptions, either in the visible light region or in the ultraviolet or infrared. Equipment using photoelectric cells existed in the 1940's, but with the widespread availability of the photomultiplier, self-balancing recording potentiometer, and servomechanisms for programming slit widths and scanning speeds, it has become possible for an investigator to accomplish in one afternoon what previously would have been two weeks' work. The field of temperature measurement and control, which is of importance in milk and food chemistry and bacteriology, has undergone a comparable refinement, in this case one of precision, with the application of electronics. From strictly manual control and bimetallic thermostats, temperature control has progressed to electronic relays and now thermistor controlled proportional regulation—systems capable of maintaining a specified temperature to within a few thousandths of a degree.

WATER SUPPLY AND WATER POLLUTION

The oldest field in sanitary engineering—water supply and water pollution—has also felt the impact of electronic instrumentation. Advances in basic analytical techniques such as spectrophotometry, particularly infrared, have permitted extensive determinations of trace-level organic contaminants—such as insecticides, petrochemical wastes, and surfactants—in water supplies. Another application of infrared spectrophotometry has resulted in a very rapid method for bacterial identification by individual comparison of a specimen's infrared absorption with that of thousands of known

spectra stored on IBM punch cards. Studies of water quality have required automatic recording devices which monitor the parameters of interest. An example here is the Thomas Dissolved Oxygen Monitor, which utilizes the paramagnetic properties of oxygen to determine the parts-per-million dissolved in a water source. Applications of analog and digital computers are being developed for use in problems of water supply and water pollution.

TRAINING

As the principal medium of transmission of new developments in all fields of Center research, the training program has elements of instrumentation representative of each. Course enrollments are limited to a maximum of 25 persons, and sufficient instrumentation is provided so that each student may have individual experience with the equipment. The large number of instruments required to make this possible has led to first-echelon maintenance within the training program, and to the appearance of an electronics scientist and a technician on the training staff. During periods between courses the facilities of the training program are utilized for research or by special project groups from other Center programs.

MODEL SHOP

The exponential growth of electronics instrumentation in the Sanitary Engineering Center, has led to the creation of a model shop staffed by an engineer, machinist, mechanic, and two electronics technicians. All of the programs consult with the model shop on problems

related to instrumentation; the model shop builds equipment to their design, in addition to handling maintenance of electronic equipment in the Center. It has become necessary to stock many hundreds of vacuum tubes and transistors and to use precision wide-range voltage measuring devices, pulse generators, fast oscilloscopes, and other elaborate and expensive service equipment. Activities of service and design are coordinated through the model shop, and a typical distribution of shop time by programs (April-September, 1958) is shown in Fig. 2.

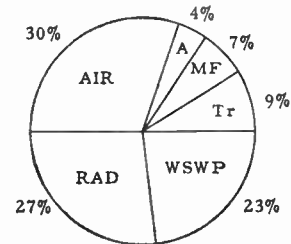


Fig. 2—Distribution of model shop time, by programs.

THE YEARS AHEAD

The growing complexity of man's environment presents the sanitary engineering field with an increasing challenge to devise and adapt new methods for safeguarding the public health. As these techniques require more and more elaborate instrumentation, it is certain that electronics will assume a role of vital importance. The impact has been made; an indispensable partnership lies ahead.

Instrumenting the Study of the Air Pollution Problem*

JOHN S. NADER†

Summary—A brief discussion is given outlining the problems of air pollution and the necessary instrumentation. The conventional types of instruments in general use are summarized, and a review of special developments is given. Attention is called to the design and development requirements of instruments needed for current and future studies of the air pollution problem.

THE increasing attention to air pollution in the past several years has culminated in the recent National Conference on Air Pollution called by the Surgeon General of the U. S. Public Health Serv-

ice.¹ The Conference asked questions: What is in the air? How does it get there? Where does it come from? What does it do to humans? What can be done? How is it to be done?

To obtain the answers to these questions, a complex array of instrumentation is necessary.

This instrumentation, ranging from the simplest manually operated grab sampler to the complicated mass spectrometer, is interdisciplinary in drawing information from a number of different fields, meteorology,

* Original manuscript received by the IRE, December 4, 1958.

† Chief of Instrumentation, Air Pollution Eng. Res., Robert A. Taft Sanitary Eng. Center, Cincinnati, Ohio.

¹ National Conference on Air Pollution, Sheraton-Park Hotel, Washington, D. C., November 18-20, 1958. Proceedings to be published.

chemistry, physics, engineering, and medicine.² In fact, it is as complex as the air pollution problem itself. This diversity of instrumentation imposes a requirement on the instrument engineer for a broad background and versatility in the application of instrumentation to air pollution measurement and control. This requirement is simplified only as a class of instruments, *i.e.*, recorders, transducers, or others, and is applicable to several disciplines. The diversity of instrumentation is best appreciated with application to specific aspects of the air pollution problem.

To determine what is in the air, it is necessary to sample the particulate and gaseous contents of the air and to analyze them for their physical and chemical properties.³⁻⁵ Supporting meteorological data, temperature, humidity, wind speed and velocity are also required to establish the environmental conditions in which pollutants are found. Here the required instrumentation includes samplers, analyzers, and meteorological equipment.⁵⁻⁸ To determine how the pollution gets into the air and to establish its source require more detailed information about the meteorological parameters of temperature, wind speed and direction, the stratification of air layers, the dynamics of air movements, and gaseous diffusion. The control of pollutant emissions to the atmosphere and the enforcement of legislation for such control necessitate instrumentation for monitoring to establish the efficiency of control and its conformance to established standards.

The effects of pollution on humans are considered from physiological and psychological aspects. Instrumentation plays an important role in studies of eye irritation, odor thresholds, lung cancer, chronic bronchitis, and other afflictions associated with air pollution.^{1,4,9,10} Since instruments provide an objective measurement, their application is most helpful in areas where human response can be correlated to a measurable physical phenomenon.

² P. L. Magill, F. R. Holden, and C. Ackley, "Air Pollution Handbook," McGraw-Hill Book Co., Inc., New York, N. Y.; 1956.

³ "Air Pollution Measurements of the National Air Sampling Network," U. S. Government Printing Office, Washington, D. C., Public Health Service Publication No. 637; 1958.

⁴ "The Air Over Louisville," Summary of a Joint Report by the Special Air Pollution Study of Louisville and Jefferson County, Ky., 1956-1957. Robert A. Taft Sanitary Eng. Center, Cincinnati, Ohio, 57 pp.; May, 1958.

⁵ G. D. Clayton, "Determination of Atmospheric Contaminants," American Gas Assoc., New York, N. Y., Rep. No. PC-53-12; May, 1953.

⁶ "Encyclopedia of Instrumentation for Industrial Hygiene, Instruments Specifically Designed for Atmospheric Pollution Evaluation and Meteorological Measurements," University of Michigan, Inst. Industrial Health, Ann Arbor, Mich. Sec. 3; 1956.

⁷ "Instruments for the Study of Air Pollution," ASME, New York, N. Y., 3rd ed.; 1958.

⁸ W. E. K. Middleton and A. F. Spilhaus, "Meteorological Instruments," University of Toronto, Toronto, Can., 3rd ed.; 1953.

⁹ A. Goetz, "Study of Properties of Aerosols with Particular Reference to the Nature of the Air-Particle Interface," Robert A. Taft Sanitary Eng. Center, Cincinnati, Ohio, Tech. Rep. A58-10, U.S.P.H.S. Contract No. SAPH-69537; August, 1958.

¹⁰ J. S. Nader, "An odor evaluation apparatus for field and laboratory use," *Amer. Ind. Hygiene Assoc. J.*, vol. 19, pp. 1-7; February, 1958.

GENERAL INSTRUMENTATION

Many of the instruments utilized in determining the air pollution problem are of conventional types usually applied in industrial process and control work with which the instrument engineer is familiar. These include standard components such as various relays; pressure, temperature, and velocity transducers; potentiometric and galvanometric strip chart recorders; small-signal dc amplifiers; and servomechanisms. These components are common to equipment routinely used in the various disciplines.

In chemical analysis the instrumentation includes many of the well-recognized laboratory-type apparatus, namely pH meters, spectrophotometers, electronic balances, infrared and ultraviolet analyzers, vapor phase chromatographs, mass spectrometers, etc. The literature covers this subject extensively with reference to current standard techniques and to new developments.^{11,12}

Much of the equipment in use by the meteorologist is directly applicable to the air pollution problem since a study of the meteorological parameters is an immediate objective. Measurement of parameters in the upper atmosphere introduces the use of telemetering instrumentation. This technique is well established in the activities of the U. S. Weather Bureau and the Air Force.

SPECIFIC INSTRUMENTATION

Special developments of instrumentation for the study of air pollution have thus far involved, for the most part, costly and delicate equipment lacking portability and flexibility. This is particularly true of instruments used in automatic monitoring of gaseous pollutants in the atmosphere, in which the analysis is achieved by colorimetric techniques involving gas absorption by selective chemical reagents. There are many aspects of the air pollution problem which can benefit greatly by instrument developments based on new and improved applications of physical methods. Physical methods are being applied, using infrared and ultraviolet absorption techniques.¹³ Radioactive isotopes have been used in a beta gauge development¹⁴ applied to a research study of air pollution control equipment. A similar beta gauge technique is the basis of an instrument under development by the author for measurement of dust loading in the ambient atmosphere. Ion current measurement has been incorporated in an in-

¹¹ S. Lewin, "Trends in laboratory instrumentation for analysis," *Analytical Chemistry*, vol. 30, pt. 1, pp. 19A-29A, pt. 11, pp. 17A-26A; June-July, 1958.

¹² H. H. Willard, L. T. Merritt, and J. A. Dean, "Instrumental Methods of Analysis," D. Van Nostrand Co., Inc., Princeton, N. J., 3rd ed.; 1958.

¹³ L. H. Rogers, "High sensitivity continuous instrumentation for atmospheric analysis," *Chem. Eng. Progress*, vol. 53, pp. 381-384; August, 1957.

¹⁴ D. G. Stephen, *et al.*, "A new technique for fabric filter evaluation," *Amer. Indus. Hygiene Assoc. J.*, vol. 19, pp. 276-284; August, 1957.

strument¹⁵ to measure specific gaseous pollutants which react with prepared reagents to form aerosols in an ion chamber. Atmospheric particle counting and particle size analysis are being accomplished by instrumentation which utilizes light scatter to generate electrical pulses that are counted and sized by pulse techniques.¹⁶

Present requirements in instrumentation for air pollution measurements, furthermore, emphasize the need for equipment which is less expensive, more portable, reliable under continuous operation, more or less automatic, and sensitive to the low levels of pollutant concentration found in community atmospheres. Where satisfactory physical methods have been applied to air pollution instrumentation, the problem of instrument improvement is essentially one of miniaturization and system engineering to obtain portability, reliability and reduced cost. In some laboratory-type instrumentation, there is a need of greatly improved sensitivity while minimizing noise levels sufficiently to maintain a satisfactory signal-to-noise ratio. This improved sensitivity may be brought about by circuit revision and/or modification by quality control of components, or by adaptation of alternative physical methods which may be less sensitive to noise and other interferences.

The need for automation extends beyond that re-

¹⁵ R. A. Morris, *et al.*, "Detection and Analysis of Gases by Ion Current Measurements," Mine Safety Appliance Co., Pittsburgh, Pa.; 1956.

¹⁶ A. L. Thomas, "Measuring and counting small particles," *Bull. Southern Res. Inst., Birmingham, Ala.*, vol. 10, no. 3, pp. 14-18; 1957.

quired for unattended performance over long periods. The accumulation of field data becomes extensive, especially in monitoring work, and the handling of these data is time consuming and subject to human errors. Therefore, there is a need to engineer the instrumentation to present data in printed digital form instead of on strip chart line recordings and in a form which lends itself to data-handling processes such as punched card and magnetic tape systems. Equipment for conversion of analog data to digital form has been developed and is used in conjunction with conventional instrumentation. In several instances a conversion system has been directly incorporated in instrumentation for air pollution measurements.

It is only within recent years that the air pollution problem has received special attention as to development of specific instrumentation engineered to the particular aspects of the problem under study. Previously, most of the instruments in use were carried over from the field of industrial hygiene with more or less modification as the occasion demanded. This situation continues to a certain extent and represents an area in which the attention and competence of more instrument engineers would be welcomed. The air pollution problem provides a widening opportunity to apply current knowledge and techniques of electronics to the instrumentation requirements presented by various aspects of the problem. Capable engineers will find their abilities well challenged and the results well worth their effort and interest.

Correspondence

The Usefulness of Transconductance as a Transistor Parameter*

In a recent communication,¹ Armstrong points out that the transconductance of a transistor in the common emitter configuration is a useful parameter and is relatively independent of frequency. From

$$\beta = \frac{\beta_0}{1 + j \frac{\omega}{\omega_\beta}} \quad (1)$$

$$z_i = \frac{r_i}{1 + j \frac{\omega}{\omega_\beta}} \quad (2)$$

where

$$r_i = r_b + (\beta_0 + 1)r_e \quad (3)$$

he shows that

$$g_m = \frac{\beta}{z_i} = \frac{\beta_0}{r_i} \quad (4)$$

and, therefore, concludes that in this approximation the transconductance is independent of frequency.

Transconductance is indeed a useful transistor parameter. However, it is worthwhile pointing out that the result expressed in (4) is somewhat oversimplified.

It can be shown² that the general expression for the transconductance is given by

$$G_m(s) = \frac{I_c(s)}{V_s(s)} = \frac{\alpha_0}{R_e + (1 - \alpha_0)R_b} \cdot \frac{1}{1 + \frac{s}{\omega_r}} \quad (5)$$

where it is related to the source voltage and where

$$\omega_r = \frac{R_e + (1 - \alpha_0)R_b}{R_e + R_b} \cdot \omega_\alpha \quad (6)$$

In these equations, $V_s(s)$ is the Laplace transform of the source voltage, R_e is the total series emitter resistance, R_b is the total series base resistance, ω_α is the cutoff frequency of the internal current generator, and the derivation of (5) is subject to the usual assumption that the collector time constant is negligible.

Eq. (5) is valid for any circuit configuration of the transistor, including circuits where none of the transistor terminals are grounded, and circuits with arbitrary source resistance (*i.e.*, not necessarily a voltage drive). Whereas (5) holds for any configuration, the input impedance is dependent on the input terminal. The input impedance for the base-input circuit is given, when related to the source voltage, by

* Received by the IRE, February 19, 1959.
¹ H. L. Armstrong, "On the usefulness of transconductance as a transistor parameter," *Proc. IRE*, vol. 47, pp. 83-84; January, 1959.

² A. Harel and J. F. Cashen, "Unified representation of junction transistor transient response," *RCA Rev.*, vol. 20, pp. 136-152; March, 1959.

$$Z_i(s) = \frac{V_i(s)}{I_b(s)} = (R_e + R_b) \frac{s + \omega_r}{s + \omega_\beta} \quad (7)$$

The analysis carried out by Armstrong may be regarded as a special case of these equations, in which only the transistor is considered, *i.e.*,

$$R_e = r_e \quad \text{and} \quad R_b = r_b \quad (8)$$

In this case, (5)–(7) become, respectively, for the transistor itself,

$$g_m(s) = \frac{\alpha_0}{r_e + (1 - \alpha_0)r_b} \frac{1}{1 + \frac{s}{\omega_r}} \quad (9)$$

$$\omega_r' = \frac{r_e + (1 - \alpha_0)r_b}{r_e + r_b} \omega_\alpha \quad (10)$$

$$z_i(s) = (r_e + r_b) \frac{s + \omega_r'}{s + \omega_\beta} \\ = \frac{[r_b + (\beta_0 + 1)r_e] + \left[(r_e + r_b) \frac{s}{\omega_\beta} \right]}{1 + \frac{s}{\omega_\beta}} \quad (11)$$

Clearly, as may be verified by (9) and (11), the transconductance is given by

$$g_m = \frac{\beta}{z_i} \quad (12)$$

It is, however, frequency dependent. The frequency independence expressed in (4) is valid only in the range

$$\omega \ll \omega_r',$$

resulting from

$$[r_b + (\beta_0 + 1)r_e] \gg \left[(r_e + r_b) \frac{s}{\omega_\beta} \right], \quad (13)$$

which is the condition that reduces (11) to (2).

As a matter of fact, the exact frequency dependence of the transconductance is given by (9) where it is seen that ω_r' is the cutoff frequency of the transconductance. The transconductance, g_m , is indeed a weaker function of frequency than the common emitter current amplification factor, β , since

$$\omega_\beta \ll \omega_r' < \omega_\alpha. \quad (14)$$

It is of interest to note that the frequency dependence of g_m indicated in (9) correlates well with the experimental results³ mentioned by Armstrong, which, however, are not quantitatively evaluated by him.

We feel that transconductance—a well-defined function of frequency, both for transistors and for tubes, with a relatively high cutoff frequency—is a useful parameter, which may be of help, particularly to students. It aids in combining transistor and tube theories, as well as in representing the response of various circuit configurations of the transistor in a unified way.²

J. F. CASHEN

A. HAREL

Radio Corp. of America
Camden 2, N. J.

An Estimate of Transistor Life in Satellite Applications*

The investigations connected with the IGY have shown¹ that there are two distinct and widely separated zones of high-intensity corpuscular radiation. The nature of this radiation is not known unambiguously, but since these high-intensity belts are believed caused by the geomagnetic field, the radiation must be primarily charged particles, *i.e.*, protons and electrons. Since all measurements thus far have been with Geiger counters, only the flux (number of protons or electrons per cm² per second) is known. An estimate of the energies involved can be obtained from the shielding present around the Geiger tubes.

In summary, the following information is available on the radiation in these high-intensity belts:

- 1) Nature—most likely protons and electrons
- 2) Energies
 - a) if protons, the energy is approximately 10 to 100 mev
 - b) if electrons, the energy is approximately 1 to 10 mev
- 3) Peak flux observed in the high-intensity belts was approximately 10⁴ particles per cm² per second.

The mechanism of damage to transistors (as well as other solid-state devices) is the displacement of atoms of the substance from their equilibrium lattice sites by collisions with the incident particles. Thus, to a first approximation, after a critical number of such displacements, the transistor will be damaged significantly. In this sense, all types of radiation are equivalent in that they tend to produce damaging displacements. There are other effects which are produced by some kinds of radiation and not others, but these can be neglected here.

For example, just before a proton stops, it will be very damaging; but since we are dealing with high-energy particles and thin targets (the base of a transistor), the proton will tend to pass right through without stopping.

Assuming that the radiation consists of protons with energies of 100 mev, the number of displacements per second per unit volume can be calculated to be $N_d = 3.5 \times 10^4$ per second per unit volume.

Now it has been experimentally demonstrated² that the γ radiation from a 100 Curie Co⁶⁰ source produces significant damage to transistors after a dose of 10⁶ roentgens. The flux from this source was 2×10^6 roentgens per hour so the time needed in this environment for transistors to be damaged significantly was approximately 5 hours. Calculation shows that the radiation from this source produces 10¹⁰ displacements per second per unit volume. Thus the "critical" number of displacements needed to produce significant transistor damage is

$$N_{\text{total}} = 1.8 \times 10^{14} \text{ displacements per unit volume.}$$

Therefore, since for the proton damage under consideration $N_d = 3.5 \times 10^4$ per second per unit volume, the time needed for damage in the proton environment is

$$t = \frac{N_{\text{total}}}{N_d} = 5 \times 10^8 \text{ sec} \approx 10^5 \text{ hours.}$$

Admittedly, the above calculation is rough, for the type of damage by gammas and protons is significantly different. Even if the above figure is off by 100, the useful life would still be 10³ hours in the region of peak radiation intensity. A typical satellite orbit might only touch this region for a short time during each revolution resulting in much longer transistor life. Also, since the radiation probably consists of both electrons and protons, a longer life may be anticipated.

It is therefore concluded from the information available on the properties of the radiation in the high-intensity belts, that the useful life of transistors in satellite applications should be longer than 10⁴ hours provided that reasonably protective circuits are used.

A. J. HEEGER

T. R. NISBET

W. W. HAPP

Lockheed Missile Systems Div.
Palo Alto, Calif.

Recordings of Transmissions from the Satellite 1958 Δ_2 at the Antenna Laboratory, The Ohio State University*

Observations of the radio transmissions from the Russian Satellite 1958 Δ_2 (Sputnik III) are currently being carried out at the Antenna Laboratory Field Station, The Ohio State University, Columbus, Ohio.

It has been reported¹ that the signal characteristics of transmissions from Sputnik I on 20.005 mc were of an unsymmetrical nature during passages of the satellite in a NNW to SSE direction near Columbus, Ohio. This effect was not found on the SSW to NNE passages, and it was postulated that the change in signal character on the NNW to SSE pass was due to an ion cone being formed ahead of the satellite. Signal strength recordings of the 20.005-mc transmission from Δ_2 , taken at the Antenna Laboratory throughout the latter half of 1958, show that a similar phenomenon exists on both directions of pass near Columbus, irregular signals being received when the satellite is north of Columbus irrespective of the direc-

* Received by the IRE, February 24, 1958.

¹ J. A. Van Allen and L. A. Frank, "Survey of radiation around the earth to a radial distance of 107,400 kilometers," in "Annals of the IGY," Pergamon Press, London, Eng.; January, 1959.

² W. W. Happ and S. R. Hawkins, "A Critical Survey of Radiation Damage to Circuits," Third Semi-Annual Radiation Effects Symp.; October, 1958.

* Received by the IRE, February 16, 1959. This work was partially supported by the USAF through the Wave and Propagation Group, Wright Air Dev. Center, Wright-Patterson Air Force Base, Ohio.

¹ J. D. Kraus and J. S. Albus, "A note on some signal characteristics of Sputnik I," Proc. IRE, vol. 46, pp. 610–611; March, 1958.

³ P. J. W. Jochems, O. W. Memelink, and L. J. Tummers, "Construction and electrical properties of a germanium alloy-diffused transistor," Proc. IRE, vol. 46, pp. 1161–1165; June, 1958. Figs. 10–12 especially illustrate the point made here.

tion of the pass. Thus, it is concluded that the irregular signals are caused by some phenomenon occurring in the area to the north of Columbus and not necessarily due to the formation of an ion cone ahead of the satellite.

Some variation in the amplitude characteristics of the recorded signals during a pass would be expected due to the presence of the earth's magnetic field. Approximate calculations of the components of the earth's magnetic field intensity for a pass over Columbus in a NNW to SSE direction indicate that the relative magnetic intensity in the direction of propagation is small in the region to the north, reaches a maximum over Columbus, and slowly decreases to the south. The component transverse to the direction of propagation is a maximum in the northern region, is extremely small over Columbus, and increases to the south. Applying these conditions to the Appleton-Hartree expression for the refractive index, it is found that the polarization of the satellite's signals is determined mainly by the transverse component when the satellite is to the far north of Columbus and by the component in the direction of propagation when it is near or to the south. If the satellite's signals are received on a linearly polarized antenna, then the voltage amplitude at the antenna terminals, when the satellite is in the northern region, will be determined by two oppositely-rotating elliptically-polarized components. In the near and southern regions the polarization ellipses become circles, resulting in a regular amplitude fading superimposed on the transmission by rotation of the polarization (Faraday rotation). The point at which the effect of the transverse component can be neglected occurs at a satellite elevation angle of approximately 30 degrees from Columbus for the NNW to SSE overhead pass at low altitude. A more detailed analysis of the polarization characteristics of satellite transmissions is in progress.

During the observations of Δ_2 transmissions, it has been noticed that large changes occur in the integrated electron density, deduced from the Faraday rotation rate, from day to day. Table I gives the value of N_h for

TABLE I

Date	Time EST	Height Miles	$N_h \times 10^{17}$
12-1-58	19.29	898	0.81
	21.16	935	0.82
12-2-58	19.25	900	>8.0
	21.12	936	>8.2
12-3-58	19.24	906	1.61
	21.11	940	1.65
12-4-58	19.20	910	>7.9
	21.06	942	8.24

sample days in December, 1958. These particular days were chosen because the satellite is approximately in the same position on successive days, the 19.29 pass being to the east and the 21.16 pass to the west of Columbus. The value of N_h was computed from the expression,

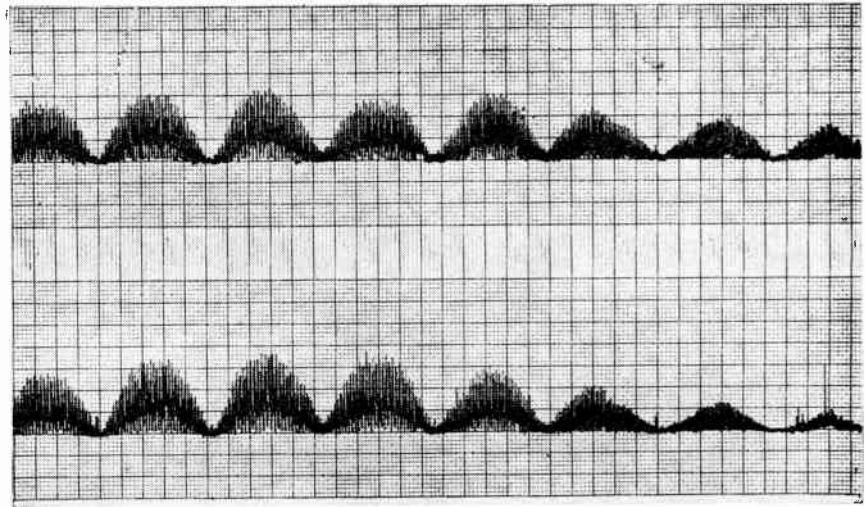


Fig. 1—Amplitude fading superimposed on 1958 Δ_2 transmission. One large division equals 10 seconds.

$$N_h \approx \frac{\pi h f^2}{2.97 \times 10^{-2} H v T},$$

where

- N_h = the electron content in a column of height h and 1 square meter in area,
- h = satellite height in meters,
- f = frequency in cps,
- H = horizontal component of the earth's magnetic field in amp/meter,
- v = satellite velocity in meters/second,
- T = time between adjacent nulls of the polarization rotation in seconds.

The above expression is only applicable for NNW to SSE passes near Columbus, since it neglects the effect of the E-W component of the earth's magnetic field and is approximate due to neglect of the time dependence of h , v , and H over the time duration of 180 degrees rotation of polarization. It will be noted that the values for N_h on December 2 and 4 are in excess of those normally expected and reported by others.² Absolute values for three of the passes are not given due to the pulsed nature of the signals, which makes it difficult to measure extremely small Faraday rotation periods accurately.

In November, 1958, it was observed that a severe amplitude fading, not due to Faraday rotation, had become superimposed on the transmission from Δ_2 . An example is shown in Fig. 1 in which two orthogonal polarizations are recorded, indicating an amplitude fade every 52 seconds in the central portion of the record. During November, 1958 to January, 1959, the period varied between 39 and 52 seconds. The rotation and precession of the satellite could be such that a considerably different aspect of the antenna pattern was observed. However, since the appearance of this effect was rather sudden, perhaps some damage or malfunction has occurred to one of the antenna elements causing the antenna pattern to have more pronounced nulls.

² R. Parthasarathy and G. C. Reid, "Signal strength recordings of the Satellite Δ_2 (Sputnik III) at College, Alaska," Proc. IRE, vol. 47, pp. 78-79; January 1959.

The recording apparatus used for the observations of Russian satellite transmissions on the 20.005-mc channel consists of two orthogonally oriented dipoles situated 0.2 λ above ground, each dipole being connected to an RME DB 23 preselector and a Collins 51J4 receiver. The AVC outputs of the receiver are recorded on a Sanborn high-speed dual-channel recorder. The second harmonic of the 20.005-mc transmission is monitored using a 40.01-mc dipole 0.2 λ above ground, a Tapetone TC40 converter, and a Collins 75A4 receiver. The AVC voltage of the 40.01 receiver is recorded on the FM channel of an Ampex twin-channel tape recorder together with 20.005 or 40.01-mc Doppler data on the AM channel.

T. G. HAME
E. M. KENNAUGH
Antenna Lab.

The Ohio State University
Columbus, Ohio

Tunable L-Band Ruby Maser*

A tunable solid-state maser amplifier using ruby¹ has been operated over a frequency range from 850 to 2000 mc. Operation to date has been achieved using high magnetic fields oriented at 90° with respect to the ruby C axis. Measurements of the voltage-gain bandwidth product were made at liquid helium bath temperatures of 4.2°K and 1.5°K. The results of these measurements are listed in Table I.

Table I shows that a voltage-gain bandwidth product of 37.5 mc was measured at 1750 mc, for a temperature of 1.5°K. Thus, for a half-power bandwidth of 6.5 mc, a gain of 15 db will be obtained. Furthermore, operation was also obtained at an operating

* Received by the IRE, February 12, 1959. This work was supported by the Dept. of Defense.

¹ G. Makhov, C. Kikuchi, J. Lambe, and R. Terhune, "Maser action in ruby," Phys. Rev., vol. 109, p. 1399; February 15, 1958.

TABLE I
SUMMARY OF L-BAND MASER PERFORMANCE

Operating Temperature (°K)	Signal Frequency (mc)	Voltage-Gain Bandwidth Product Measured (mc)	Voltage-Gain Bandwidth Product Determined from External Q Measurement (mc)	Pump Frequency (mc)	Magnetic Field (Gauss)
4.2	2000	*	*	12,150	2260
4.2	1815	20	22	11,860	2150
4.2	1200	11	13	10,950	1740
4.2	1010	9	12	10,700	1650
4.2	850	*	*	10,460	1510
1.5	1750	37.5	37.5	11,795	2100
1.5	1200	19	20	10,845	1740

* Not measured because of equipment limitation in initial setup.

temperature of 4.2°K, where a voltage-gain bandwidth product as high as 20 mc was measured. The pump power that was required varied between 5 and 150 mw for the various measurements.

The voltage-gain bandwidth product was determined by two methods. In the first method, gain and half-power bandwidth were measured directly using the setup shown in Fig. 1. In this arrangement, maser

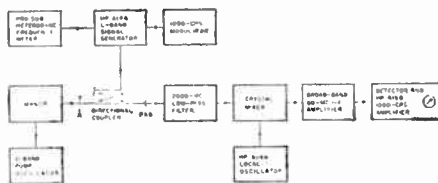


Fig. 1—Setup for measurement of maser voltage-gain bandwidth product.

gain was measured by noting the increase in generator output required to maintain a constant output meter reading when the maser is disconnected from the circuit and replaced by a short at point A of Fig. 1. In the second method, the external Q of the maser cavity is measured with the pump power and magnetic field turned off.

The voltage-gain bandwidth product can then be calculated since, for the condition of high gain, it is approximately equal to twice the signal frequency divided by the external Q.²

The maser cavity operated in the TEM mode at the signal frequency and the TE₁₀₇ mode at the pump frequency. The filling factor (defined here as the fraction of the total magnetic energy at the signal frequency that is stored within the volume occupied by the ruby crystal) is estimated to be 85 per cent. The 100-carat ruby crystal had a residual chromium content of 0.05 per cent. The signal and pump frequency cavity modes are independently tunable over thousands of megacycles, so that maser operation is quickly obtainable at any frequency within the operating range. The 90° crystal orientation used was attractive, because the calculated maximum signal transition probability is high for this orientation.³

The 850- to 2000-mc tuning range does not represent the maximum tuning range of

this maser. It is believed that a tuning range of at least two octaves can be achieved.

The advice and aid of Prof. C. H. Townes and A. Penzias of Columbia University are gratefully acknowledged.

F. R. ARAMS
S. OKWIT
Airborne Instruments Lab.
Div. of Cutler-Hammer, Inc.
Mineola, N. Y.

Half-Cycle Resonant Delay Circuit*

The circuit shown in Fig. 1 can be made to exhibit several of the waveforms usually associated with a one-shot multivibrator, or univibrator. In particular it produces a rectangular output pulse which goes positive with the input pulse and which returns to its baseline after a fixed delay time. The circuit differs from most univibrator circuits, however, in that the delay time τ is equal to the natural half period of a resonant circuit, consisting of L and C in the diagram. The half cycle of oscillation is initiated by a positive input pulse, acting through a diode D_1 and is concluded abruptly when current through the inductance coil L tries to reverse and is interrupted by the diode D_2 . With the interruption of this current, the output voltage returns rapidly to its initial value, according to a transient determined by the parallel resonant circuit consisting of L , R and $(C_1 + C_2)$. Here C_1 is the distributed capacitance of the coil, and C_2 is the input stray capacitance of the cathode follower. In practice R is chosen to give critical damping, and C_1 and C_2 are made as small as possible. The fall time of the output signal is approximately given by

$$\tau \approx \sqrt{\frac{C_1 + C_2}{C}}$$

and can be made less than one fiftieth of the delay time τ .

It is believed that this circuit has several advantages. For one thing, the delay time is controlled by passive circuit elements as in a delay-line-controlled multivibrator and is insensitive to variations in the characteristics of such active elements as are in fact

necessary to transmit the output signal, or to variations in the supply voltages thereto. Secondly, it will be noticed that, with the cessation of the current interruption transient, the voltage across C is essentially zero and there is no current flowing through the inductance coil. In other words, the circuit is already completely recovered, and ready to accept another pulse at any time. It may therefore be driven at rates up to very nearly the theoretical maximum for such a delay circuit, *viz.*, $1/\tau$, and the delay time will be essentially independent of frequency throughout this range. To summarize the advantages of the circuit: the delay time is essentially independent of tube characteristics, power supply voltages, and the frequency of pulsing or the time since last pulsed, whichever one wishes to consider. A possible disadvantage of the circuit is that the output pulse size is proportional to the input pulse height and if several such circuits are to be cascaded, some form of clipping or gain stabilization must be employed. A further disadvantage is that the circuit operation will be disturbed by an input pulse that arrives at any time before the delay cycle is completed.

A practical version of this circuit has been constructed to be used as a part of a precision time-gated sampling circuit at Lincoln Laboratory. Ten such circuits are connected in cascade and driven from a precision frequency standard at a rate of either 15 or 30 cps to provide a selection of pulses at integral millisecond delays with respect to the first. The unit millisecond delay circuit is shown in Fig. 2. Here a pentode is

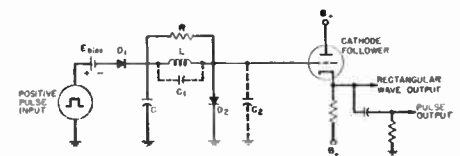


Fig. 1—Half-cycle resonant delay circuit.

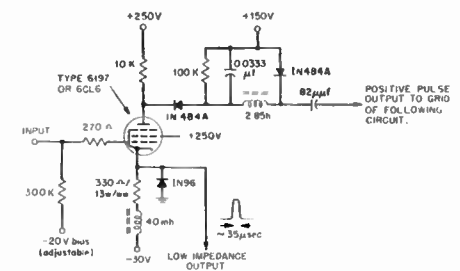


Fig. 2—Unit 1-millisecond delay circuit.

used to restore the pulse and correct the polarity in each stage; a current-limiting cathode circuit stabilizes the amplitudes of the pulses throughout the chain. The pulse output from each stage is available at low-impedance level at the cathode; the width of this pulse is approximately 35 μ sec. In this particular version it was not necessary that the circuit operate anywhere near the theoretical maximum rate of 1000 cps, and this feature was deliberately sacrificed in order to provide greater stage gain. The inductance

² J. O. Artman, "The solid-state maser," *Proc. Symp. on Role of Solid-State Phenomena in Electric Circuits*, Polytech. Inst. Brooklyn, sec. 3, p. 77; April, 1957.
³ W. S. Chang and A. E. Siegman, "Characteristics of Ruby for Maser Applications," *Electron Devices Lab., Stanford Univ., Stanford, Calif., Tech. Rep. 156-2*; September 30, 1958. See Figs. 14 and 15.

* Received by the IRE, December 9, 1958. The work reported here was performed at Lincoln Lab., M.I.T., Lexington, Mass., with the joint support of the Army, Navy, and Air Force, under contract.

coils were specially wound for low distributed capacitance (in this case approximately 15 μmf) and temperature-stabilized molybdenum permalloy powder cores of $\mu=125$ were used. Trimming of the individual delay circuits was effected by adjustment of the 0.0333- μf condensers. No significant variation of delay time with signal level was observed with this circuit, indicating that the powder cores were operating well within their linear range. The stability of this circuit, when operated entirely from unregulated power supplies and starting from a cold start, is a few parts in a thousand.

It will be noticed that by differentiation one may obtain a pulse output from such a circuit the length of which can be made shorter than the input pulse which initiates the cycle of delay. Thus, such a circuit can act as a pulse sharpener (with a built-in delay) and by the suitable cascading of several such circuits it should be possible, starting with any sort of pulse, to generate pulses with lengths measured in millimicroseconds.

BRADFORD HOWLAND

Lincoln Lab.
Mass. Inst. Tech.
Lexington, Mass.

Alternative Detection of Cochannel FM Signals*

In his recent correspondence,¹ Farris gave a sketchy description of a system² that originated with the present writer to facilitate the extraction of the message of the weaker of two cochannel FM signals. Unfortunately, this description has omitted a great deal of vital information about the "proposed method," including some necessary references.^{3,4}

The system that Farris presents in his Fig. 1 is an example of a technique that I conceived in 1955 for capturing the weaker, or enhancing the relative amplitude of the stronger, of two cochannel FM signals. The name chosen for this system is "feedforward." The problem that led me to this technique was that of devising a means for improving the interference-suppression capability of a narrow-band limiter "... by some circuit modification that does not alter the character of the narrow-band limiter *per se*, but instead uses such a limiter as a central building block in a more elaborate

scheme. One possibility (was) the use of feedback at a proper angle and through a proper feedback bandwidth, from the output of a limiter to its input."⁵ Another possibility was the use of "feedforward" at a proper angle and through a proper bandwidth, from the input of the limiter to its output.

From the theory of feedforward across the limiter it has been shown that this technique could be used for reliable capture of the weaker of two cochannel FM signals, or for improving the stronger-signal capture capabilities of an FM receiver. Indeed, theory and experiment at the M.I.T. Research Laboratory of Electronics have shown that the feedforward technique can be incorporated (through economical modifications) in a conventional FM receiver to enable the receiver to deliver high-quality reproductions of programs that are carried by two independent FM carriers whose frequencies sweep contiguous parts of an assigned frequency band. The quality of weaker-signal reception deteriorates if the two carriers are allowed to sweep overlapping portions of the band, but useful reception of the weaker signal is possible even when the carrier frequencies sweep the same frequency band.

The feedforward system has been described and demonstrated to the public repeatedly within the last year. In particular, it was presented at the M.I.T. Electrical Engineering Colloquium of April 9, 1958, and at the M.I.T. Industrial Liaison Symposium on Communication: Signaling and Detection, on May 15, 1958. Nearly one hundred representatives from twenty industrial concerns were present at this Symposium. The feedforward system has also been described in more recent publications.⁶⁻⁸

The feedforward arrangement shown in Farris's Fig. 1 is undesirable for practical realization. The ratio of the outputs of paths *A* and *B* will vary with the input signal level. Consequently, for certain signal levels, the system will suppress the stronger signal; for others, it will suppress the weaker signal. Moreover, the quality of signal suppression will also depend upon the input signal level. Thus, the system performance cannot be optimized with one setting for a significant range of signal levels. Indeed, this dependence of the performance upon the input signal level can cause the capture to switch from one signal to the other even for relatively slight changes in the input signal level. This difficulty can be remedied by contriving by any suitable means to regulate the level of the signal that drives paths *A* and *B* in Farris's diagram. The simplest and most reliable way to do this is to place an amplitude limiter of appropriate bandwidth before paths *A* and *B*. If path *A* contains only one

narrow-band limiter, then the limiter that drives the combination must also be narrow-band. If, however, two or more narrow-band limiters are cascaded in path *A*, then the limiter that drives the combination need not be narrow-band. Incidentally, narrow-band limiting is not absolutely necessary in path *A*. The narrow-band filtering can be carried out in the adder.

ELIE J. BAGHDADY
Res. Lab. of Electronics
Mass. Inst. Tech.
Cambridge, Mass.

Author's Comment⁹

The first reference which the writer has been able to find in print on the subject of weak-signal detection by a subtraction scheme resulted from work at the Georgia Institute of Technology in 1956.¹⁰ Because the method appeared needlessly cumbersome, in the light of extant literature on the subject of strong-signal enhancement, the alternative method of the earlier letter¹ was developed independently at the University of Michigan.

The method is indeed diagrammatically similar to that reported by Baghdady⁴ with the notable exception that this method requires a linear path (*B*) instead of a second nonlinear path. Experience in the laboratory shows this to be necessary; it is well known that nonlinearities will lead to the weak-signal suppression effect.

The combination of the limiter path (*A*) and the linear path (*B*) was successfully implemented and demonstrated in February, 1958 and thereafter at the University of Michigan. It does suffer the dependence on input level and amplitude ratio which Baghdady mentions but should be applicable to areas enjoying locally strong signal strengths and an absence of multipath effects. The incorporation of AGC within the linear path for the stabilization of output level is one solution to part of the problem. A similar application of AGC around the common driver stage would be preferable to the introduction of a limiter stage at this point; any nonlinear device will operate to the detriment of the weaker signal which is sought.

The purpose of the first letter on this subject was to record our weak-signal detection method which has the attribute of simplicity. It was not intended to detract in any way from Baghdady's excellent work in the field of FM interference. His publication record in this area stands him in good stead. However, we do believe the generation of ideas to be a wide-open field which all of us may till.

H. W. FARRIS
Elec. Eng. Dept.
University of Michigan
Ann Arbor, Mich.

* Received by the IRE, November 24, 1958. This work was supported in part by the U. S. Army (Signal Corps), the U. S. Air Force (Office of Sci. Res., Air Res. and Dev. Command), and the U. S. Navy (Office of Naval Res.).

¹ H. W. Farris, "Alternative detection of cochannel FM signals," *Proc. IRE*, vol. 46, pp. 1876-1877; November, 1958.

² An application for a patent that covers this system has been filed by E. J. Baghdady.

³ E. J. Baghdady, Classified Quarterly Progress Reports on Contract DA-36-039sc-73195, General Electronic Labs., Inc., Cambridge, Mass.; May, 1957, September, 1957, December, 1957, *et seq.*

⁴ E. J. Baghdady, "Feedforward Across the Limiter," *Quart. Prog. Rep.*, Res. Lab. of Electronics, M.I.T., Cambridge, Mass., pp. 52-53; October 15, 1957.

⁵ E. J. Baghdady, "Theory of feedback around the limiter," 1957 IRE NATIONAL CONVENTION RECORD, pt. 8, pp. 176-202.

⁶ R. H. Small, "Single-channel, two-carrier FM multiplex system," M.S. thesis, Dept. of Elec. Eng., M.I.T., Cambridge, Mass.; June 3, 1958.

⁷ E. J. Baghdady, "Capture of the Weaker Signal: Feedforward Technique," *Quart. Prog. Rep.*, Res. Lab. of Electronics, M.I.T., Cambridge, Mass., pp. 56-62; July 15, 1958.

⁸ E. J. Baghdady, "Signal-cancellation techniques for capturing the weaker of two cochannel FM signals," presented at the Internat. Conf. on Electromagnetic Wave Propagation, Liege, Belgium; October 6-11, 1958. To be published in the Proceedings of the Conference (Academic Press, London, New York).

⁹ Received by the IRE, December 15, 1958. This work was supported in part by the U. S. Army Signal Corps.

¹⁰ *Prog. Letter No. 38*, Contract No. AF-30(602)-673, "Study Program for Investigation to Aid in Reduction and Prevention of UIIF Interference," September 1, 1956.

Low-Frequency Prototype Traveling-Wave Reactance Amplifier*

A low-noise UHF traveling-wave reactance amplifier that used special junction-diode nonlinear capacitors mounted in a semidistributed transmission line was recently described.^{1,2} The purpose of this letter is to compare theory and experimental results obtained with a low-frequency prototype of a similar traveling-wave reactance amplifier. This prototype amplifier was built because of the availability of suitable commercial junction-diode nonlinear capacitors,

Furthermore, the pump line has a balanced configuration and a low characteristic impedance to minimize interaction with the signal line (but at the expense of increased pump power).³

Fig. 1 is a general schematic diagram of the amplifier. Both transmission lines are constant- K types with m -derived terminal half-sections to approximate true image terminations. The N pairs of junction-diode nonlinear capacitors are reverse biased with the same magnitude of voltage V . The diodes in each pair are effectively in parallel to constitute the entire shunt capacitance of

plished by having the cut-off frequency of the pump line (f_p) equal to twice that of the signal line (f_s), and by also having two sections of pump line per section of signal line. In this way, also, the pump line is operated well below cut-off, which results in almost lossless propagation of the pump voltage and uniform pumping of the diode capacitors.

The experimental amplifier used the parameter values given in Table I (refer also to Fig. 1). Measurements on the experimental amplifier, together with theoretical values from the relations⁴ in Fig. 1, are summarized in Table II.

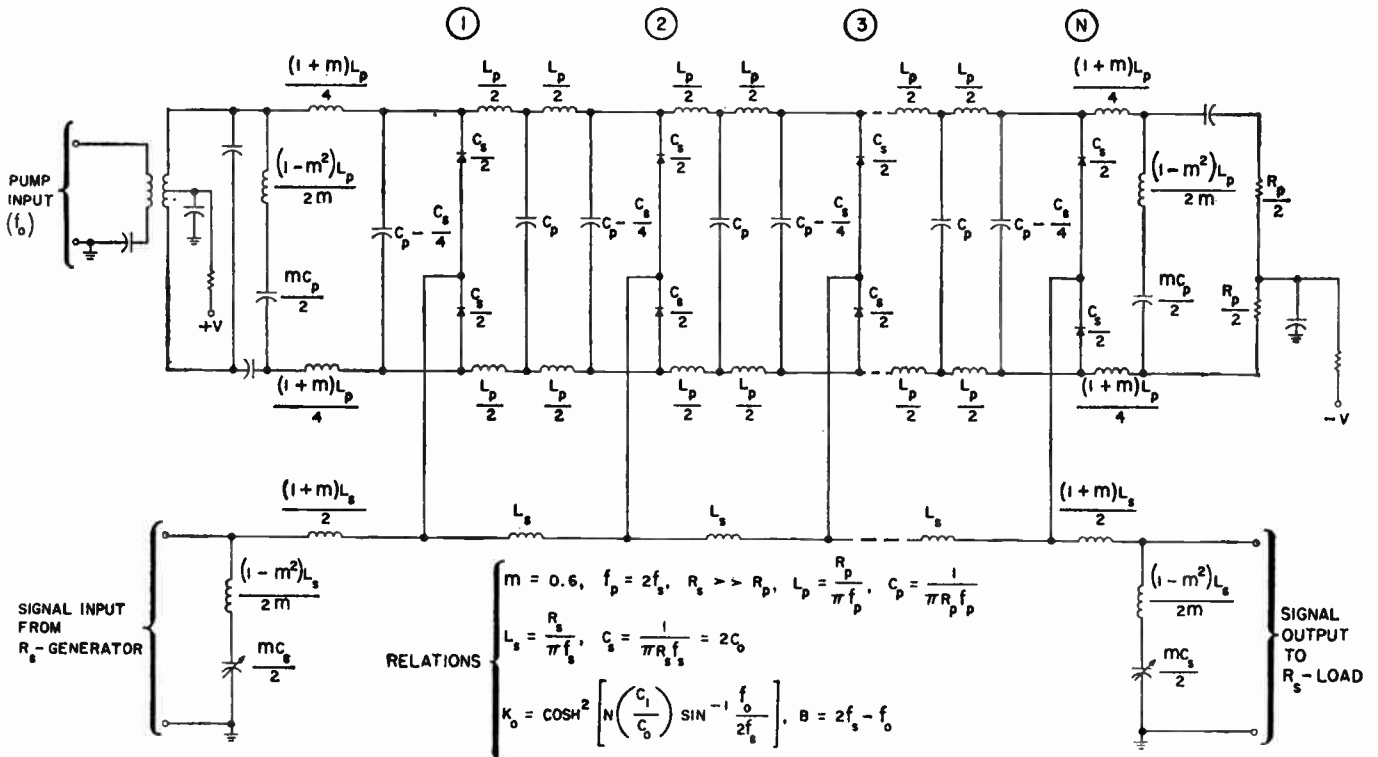


Fig. 1— N -section prototype traveling-wave reactance amplifier.

TABLE I
PARAMETER VALUES OF EXPERIMENTAL AMPLIFIER

f_0 (pump frequency)	5.0 or 6.1 mc
$f_s = 4.5$ mc ($f_p = 9.0$ mc)	
R_s (characteristic impedance of signal line)	750 ohms
R_p (characteristic impedance of pump line)	50 ohms
N	12
V	3.0 volts ($f_0 = 5.0$ mc) or 2.5 volts ($f_0 = 6.1$ mc)
Diodes	matched pairs of Pacific Semiconductors Inc., Type V33 Varicap
$C_1/C_0 \approx 0.25$	(The instantaneous capacitance of each diode as a function of pump voltage is considered to be $c(t) = C_0 + 2C_1 \cos(2\pi f_0 t - 2n\theta_0)$, where n is the number of the diode pair, and the phase shift per section of pump line is $\theta_0 = 2 \sin^{-1} [f_0/f_p]$).

and because of the relative ease of measurements at lower radio frequencies. Two separate low-pass artificial transmission lines are used in the amplifier, one for the pump and the other for the signal and idler frequencies.

* Received by the IRE, February 12, 1959.
¹ R. S. Engelbrecht, "A low-noise nonlinear reactance traveling-wave amplifier," presented at the Solid-State Device Res. Conf., The Ohio State University, Columbus, Ohio; June, 1958.
² R. S. Engelbrecht, "A low-noise nonlinear reactance traveling-wave amplifier," Proc. IRE, vol. 46, p. 1655; September, 1958.

TABLE II
THEORETICAL AND MEASURED PERFORMANCE OF EXPERIMENTAL AMPLIFIER

	$f_0 = 5.0$ mc		$f_0 = 6.1$ mc	
	Measured	Theoretical	Measured	Theoretical
Midband available power gain (K_0)	6.1 db	9.6 db	8.4 db	13.5 db
Bandwidth (B)	3.8 mc	4.0 mc	2.7 mc	2.9 mc
Effective input noise temperature	$45 \pm 20^\circ K$	$0^\circ K$	$65 \pm 20^\circ K$	$0^\circ K$
Over-all effective input noise temperature	$120 \pm 10^\circ K$	$27^\circ K$	$117 \pm 10^\circ K$	$11^\circ K$
Pump power	90 mw	—	60 mw	—

the signal line, and are effectively in series to constitute only a small portion of the shunt capacitance of the pump line. As required for wideband amplification, the effective velocity of propagation on the pump line is, in general, approximately equal to that on the signal line, and it is theoretically exactly equal for a signal frequency equal to half of the pump frequency. This is accom-

It is seen from Table II that raising the pump frequency increases the gain at the expense of decreasing the bandwidth, as predicted. Furthermore, the measured and theoretical values of bandwidth are in good agreement. By contrast, the agreement is poor between measured and theoretical values of gain. This can be attributed to slight variations in C_1/C_0 and slight reflections from section to section of the signal line. An additional factor may be the slight

³ A similar balanced configuration is shown by R. S. Engelbrecht, "Nonlinear-reactance (parametric) traveling-wave amplifiers for UHF," *Digest of Technical Papers*, presented at the Solid-State Circuits Conf., University of Pennsylvania, Philadelphia, Pa., p. 9, Fig. 7; February, 1959.

⁴ The gain and bandwidth relations in Fig. 1 are based on P. K. Tien and H. Suhl, "A traveling-wave ferromagnetic amplifier," Proc. IRE, vol. 46, pp. 700-706; April, 1958.

interaction between signal and pump lines, because the low-impedance pump line is connected in the ground return circuit of the high-impedance signal line in a complex manner. The effective input noise temperatures listed in Table II require some discussion.

In stating that the theoretical effective input noise temperatures are 0°K , it is assumed that the inductances in the signal line are lossless and the diodes used are ideal nonlinear capacitors. Actually, losses in the inductances and the diodes, although very small at the operating frequency, do contribute to the noisiness of the amplifier. A theory for this has not yet been worked out.

With regard to the measured effective input noise temperatures in Table II, these are actually calculated from separate spot noise measurements near midband ($f_1 = 2.15$ mc) on a postamplifier and on the over-all combination of the traveling-wave reactance amplifier and the postamplifier. The relation used for this calculation is

$$T_1 = T_{12} - \frac{T_2}{K_1 + K_2}, \quad (1)$$

where T_1 , T_2 , and T_{12} are, respectively, the effective input noise temperatures of the traveling-wave reactance amplifier, the post-amplifier, and the over-all amplifier. Also, K_1 and K_2 are, respectively, the available power gains of the traveling-wave reactance amplifier from signal input to signal output frequencies and from idler input to signal output frequencies. Similarly, the theoretical values of over-all effective input noise temperature listed in Table II are computed from (1). Here it is assumed that $T_1 = 0^\circ\text{K}$, $K_1 = K_2 =$ theoretical gain listed in Table II, and $T_2 = 490^\circ\text{K}$ (a measured value). It should be noted that the above noise measurements are all made with a broad-band noise source.⁵ Thus, the actual signal-to-noise ratio for a narrow-band signal will be further degraded by any idler noise from the generator and the reactance amplifier.

Figs. 2 and 3 are photographs of the detected output of the amplifier in response to a swept frequency input for $f_0 = 5.0$ mc and 6.1 mc, respectively. Because the traveling-wave reactance amplifier generates an idler frequency output ($f_2 = f_0 - f_1$) as well as a signal frequency output (f_1) in response to a signal frequency input (f_1), it was not sufficient to connect a single broad-band detector to the output terminals to view the swept response. Instead, a pair of detectors was used, one preceded by the low-pass section and the other preceded by the high-pass section of a complementary filter having a crossover frequency equal to half the pump frequency. The two detector outputs were then displayed on a dual-trace oscilloscope (Tektronix 545 with 53/54C amplifier) with the low-pass output on the upper trace and the high-pass output on the lower trace. This separation of the output into low-pass and high-pass channels is shown in the upper pair of traces in both Figs. 2 and 3, which show the output with the pump voltage removed. The middle pair of traces show the

⁵ The broad-band noise source used was a temperature limited diode, whose resistive termination was immersed in a liquid nitrogen bath for greater accuracy of measurement.

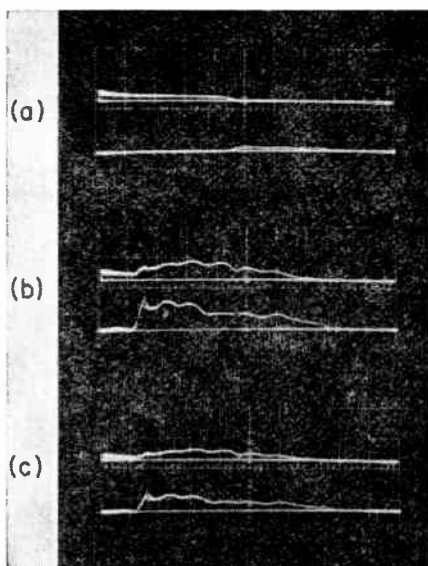


Fig. 2—Frequency response of prototype traveling-wave reactance amplifier. The pump frequency is 5.0 mc. The upper trace in each pair shows the output from 0 to 2.5 mc. The lower trace in each pair shows the output above 2.5 mc.
(a) Pump voltage removed.
(b) Normal pump voltage applied. Frequency markers appear at 0.6 and 4.4 mc.
(c) Same as (b) with 3 db less signal input.

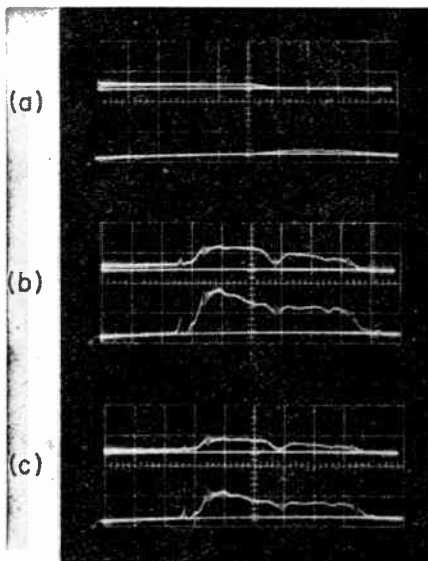


Fig. 3—Frequency response of prototype traveling-wave reactance amplifier. The pump frequency is 6.1 mc. The upper trace in each pair shows the output from 0 to 3.05 mc. The lower trace in each pair shows the output above 3.05 mc.
(a) Pump voltage removed.
(b) Normal pump voltage applied. Frequency markers appear at 1.8 and 4.3 mc.
(c) Same as (b) with 3 db less signal input.

output under normal operating conditions, with the full pump voltage applied. Frequency markers are shown to define the approximate band limits. In addition, there appears a pseudo-marker produced by the coalescing of signal and idler frequencies at half the pump frequency. The lower pair of traces are the same as the middle pair, but with 3 db less signal applied.

Figs. 2 and 3 tend to exaggerate the bumpiness of the frequency response of the traveling-wave reactance amplifier. This is because the detectors are operating in a

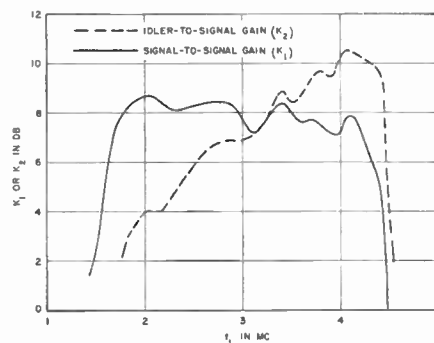


Fig. 4—Plots of signal-to-signal and idler-to-signal power gains vs signal input frequency for prototype traveling-wave reactance amplifier (pump frequency is 6.1 mc).

square-law region, and the crossover characteristics of the complementary filters are not completely abrupt. In contrast, therefore, Fig. 4 shows point-by-point plots of the signal-to-signal and idler-to-signal power gain of the amplifier vs frequency for $f_0 = 6.1$ mc.

The work described here was performed under Signal Corps Contract DA-36-039-sc-78161. We gratefully acknowledge helpful conversations with R. S. Engelbrecht and W. W. Mumford of Bell Telephone Laboratories, Inc., Whippany, N. J., and with J. C. Greene and B. Salzberg of this laboratory. We also acknowledge the help of A. Barone of this laboratory in construction and measurements.

P. P. LOMBARDO
E. W. SARD
Airborne Instruments Lab.
Div. of Cutler-Hammer, Inc.
Mineola, N. Y.

Note on Binary Decoding*

In all the literature on information coding, relatively little has been said about the quantity of logical elements and logical operations required for coding and decoding. Thus, a recent article devoted to this subject¹ breaks fairly virgin ground. A further reference to this article will be made later on, but certain aspects of the considerably more involved operation required for decoding than for coding will be reviewed briefly.

For message correction coding a coding book is required, which lists against an ordinal number the several e -error correcting messages of k bits each (which may also be $(e+1)$ -error detecting). Two general decoding procedures may be applied to the received messages.

For the first procedure, which will be termed the decoding book procedure, a decoding book is required, which is much larger than the coding book, for it contains 2^k en-

* Received by the IRE, January 7, 1959.
¹ J. H. Green, Jr. and R. L. San Soucie. "An error-correcting encoder and decoder of high efficiency." Proc. IRE, vol. 46, pp. 1741-1744; October, 1958.

tries. Against each entry there is given the ordinal number of the message transmitted, or, when extra redundancy is present, an indication may be given instead that more than ϵ errors must have occurred in transmission. Each received k -bit message is looked up in the book, and this operation requires only k successive decisions. Thus, decoding with this procedure is expeditious, but a large computer memory with 2^k entries is required.

For the second, or message matching procedure, computer memory in the form of the coding book only is required, and the message received is compared in turn to each possible transmitted message, in order to determine to which it is the closest. Alternately, and at some expense in error-detecting ability, the search may be interrupted as soon as a transmitted message at a distance equal to or smaller than ϵ has been found, without which the search is carried through to the conclusion that more than ϵ errors have occurred. This second procedure requires less computer memory than the first, but a greater number of logical operations must be performed. Letting N designate the number of possible transmitted messages, Nk modulo 2 additions of bit pairs must be made and tallied for each message.

For symbol correction coding an $n \times r$ coding matrix² is required for determining r parity checks from the n information bits by means of r additions modulo 2. Here also, two general decoding procedures, similar to those examined above, may be followed.

For the decoding book procedure, a decoding book with 2^r entries is required. For each message received an r -bit corrector is calculated by means of the coding matrix, and the corresponding entry in the decoding book indicates which bits arrived in error, or, alternately, indicates that more than ϵ errors have occurred. For certain cases of one-error correction, a simple one-to-one correspondence can exist between the corrector and the ordinal number of the erroneous bits. As for the case of message coding, this first procedure requires a relatively large computer memory, but is also very expeditious, since the calculation of the corrector constitutes a rapid, linear operation, after which only r successive decisions are required to determine the incorrectly received bits.

The second, or message matching procedure, is exactly the same as for message coding, and is subject to the same remarks.

While this second procedure appears awkward and lengthy at first glance, there are cases in which it requires appreciably simpler equipment than the first procedure, and is not inordinately lengthier. A case in point is that treated in the article cited (although the decoding procedure described in that article constitutes an interesting particular case which does not fit within the broad classifications considered above, as will be seen later on). Thus, when the messages are Huffman's m -sequences of length 15 and the all-zero sequence, or their complements, the second procedure requires specifically the following essential logical com-

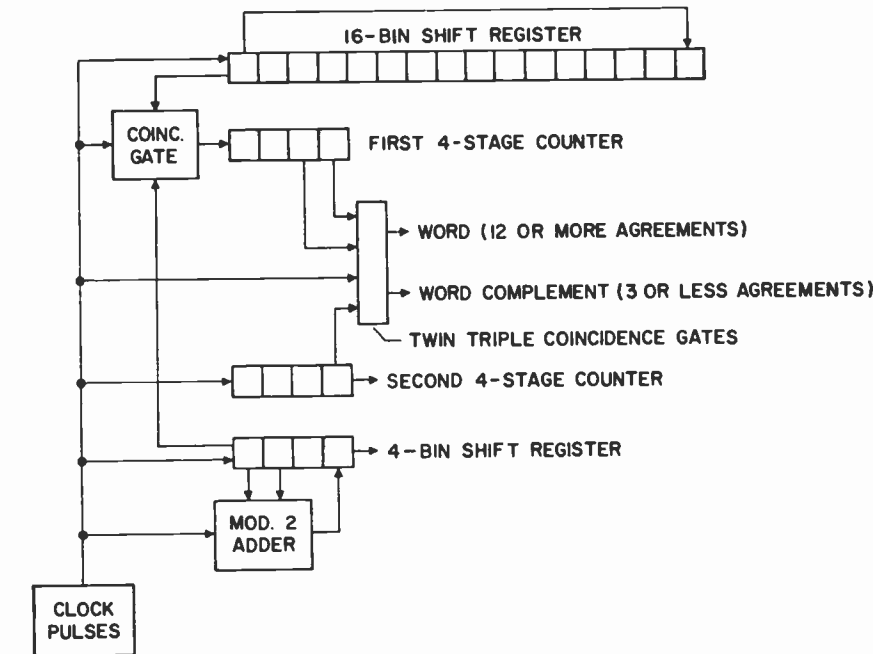


Fig. 1—Synopsis of logical elements required for a case of decoding by word matching.

ponents: a four-bin shift register, with feedback through a modulo 2 adder, a sixteen-bin recycling shift register in which the 15 received bits of a message are stored with a separating blank, two four-stage binary counters, and a few coincidence gates (Fig. 1). In operation, the sixteen-bin shift register and the four-bin shift register are actuated 15 times, as a received word is being stored in the sixteen-bin shift register, and a tally is kept in the first counter of the number of agreements noted between the states of the first bins of the two shift registers. On the fifteenth count, the second counter, which times the operation, sends a pulse to a triple coincidence gate which serves to determine whether there has been a tally of 12 or more agreements or of 3 or fewer agreements, indicating that the first word tried, or its complement, is indeed the word received, with at most 3 errors. In the absence of an output of one of the triple gates, the word stored in the sixteen-bin shift register is shifted back in the starting position at the sixteenth count, and the operation is repeated with the Huffman m -sequence shifted one bin. An average of 128 and a maximum of 256 time units are required to decode a message, and received messages with the distance 4 or more to all transmitted messages can be detected, if additional gating arrangements are provided.

It will be noted that the decoding procedure described in the article cited, while based on a completely different principle, is fairly close to the one just described when basic logical components and number of operations required are considered. These logical components are: a sixteen-bin recycling shift register, eight adders modulo 2, a two-stage and a four-stage binary counter, a five-bin shift register and an eight-bin shift register used as a ring circuit. In operation, the seven modulo 2 additions, $a_2 + a_3$, $a_3 + a_9$, etc., are made successively at the command of the ring circuit, and the count

of 1's tallied in the two-stage counter. As soon as a count of 4 is reached, a "1" is stored in the first bin of the five-bin shift register, otherwise a "0" is left there. At the eighth position of the ring circuit, the eighth modulo 2 adder compares the binit just stored in the five-bin register to the state of the first bin of the sixteen-bin register, and the result is tabulated in the four-stage counter.

As the ring circuit returns to its original position, both shift registers are shifted one bin, the two-stage counter is reset and the operation is repeated. After the fifteenth operation, the count tabulated by the four-stage counter should be either 3 or less, or 12 or more, if no more than three errors have occurred in transmission, and whichever is which determines whether the number stored in the five-stage shift register, or the complement of the number, defines the last 5 bits of the message transmitted, and hence that entire message.

It can be verified that the operation just described requires always 120 time units, as against the 128 time units required on an average for the procedure first described, but that approximately half again more logical elements are required. Received messages with the distance 4 or more to all transmitted messages are also detected, but this decoding process could be shortened to from 40 to 56 time units if this 4 error detection feature were given up and the assumption were made that no more than 3 errors will ever occur.

It is concluded that while the two general decoding procedures described at the beginning of this note are basic, other procedures such as the one described in the article cited, and tailor-made to a particular code, may provide interesting compromise solutions.

MARCEL J. E. GOLAY
Philco Corp.
Philadelphia, Pa.

² M. J. E. Golay, "Notes on digital coding." PROC. IRE. vol. 37. p. 657; June. 1949.

Diurnal Change of Ionospheric Heights Deduced from Phase Velocity Measurements at VLF*

The variation of the phase of VLF waves propagated over great distances has been reported recently by Pierce^{1,2} and Crombie, *et al.*^{3,4} In these investigations the relative phase of the carrier signal from the radio station GBR in Rugby was recorded at great distances, *i.e.* Cambridge, Mass. and Wellington, N. Z., respectively. In both cases, the received signals are compared with the output of a stable standard crystal oscillator. The experimental results show that the carrier phase shift is almost directly proportional to the length of the path which is in daylight.

The purpose of this note is to discuss the experimental data of Pierce and Crombie from the viewpoint of the mode theory of VLF propagation. The theoretical model, employed previously with some success, is a sharply bounded and isotropic ionosphere which is taken to be concentric with the surface of the earth.⁵⁻¹⁰ The effective electrical constants of the ionosphere, when chosen to give best agreement with experiment, yields a reflection coefficient at highly oblique incidence, having a magnitude near unity and phase near 180°. The corresponding phase velocity v for a mode of order n (1, 2, 3, ...) in such a waveguide can be approximated by⁵

$$v \cong c(1 - C_n^2)^{-1/2} \left(1 - \frac{h}{2a}\right) \quad (1)$$

where

- $C_n = (n - \frac{1}{2})\lambda / (2h)$,
- c = velocity of light in vacuo,
- h = height of ionosphere reflecting layer,
- a = radius of the earth, and
- λ = wavelength.

At great distances (>3000 km) the attenuation of the higher order modes is sufficiently great that only the mode corresponding to $n=1$ need be considered. Thus

$$v \cong c \left(1 - \frac{\lambda^2}{16h^2}\right)^{-1/2} \left(1 - \frac{h}{2a}\right) \cong c \left(1 - \frac{h}{2a} + \frac{\lambda^2}{32h^2} + \text{small order terms}\right). \quad (2)$$

Taking differentials, it is seen that

$$\frac{\Delta v}{c} \cong - \left[\frac{h}{2a} + \frac{\lambda^2}{16h^2} \right] \frac{\Delta h}{h} \quad (3)$$

where Δv is the change in phase velocity resulting from a change of height Δh . The left-hand side can be replaced by the ratio $\Delta t/t$ where Δt is the corresponding change of phase (measured in seconds) and $t=s/c$ where s is the length of the path under consideration. The quantity $\Delta t/t$, which is available from the experimental data, can thus be converted to the fractional change of height $\Delta h/h$, provided a reasonable assumption is made of the mean value of the factor

$$\left[\frac{h}{2a} + \frac{\lambda^2}{16h^2} \right].$$

From the examination of the amplitude data, the mean height h between day and night is about 80 km.⁶ Thus for a frequency of 16 kc the factor of proportionality should be replaced by 0.0097 and then

$$\frac{\Delta h}{h} \cong -1.03 \times 10^{-2} \times (\Delta t/t). \quad (4)$$

From Pierce's data,^{1,2} it is seen that the phase shift Δt from an all daylight path to all night path across the Atlantic ($d \cong 5200$ km) is about $-34 \pm 1 \mu\text{sec}$ throughout the year and is "surprisingly constant." The corresponding fractional change of height is thus

$$\frac{\Delta h}{h} \cong 0.203 \quad \text{or} \quad \Delta h \cong 16 \text{ km}. \quad (5)$$

Pierce quotes values of Δh of the order of 10 to 12 km which he says can be deduced from an analysis of a three-hop mode. In the same paper he writes, "Calculations too intricate to be worth reproducing here indicate that, at 16 kc, a 35- μsec diurnal range might be produced by a change in virtual height of reflection as small as 4 or 5 km, or as large as 16 or 18 km." Our value of 16 km, having an uncertainty of about ± 1 km, is based on the waveguide mode concept where it is not necessary to reject or select a discrete sequence of hops.

Similar deductions can be made from the data of Crombie, *et al.*^{3,4} From the curves shown in Fig. 3 in the second paper,⁴ it can be seen that the Δt is about $-70 \pm 2 \mu\text{sec}$ when an 11,000-km portion of the path changes from daylight to night. This readily leads to $(\Delta h/h) \cong 0.19$ or $\Delta h \cong 15$ km which would have an uncertainty of ± 1.5 km as a result of inaccuracies in reading the curves in the paper. This value for Δh is somewhat larger than the 10 km deduced by Crombie, *et al.*⁴ from their data. Two explanations for the discrepancy are offered. The first is that Crombie assumes a reflecting layer with a zero phase shift on reflection; secondly, he neglects the curvature of the earth.

The 15- or 16-km increase of effective reflection height from night over day has been

deduced here using a rather simple model which, nevertheless, is consistent with the interpretation of field strength vs distance data in the same frequency range.^{6,7} Furthermore, the diurnal change of reflection height is in good agreement with values deduced by the Cavendish group from the analysis of the interference pattern of the ground wave and the first-hop sky wave at 16 kc.¹¹

A possible objection to the present analysis is that the ionospheric reflection coefficient is assumed to be the same for day and night. In this sense, the deduced change of height should be regarded as an effective value. More realistically, the effective ionospheric conductivity at night is increased by about a factor of 3 over that for daytime. For 16 kc such an increase of conductivity has the effect of increasing the phase velocity by about 0.02 per cent.⁸ The modified form of (4) thus reads

$$\frac{\Delta h}{h} \cong 1.03 \times 10^{-2} \left[\frac{\Delta t}{t} + \delta \right] \quad (6)$$

where δ is the fractional increase of phase velocity corresponding to the increase of conductivity. From above it was found that $\Delta t/t$ is of the order of 2×10^{-3} and if δ is 2×10^{-4} the previously deduced values of the height increment should be increased by 10 per cent leading to $\Delta h \sim 17 \pm 1$ km. It thus appears that the assumption of a constant 180° phase shift on reflection is to underestimate slightly the diurnal height change.

JAMES R. WAIT
National Bureau of Standards
Boulder, Colo.

¹¹ W. C. Bain, R. N. Bracewell, T. W. Straker and C. H. Westcott, "The ionospheric propagation of radio waves of frequency 16 kc/s over distances of about 540 km," *Proc. IRE*, vol. 99, pp. 250-260; February, 1952.

* Received by the IRE, January 5, 1959.
¹ J. A. Pierce, "The diurnal carrier phase variations of a 16-kc transatlantic signal," *Proc. IRE*, vol. 43, pp. 584-588; May, 1955.
² ———, "Intercontinental frequency comparison by very low-frequency radio transmission," *Proc. IRE*, vol. 45, pp. 794-803; June, 1957.
³ A. H. Allen, D. D. Crombie, and W. A. Penton, "Frequency variations in New Zealand of 16 kc/s transmissions from G.B.R. Rugby," *Nature*, vol. 177, pp. 178-179; January 28, 1956.
⁴ D. D. Crombie, A. H. Allan, M. Newman, "Phase variations of 16 kc/s transmissions from Rugby as received in New Zealand," *Proc. IEE*, pt. B, pp. 301-304; May, 1958.
⁵ J. R. Wait, "Propagation of VLF pulses to great distances," *J. Res. NBS*, vol. 61, pp. 187-203; September, 1958.
⁶ J. R. Wait, "Transmission loss curves for VLF propagation," *IRE TRANS. ON COMMUNICATION SYSTEMS*, vol. CS-6, pp. 58-61; December, 1958.
⁷ W. Taylor and J. Lange, "Some characteristics of VLF propagation using atmospheric waveforms," in "Proc. Second Conference on Atmospheric Electricity," Pergamon Press, London, England; in press.
⁸ A. G. Jean, "Spherics" *Proc. 12th URSI General Assembly*, Boulder, Colo., pt. IV, pp. 1-50; September, 1957.
⁹ J. R. Wait and A. Murphy, "Multiple reflections between the earth and the ionosphere in VLF propagation," *Geofisica Pura e Applicata*, vol. 35, pp. 61-72; 1956. (Discusses the influence of the earth's magnetic field.)
¹⁰ J. R. Wait, "A study of VLF field strength data—both old and new," *Geofisica Pura e Applicata*. (In press.)

On the Very-Wide-Band Balun Transformer*

In the paper by O'Meara and Sydnor,¹ the use of ferrite material in the construction of a very wide balun has been described. In our study on broadband interference problems, the development of a very wide balun was also required, and a similar use of ferrite has been made which resulted in satisfactory characteristics. Although the basic principle involving the use of ferrite is the same in our application, the construction of the balun itself is different, and it might be worthwhile to comment on our results for those interested.

Considering a simple phase inverter as in Fig. 1(a), our application has the transition between the two coaxial circuits at a point just inside the balun box, whereas the transition is made in the center of the box in O'Meara and Sydnor's paper. An equivalent

* Received by the IRE, December 24, 1958.
¹ T. R. O'Meara and R. L. Sydnor, "A very-wide-band balun transformer for VHF and UHF," *Proc. IRE*, vol. 46, p. 1848; November, 1958.

circuit for the inverter is indicated in Fig. 1(b), where T represents an ideal 1:1 phase inverting transformer, Y_s is an admittance shunting between points a and b . This shunting admittance results from the necessary physical structure which consists of the outer surface of the outer conductor of cable bc and the inside surface ca of the box. The stray inductance between a and b is neglected.

The value of Y_s can be measured at the terminal plane aa' of Fig. 2 by means of a standing wave meter or an admittance meter. Our objective was to develop a configuration that maintains a much smaller value of Y_s than the characteristic admittance of the unbalanced circuit in the frequency range from 60 mc to 1000 mc which was our range of interest.

After several experiments, the use of ring shaped ferrite material placed around the cable between bc as indicated in Fig. 2, produced the best data. This is presented in Fig. 3. In the frequency range from 30 mc to 1000 mc, Y_s is much smaller than the characteristic admittance of the 50-ohm cable which is 20 millimhos.

Originally, the variation of Y_s with frequency was expected to move along a spiral locus which converges into the characteristic admittance of the construction, but in the measured data, the movement of Y_s is very slow and this can be interpreted that the equivalent length is long, and the value of Y_s at each frequency corresponds to the characteristic admittance at that frequency.

For the ferrite rings, four cup core type transformer cores made of General Ceramic Q material are placed around the cable which is three inches long. Each core has a length of five-eighths inch and a diameter of one and three-eighths inches. Because our lower frequency of interest is 60 mc, the values of Y_s below 30 mc have not been investigated, but it is expected that low values will be maintained well below this frequency.

Using the above techniques, it is easy to construct wide band balun transformers with impedance ratios of 1:1 or 4:1 by appropriate connection of the coaxial cables at the terminal point. For a balun with a 1:1 ratio, the construction of Fig. 4(a) can be used. Its equivalent circuit is given in Fig. 4(b) in which the impedance of the coaxial cables inside the box should be one half the impedance of the unbalanced side. In Fig. 4(b), θ indicates the transmission line length.

No current exists on the outside of the outer conductor of the cable between a and b in this figure, so it can be arranged in any convenient manner within the balun box. Cable cd may be arranged in a curved form to keep its length the same as cable ab . The ferrite rings are placed around cable cd . It is essential that the lengths and inside characteristics of cables ab and cd be identical.

A balun with a 4:1 impedance ratio is illustrated in Fig. 5(a) with its equivalent circuit given in Fig. 5(b). In this case, the characteristic impedance of the inner coaxial cables should each be twice that of the unbalanced circuit.

The data taken applies to a 4:1 balun. The insertion loss for the desired balance to unbalance transmission is shown in Fig. 6(a),

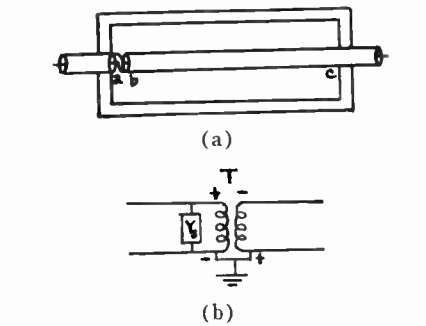


Fig. 1—Phase inverter.

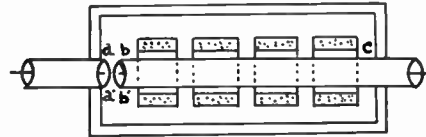


Fig. 2—Measurement of shunting admittance.

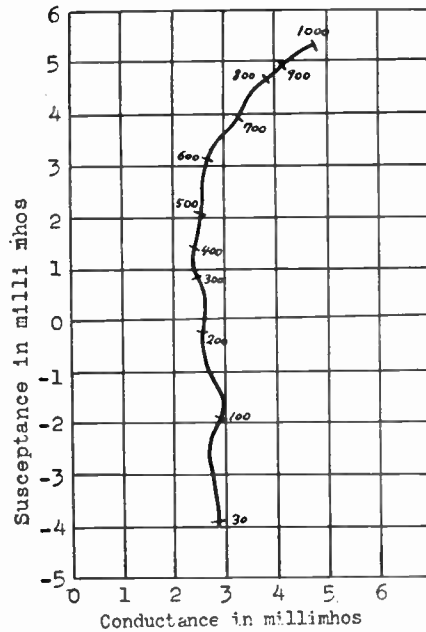


Fig. 3—Variation of shunting admittance with frequency (numbers along curve indicate mc).

and the transmission of the undesired unbalance to unbalance mode is illustrated in Fig. 6(b). The insertion loss is somewhat larger than the value expected from Y_s of Fig. 3, but this difference is considered to have come from some difference of constructional dimensions. There should be no transmission of the undesired mode from the nature of construction if the inside characteristics of the two balanced side cables are identical and if the stray inductances at the transition point are negligible. However, this ideal is not achieved and an undesired transmission exists, although it is less than -30 db over the frequency range of our interest.

To decrease stray inductances and stray capacitances, the use of small cable is sug-

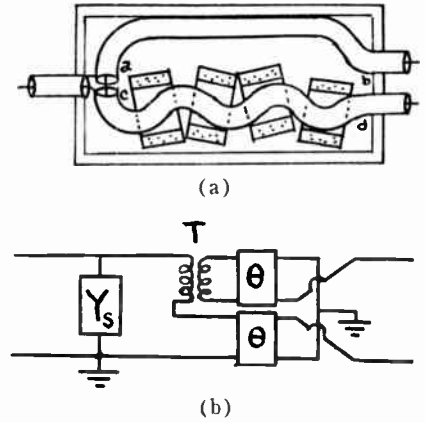


Fig. 4—1:1 Ratio balun.

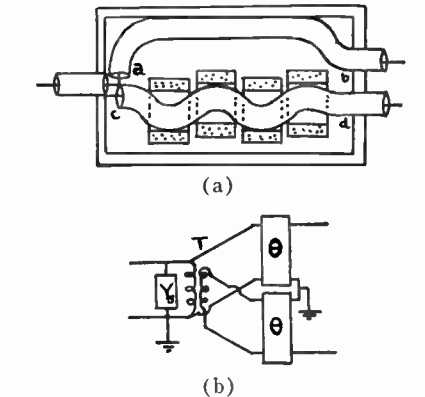


Fig. 5—4:1 Impedance ratio balun.

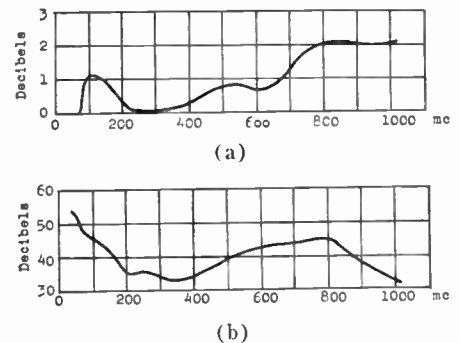


Fig. 6—(a) Insertion loss of balance to unbalance transmission. (b) Attenuation of unbalance to unbalance transmission.

gested. Also these effects might be controlled by suitable values of inductance and capacitance as mentioned in O'Meara and Sydnor's article.

Summarizing, our method of constructing wide-band baluns appears to be simpler compared to that mentioned in O'Meara and Sydnor's paper in the sense that only one side of the coaxial structure is required to have a small value of shunt admittance. This can be achieved by placing the rings of ferrite material around one of the inner coaxial cables.

MASAKAZU OSHIMA
Electrical Eng. Dept.
Rensselaer Polytech. Inst.
Troy, N. Y.

A Medium Power Ferrimagnetic Microwave Limiter*

A medium power microwave ferrimagnetic limiter has been built using a magnesium manganese ferrite, General Ceramics R-1. The limiter was developed for X band and was operated over a narrow band of frequencies centered at 9360 mc. By variation of the magnetic field, the limiter could be adjusted to attenuate incoming signals lying between 30 and 1000 watts peak to an output of 100 mw. The untransmitted portion of the input power was dissipated by absorption in the ferrite. The absorption loss was presumed to be produced by microwave coupling to a long wavelength spin wave.

The ferrite was used in plate form with a broad face of the plate in contact with the narrow waveguide wall. The width of the plate was approximately 0.13 inch, which was less than the minimum thickness necessary to support the lower order dielectric waveguide modes. The external static magnetic field was skewed. The field skewing subjected the ferrite to two principal static magnetic field components. One magnetic field component was normal to the narrow waveguide wall, the other field component was parallel to the direction of wave propagation. The field component normal to the broad face of the ferrite was variable from 200 to 1100 oersteds, depending upon the desired threshold power setting. The power threshold is here defined as the input power at which the limiter insertion loss begins to deviate from the low power value. The magnetic field component, parallel to the direction of propagation, was somewhat less than 100 oersteds.

A set of performance curves for a limiter containing ferrite R-1 is shown in Fig. 1. The transmission loss in db is plotted vs input power in db above 1 mw (dbm). The four curves correspond to differing adjustments of the intensity and direction of the skewed magnetizing field. These measurements were made using 0.8 μ sec input pulses. A very low repetition frequency < 100 cps was used to avoid heating the ferrite. For a selected input power the magnetic field was adjusted to approximately maximize the absorption loss and to provide a net attenuation that yielded a transmitted power level of 20 dbm. In Fig. 2 are plotted transmission losses at 60 dbm, 0 dbm and the VSWR at 0 dbm vs frequency for the condition corresponding to curve D of Fig. 1.

Curve D of Fig. 1, showing a maximum transmission loss at 60-dbm input power, was obtained with a static magnetic field component perpendicular to the narrow waveguide wall of approximately 1100 oersteds. The required perpendicular static magnetic field component was found to decrease as the input power, at which peak absorption occurred, decreased. Curve A of Fig. 1 was obtained using a field component normal to the plate of 200 oersteds, curve B, using approximately 500 oersteds and curve C, using approximately 800 oersteds. Two absorption maxima are seen in curve A, one occurring at 45 dbm and the other at 64 dbm. The two loss maxima are presumed

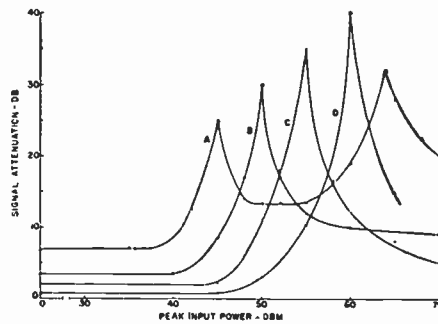


Fig. 1—Limiter transmission loss as a function of peak input power for various adjustments of the static magnetic field.

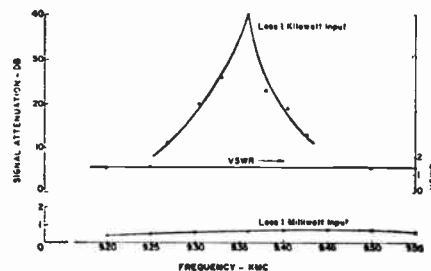


Fig. 2—Frequency response for a limiter containing ferrite R-1—corresponding to curve D of Fig. 1.

to be associated with separately excited spin wave disturbances. The loss maximum occurring at the lower input power could be attributed to a long wavelength spin disturbance. The higher power loss peak would then be ascribed to a shorter wavelength spin wave with a higher power excitation threshold.

In Fig. 1 only a single absorption maximum was observed for curves B and C up to 10-kw input power and for curve D up to 3-kw input.

As is clear from Fig. 1, the low power insertion loss at 0 dbm is an inverse function of the power threshold. The low power loss, although of magnetic origin, is not attributable to the low field loss effect, described by Polder and Smit,¹ which arises in a partially saturated ferrite sample since the removal of the magnetic field reduces the low power limiter loss to a small fraction of a db. When polycrystalline yttrium garnet of approximately 50-oersteds uniform precessional line width and dielectric loss tangent of 0.003 at 20 mc was substituted for R-1, the limiter performance was observed to be similar with the exception that the low power loss was noticeably higher. As with ferrite R-1, the low power loss of the garnet limiter virtually vanished with the removal of the magnetizing field, indicating the loss to be magnetic in origin.

In Fig. 2 it is noticed that the limiter is reasonably well matched throughout the frequency band. However, the transmission loss at 60-dbm input power is markedly dependent upon frequency, while the loss at 0 dbm is independent of frequency. The resonance seen in the 60-dbm input power curve results from a spin wave interaction with the microwave signal.

The intrinsic line width of the spin wave, excited in the limiter at high power, can be estimated with the aid of the data of Fig. 2. From the curve of Fig. 2, plus the uniform precessional resonance curve obtained for the same ferrite sample, the ratio of effective line width of the spin wave mode and uniform precessional mode is obtainable. The intrinsic line width of the excited spin wave can then be estimated, knowing the intrinsic uniform precessional line width of the material. For ferrite R-1 used in the present limiter the estimated value of the intrinsic spin wave line width occurring at 60-dbm input power is 27 oersteds based upon a uniform precessional line width of 450 oersteds.

Parenthetically, the frequency response data taken at 60-dbm input power upon a limiter using yttrium garnet were nearly identical to those of Fig. 2 for ferrite R-1. Thus, the spin wave excited in the yttrium garnet exhibited a line width at 60-dbm input power approximately equal to the uniform precessional line width. This result suggests that the spin wave, excited at 60-dbm input power in the polycrystalline yttrium garnet, is very tightly coupled to the uniform precession motion; consequently, the garnet displays a loss which is characteristic of the uniform precessional motion only.

It has been observed that the magnitude of the absorption peaks of Fig. 1 is dependent upon the intensity of the transverse component of the RF magnetic field at the surface of the ferrite nearest the center of the waveguide. Also, the differential loss, *i.e.*, the difference between the high power transmission loss and low power loss at 0 dbm, is a direct function of the transverse microwave magnetic field component, h_t . Consequently, most design changes which enhance h_t have been found to enhance the limiting. Figs. 3 and 4 present information in support of this point. The RF magnetic field component transverse to the propagation direction was measured with a magnetic field probe consisting of a coupling iris and polarized detector mounted on the moveable top wall of a rectangular waveguide. The garnet samples used in the studies of Fig. 3 and 4 were uniformly magnetized by a homogeneous static field. The use of a uniform magnetic field reduced the differential loss relative to that achievable with a skewed magnetic field but did not suppress the limiting effect. The uniform magnetic field was employed to increase the precision and reproducibility of the data. In Fig. 3(a) a sample of polycrystalline yttrium garnet of width (0.140 inch) and length (0.7 inch) was varied in height from 0.384 to 0.399 inch. The value of h_t at 60-dbm input power for a frequency of 9360 mc was measured at the slab surface nearest the waveguide center as a function of the plate height. The detected signal plotted in Fig. 3 is normalized to the signal for a plate height of 0.384 inch. In Fig. 3(b) the difference in transmission loss measured at 60-dbm input power and at 0-dbm input (the differential transmission loss) is recorded vs h_t with the same static magnetic field as used for the data of Fig. 3(a). The differential loss is seen to rise rapidly with an increase in h_t .

In Fig. 4(a) h_t at the slab surface nearest the waveguide center is plotted vs the trans-

¹ D. Polder and J. Smit, *Rev. Mod. Phys.*, vol. 25, p. 89; January, 1953.

* Received by the IRE, January 13, 1959.

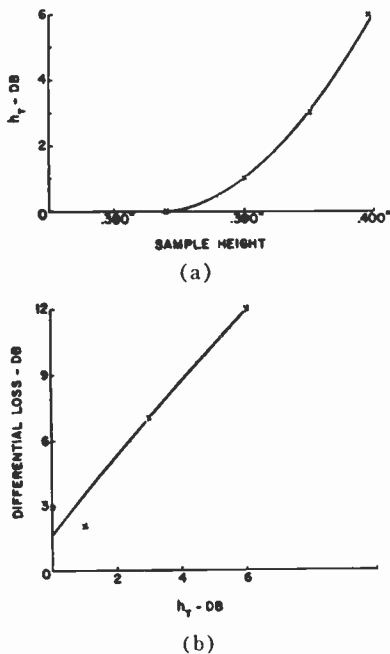


Fig. 3—Dependence of differential transmission loss upon h_t and sample height.

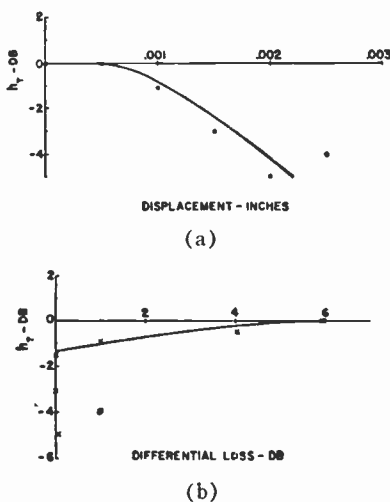


Fig. 4—Dependence of transmission loss upon h_t and sample position.

verse position of the yttrium garnet slab from the waveguide wall.

In Fig. 4(b) the differential loss is plotted vs h_t . The magnetizing field was uniform and the frequency was 9360 mc. Once again the differential loss is seen to be dependent upon h_t which decreases sharply as the slab is displaced from the narrow wall of the waveguide.

Studies of the dependence of transmission loss upon sample thickness displayed the following characteristics: as the sample thickness was increased from zero, the transmission loss at all input power levels remained approximately zero until at a critical thickness the high power transmission loss began to increase. The critical thickness lay in the range 0.1 to 0.13 inch for ferrite R-1 and polycrystalline yttrium garnet at frequencies of 9 to 9.5 kmc. For sample thicknesses above the critical value, the differential loss increased rapidly with an increase

in sample thickness, ultimately reaching a maximum value. Further increases in sample thickness produced large increases in the low power transmission loss and a reduction of the differential transmission loss.

The value of the total spike leakage was measured as 150 ergs for a 1000-watt input power in a limiter using ferrite R-1 and adjusted corresponding to curve D of Fig. 1. The total spike leakage was composed of nearly equal amounts of energy contributed by the leading and traveling edge spikes. For curves C, B and A of Fig. 1, the total spike leakage energy at input power levels of 55, 50 and 45 dbm was approximately 50, 15 and 5 ergs, respectively. The variation in spike leakage was observed to be nearly proportional to the change in input power.

E. N. SKOMAL
M. A. MEDINA
Sylvania Electric Products, Inc.
Mountain View, Calif.

Frequency Stability Requirements on Coherent Radar Oscillators*

In coherent radar systems where accuracy of velocity information is desirable, the frequency stability of the RF oscillator is of interest. At first glance, the familiar Doppler radar expression appears to serve as a basis for a stability criterion.

$$f_D = \frac{2v}{\lambda}$$

where f_D is the Doppler frequency, v is the radial velocity of a moving target, and λ is the transmitter wavelength.

If this expression were to serve as a basis for a stability criterion, in its simplest form it would demand that the transmitter frequency be held absolutely constant to assure that there would be no error in the Doppler frequency. This is not only practically impossible, but, as is shown below, is not entirely necessary.

In an ordinary pulsed radar system, the echo pulses from a stationary target bear a constant phase relationship to the transmitted pulses, the phase shift being dependent on the range of the target. In a pulsed-Doppler radar system, velocity information is obtained by measuring the change in phase shift in successive echo pulses from a target whose range is changing. The way this phase shift changes in a series of repetition periods determines the Doppler frequency. The video output of the synchronous detector on a fixed target is merely a string of pulses like the output of an ordinary radar detector. The output on a moving target, however, is a bi-polar video signal whose envelope is the Doppler

frequency. The Doppler filter following the synchronous detector can have a variable integration time to accept a certain pre-selected minimum Doppler frequency corresponding to a certain minimum radial target velocity.

The expression for the phase shift ϕ of the echo pulses in terms of the echo return time and the oscillator frequency f is:

$$\phi = 2\pi \int_0^{t_d} (f) dt = 2\pi f_{av} t_d$$

where f_{av} is the average oscillator frequency during the echo return period t_d .

Assume a drifting oscillator and a stationary target. A Doppler system will process the received information to indicate a false Doppler frequency f_f :

$$f_f = \frac{1}{2\pi} \frac{d\phi}{dt} = t_d \frac{df_{av}}{dt}$$

If df_{av}/dt is not zero but is a constant during the integrating time of the Doppler filter, then f_f represents an error in the apparent Doppler frequency observed. If df_{av}/dt is not a constant, the envelope of the detector output is not a sinusoid. Two effects can then be defined: the short-term effect in which the changes of f_{av} are such that df_{av}/dt is not constant throughout the Doppler filter integration period, and the medium-term effect in which changes of f_{av} are such that df_{av}/dt is appreciably different from zero, yet essentially constant over the integrating period. The short-term instabilities will give rise to false moving target indications while the medium-term instabilities will give an incorrect rate of change of the angle ϕ , thus giving an error in an observed velocity of an actual target. There is, of course, also the long-term effect (changes of f_{av} over periods of minutes, hours, or even days) in which the radar data can be in error due to a slowly accumulated error in the transmitter frequency; the criterion for long-term stability is simply the $f_D = v/\lambda$ relationship. With present-day transit-time oscillators, these long-term stability problems are relatively minor with respect to the above mentioned short- and medium-term stability problems.

The rate of change of phase difference between a returning echo and a drifting oscillator can be said to be equal to the sum of the actual Doppler frequency and the false Doppler frequency. Thus, if a single scatterer is assumed, the envelope of the output of the synchronous detector has an instantaneous frequency equal to the difference of the two input frequencies and thus to the sum of the actual Doppler frequency and the drift frequency. Hence:

$$a = a_e \cos 2\pi t \left(f_D + t_d \frac{df_{av}}{dt} \right)$$

where a is the ordinate of the detector output envelope and a_e is the amplified amplitude of the echo signal. This, then, establishes the criteria for the short- and medium-term stabilities.

If the system used indicates velocity through using the $f_D = 2v/\lambda$ relationship, and an acceptable per-unit velocity error is N , then the medium-term stability criterion becomes:

* Received by the IRE, January 13, 1959. This work is based on part of a thesis submitted for the M.S. degree to the Dept. of Elec. Eng., Mass. Inst. Tech., Cambridge, Mass.; June, 1958. The thesis research was supported by the U. S. Air Force (Weapons Guidance Lab., Wright Air Dev. Center) under Contract No. AF-33(616)-5489 with the M.I.T. Servomechanisms Lab.

$$\frac{df_{nv}}{dt} \approx \frac{Nv f_{nv}}{\lambda_d}$$

The criterion for the short-term stability is that if there are short-term instabilities which cause false target indications, these indications should either fall sufficiently close to an actual target indication so as to not be resolved by the radar as separate moving targets or be of sufficiently small amplitude so as not to be detected by the data-processing sections of the system. The instantaneous frequency output of the detector when there is no actual moving target is equal to $u(df_{nv}/dt)$. The requirements on short-term stability are then fixed by the resolving capabilities of the radar system. In general, the short-term stability requirements essentially restrict the character of the short-term FM type instabilities in that both the frequencies and the amplitudes of the Fourier components of the instantaneous frequency are restricted.

It is thus seen that examining Doppler errors by assuming variations of λ in the $f_D = 2v/\lambda$ expression gives an inadequate picture. Errors of this type are actually of little interest compared to the velocity-error and false-target criteria determined by the medium- and short-term instabilities mentioned.

M. MICHAEL BRADY
Norwegian Defense Res. Establ.
Radar Div.
Bergen, Norway
Formerly with Servomechanisms Lab.
Mass. Inst. Tech.
Cambridge, Mass.

WWV Standard Frequency Transmissions*

Since October 9, 1957, the National Bureau of Standards radio stations WWV and WWVH have been maintained as constant as possible with respect to atomic frequency standards maintained and operated by the Boulder Laboratories, National Bureau of Standards. On October 9, 1957, the USA Frequency Standard was 1.4 parts in 10^9 high with respect to the frequency derived from the UT 2 second (provisional value) as determined by the U. S. Naval Observatory. The atomic frequency standards remain constant and are known to be constant to 1 part in 10^9 or better. The broadcast frequency can be further corrected with respect to the USA Frequency Standard, as indicated in the table. This correction is *not* with respect to the current value of frequency based on UT 2. A minus sign indicates that the broadcast frequency was low.

The WWV and WWVH time signals are synchronized; however, they may gradually depart from UT 2 (mean solar time corrected for polar variation and annual fluctuation in the rotation of the earth). Corrections

are determined and published by the U. S. Naval Observatory.

WWV and WWVH time signals are maintained in close agreement with UT 2 by making step adjustments in time of precisely plus or minus 20 msec on Wednesdays at 1900 UT when necessary; one retarding time adjustment was made during this month at WWV and WWVH on February 25, 1959.

WWV FREQUENCY*

1959 February 1600 UT	Thirty-day moving average, seconds pulses on 15 mc, parts in 10^{10}
1	-29
2	-30
3	-30
4	-30
5	-30
6	-30
7	-30
8	-31
9	-31
10	-31
11	-31
12	-31
13	-32
14	-33
15	-33
16	-33
17	-34
18†	-34
19	-39
20	-39
21	-39
22	-38
23	-38
24	-38
25	-38
26	-37
27	-37
28	-37

* WWVH frequency is synchronized with that of WWV.

† Decrease in frequency, ≈ 5 in 10^{10} at 1900 UT at WWV.

NATIONAL BUREAU OF STANDARDS
Boulder, Colo.

A Straightforward Description of Temperature Effects on Transistor Stages*

Recently the author published some comments on the terminology and notation used for transistors and suggested that certain changes and simplifications would be desirable.¹ Comments that have since been received suggest that quite a few other people have had similar things in mind. There was one matter, scarcely mentioned in the original remarks, on which something might be said.

Our purpose here is to make some suggestions for improvement also. The matter in question is that of the effect of temperature on the operating point of a transistor amplifier stage.

Consider this matter from the viewpoint of a student, just learning about transistor circuitry. When this matter of temperature effects comes up, he encounters something called a "stability factor." After the initial shock of discovering that the "stability

factor" is really an instability factor, he goes on to use this concept in figuring the effects of temperature changes. Typically, he will select an operating point for the stage, then judge how much the operating point may be allowed to shift under extremes of temperature, etc. From this information he will calculate a stability factor, including a safety factor, and then will proceed to insert the stability factor into various design equations.

In these last remarks lies the point from which the present objections arise. One might argue that the name "stability factor" is a misleading one, and even if it were called an "instability factor" this would still not be quite correct. By "instability," one usually means a tendency to change spontaneously, and this is not what one usually has in mind in discussing the transistor stage. In fact, what is involved is the tendency of the operating point to move due to change in temperature, etc.; thus, the number describing this might better be called a "mobility factor." This would, however, be clearly unsuitable, since the word "mobility" has a different and well-established use in semiconductor work. This matter of names, in any event, is perhaps secondary. The main objection is just this: the stability factor is a sort of figure of merit (or, rather, demerit) for a stage. Now, a designer is ordinarily not interested in figures of merit, or of demerit either, for their own sake. He is interested in the behavior of the operating point and in how it shifts as conditions change. The question might be asked, then, would it not be simpler to deal with the problem directly in terms of the operating point, and its shifting? The writer believes that it would indeed and offers the following thoughts on possible ways of doing this.

One might select a nominal operating point at collector voltage V_{cq} and current I_{cq} . (The subscript q is used because this is often called the quiescent point.) Choose an extreme operating point at collector voltage V_{cx} and current I_{cx} . (It may or may not be possible to choose I_{cx} arbitrarily, according to the type of circuit. On the other hand, V_{cx} can always be chosen arbitrarily, and usually it is the thing of greater importance.) The line joining the nominal and extreme operating points will just be the static load line; thus, supply voltages, load resistances, etc., can be determined.

According to this idea, then, one would select the nominal and extreme operating pairs by the criteria ordinarily used in selecting operating points. Then the design would proceed directly from the choice of these points, and the method would be such as to lend itself to graphical work on a family of collector characteristics. By working directly with the characteristics one can, of course, take nonlinearities into account directly, a thing that is rather difficult with purely analytical methods. Of course, this is nothing new; the utility of graphical calculations, carried out on a family of characteristics, has been demonstrated by many years of experience with vacuum-tube stages.

In conclusion, it is suggested that the more natural way of thinking about the effects of temperature changes on transistor

* Received by the IRE, January 23, 1959.

¹ H. L. Armstrong, "On the need for revision in transistor terminology and notation," *Proc. IRE*, vol. 46, pp. 1949-1950; December, 1958.

* Received by the IRE, March 26, 1959.

amplifier stages is in terms of these concepts of nominal and extreme operating points. With the help of these, many design problems can be managed rather easily by graphical constructions done on a family of characteristics.

It is hoped that a step-by-step method of design, using these ideas, will be published elsewhere.

H. L. ARMSTRONG
Queens University
Kingston, Ont., Can.

On the Computer Realization of an Orthonormal Spectrum of a Given Signal Function*

1) A given signal function can be approximated by an orthonormal series of polynomials in t of ascending powers. The exact form of these polynomials depends upon the range of definition of the analyzed function and weight attached to the error, between the actual function and its approximation, over the various intervals of the range of definition of the function. The coefficients preceding these polynomials in the over-all series form a spectrum by which the analyzed function can be identified.

When the range of definition of the function $f(t)$ that is to be analyzed is $0 < t < \infty$, and the weighting factor e^{-t} is to accentuate the discrepancy or error between $f(t)$ and the approximating sum S_n of n polynomials in t , at the origin $t=0$, while neglecting the error at large values of t , one obtains the familiar Laguerre^{1,2} expansion,

$$f(t) = \sum_{i=0}^{\infty} C_i L_i(t) \tag{1}$$

where

$$\begin{aligned} L_0 &= 1; & L_i(t) &= -t + 1; \\ L_2(t) &= \frac{1}{2} t^2 - 2t + 1; \dots \\ L_n(t) &= \frac{1}{n!} e^t \frac{d^n}{dt^n} (t^n e^{-t}). \end{aligned} \tag{2}$$

The Laguerre polynomials given by (2) are orthonormal over their range of definition,

$$\int_0^{\infty} L_i(t) L_j(t) e^{-t} dt = \delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j. \end{cases} \tag{3}$$

The C_i coefficients of the expansion (1) are given by

$$C_i = \int_0^{\infty} f(t) L_i(t) e^{-t} dt. \tag{4}$$

The computer realization of the Laguerre spectrum C_i , of the function $f(t)$ consists of synthesizing active ladder networks whose

impulse response is the cofactor of $f(t)$ in the integrand of (4). In another method, outlined at the end of this article, the impulse response of the consecutive parts of the ladder network is the mirror image of the $L_i(t)e^{-t}$ function.

2) The Laplace transform of the n th Laguerre polynomial, defined by (2), is

$$\begin{aligned} L_n(s) &= \frac{1}{n!} \int_0^{\infty} e^t \frac{d^n}{dt^n} (t^n e^{-t}) e^{-st} dt \\ &= \frac{(s-1)^n}{s^{n+1}}. \end{aligned} \tag{5}$$

Define

$$L_n'(t) = L_n(t) e^{-t}; \tag{6}$$

its Laplace transform is

$$L_n'(s) = \frac{s^n}{(s+1)^{n+1}}. \tag{7}$$

This is the transmittance between the input and the C_n terminals of the network shown in Fig. 1. When the function $f(t)$ is applied

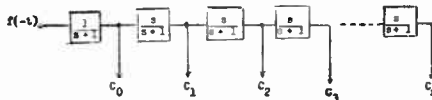


Fig. 1—The Laguerre network for the characterization of $f(t)$.

at the input terminals of the ladder network, the signal at the output of the i th section is

$$C_i(s) = F(s) \frac{s^i}{(s+1)^{i+1}},$$

$$C_i(t) = \int_0^{\infty} f(t-\tau) L_i(\tau) e^{-\tau} d\tau, \tag{8}$$

and

$$C_i(0) = \int_0^{\infty} f(-\tau) L_i(\tau) e^{-\tau} d\tau. \tag{9}$$

This integral is identical with the definition of the Laguerre coefficient, given in (4), when $f(-\tau)$, for $0 < \tau < +\infty$, has the same values as $f(t)$, for $0 < t < +\infty$. The computer network shown in Fig. 1 yields directly the Laguerre spectrum C_0, C_1, C_2, \dots , while the function applied at its terminals presents at all times—prior to the instant of measuring C_i , which is taken as $t=0$ —a mirror image of the function $f(t)$ that is to be analyzed. This implies, in practice, recording, by some method, the function $f(t)$ and then “playing” it back so as to achieve a reversed sense of time.³

The realization of the $1/s+1$ and $s/s+1$ terms is shown in Fig. 2.

3) The practical difficulties associated with the feeding in of the analyzed signal in some memory device and the later reading of it out in a reversed order are obvious, especially when the signal and its analyzer form part of a dynamic system. It is proposed in what follows to utilize the alternate form of the convolution integral (8) which, as is shown below, leads to a computer-

analyzer that accepts directly the signal $f(t)$ achieving thereby an appreciable simplification of the over-all spectrum analysis problem.

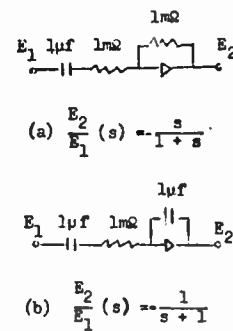


Fig. 2—The realization of $-s/(s+1)$ and $-1/(s+1)$ on an analog computer.

In practice, the signal to be analyzed is defined over a finite length of time t . Assume that the analyzed signal $f(t)$ starts at $t=0$ while its Laguerre spectrum is recorded t units of time later. The signal $f(t)$ is applied at the input terminals of a network with an impulse response $h_k(t)$. The output $C_k(t)$ is

$$\begin{aligned} C_k(t) &= \int_0^t f(t-\tau) h_k(\tau) d\tau \\ &= \int_0^t f(\tau) h_k(t-\tau) d\tau; \\ f(t) &= h_k(t) = 0 \quad \text{for } t < 0. \end{aligned} \tag{10}$$

When

$$h_k(t-\tau) = L_k'(\tau) = L_k(\tau) e^{-\tau},$$

or

$$h_k(\tau) = L_k(t-\tau) e^{-(t-\tau)}, \tag{11}$$

then C_k is the k th Laguerre coefficient of $f(t)$. From the first Laguerre polynomials L_0, L_1, L_2, \dots given by (2) and the last expression, we obtain

$$\begin{aligned} h_0(\tau) &= e^{-t\tau} \\ h_1(\tau) &= (-t + \tau + 1) e^{-t\tau}; \\ h_2(\tau) &= [\frac{1}{2}(t-\tau)^2 - 2(t-\tau) + 1] e^{-t\tau}; \\ h_3(\tau) &= [\frac{1}{6}(t-\tau)^3 + \frac{1}{2}(t-\tau)^2 \\ &\quad - 3(t-\tau) + 1] e^{-t\tau}. \end{aligned} \tag{12}$$

The respective Laplace transforms are:

$$\begin{aligned} H_0(s) &= \frac{e^{-t}}{s-1} \\ H_1(s) &= H_0(s) \frac{s}{s-1} - tH_0(s); \\ H_2(s) &= H_1(s) \frac{s}{s-1} + (\frac{1}{2}t^2 - t)H_0(s) \\ H_3(s) &= H_2(s) \frac{s}{s-1} + (-\frac{1}{6}t^3 + t^2 - t)H_0(s). \end{aligned} \tag{13}$$

The network shown in Fig. 3 realizes the successive transmittances of (13). The synthesis of the individual $1/s-1$ and $s/s-1$ blocks is shown in Fig. 4. The network of Fig. 3 is unstable and has to remain unexcited until the very moment when computation starts. When the duration t of the analyzed signal $f(\tau)$ is known in advance,

* Received by the IRE, December 18, 1958.
¹ E. N. Laguerre, “Sur l’intégral $\int_0^{\infty} x^{-1} e^{-x} dx$,” *Bull. Soc. Math. France*, vol. 7, pp. 72-81; 1879. Also see “Oeuvres de E. N. Laguerre,” Gauthier Villars, Paris, France, vol. 1, pp. 104-107, 428-437; 1898.
² R. Conrart and D. Hilbert, “Methods of Mathematical Physics,” Interscience Publishers Inc., New York, N. Y., vol. 1, pp. 93-95; 1953.
³ W. H. Huggins, “Representation and Analysis of Signals,” School of Eng., The Johns Hopkins University, Baltimore, Md., pt. 1, Rep. No. AF CRC TR-357; September 30, 1957.

the factors $-t, \frac{1}{2}t^2-t, -(\frac{1}{6})t^3+t^2-t, \dots$ are just constants. The Laguerre spectrum C_0, C_1, C_2, \dots is multiplied by e^t which can be considered as a scaling factor.

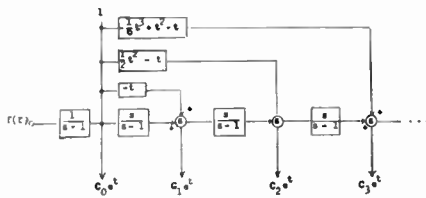


Fig. 3—Block diagram representation of a Laguerre spectrum computer.

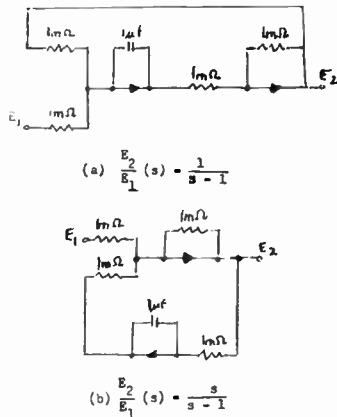


Fig. 4—Synthesis of $1/(s-1)$ and $s/(s-1)$ functions.

The last method avoids the necessity of first recording and then reading out backwards the analyzed signal $f(\tau)$. The convolution integral (10) involves integrating the product of two functions, one folded and shifted with respect to the other. The last method is based upon folding the impulse response while the first one postulates the folding of the analyzed signal. The Laguerre orthonormal series served to illustrate the method which can be expanded to other orthonormal series expansions especially to the exponential ones.

E. MISHKIN
Microwave Research Inst.
Polytech. Inst. Brooklyn
Brooklyn 1, N. Y.

What Is an Engineer?*

In regard to a recent editorial¹ on the title, "engineer," it is suggested that the word be replaced by "sciencier." A broad definition of this word is a person directly concerned with the field of applied science.

It is the opinion here that engineering education has and still is ignoring the field of semantics to the detriment of all.

ARTHUR WEINER
Sylvania Electronic Defense Lab.
Mt. View, Calif.

* Received by the IRE, January 29, 1959.
¹ "Poles and Zeros," Proc. IRE, vol. 47, p. 1; January, 1959.

Effect of Beam Coupling Coefficient on Broad-Band Operation of Multicavity Klystrons*

In my previous paper¹ the curve shown in Fig. 1 (namely, plot of normalized gap impedance vs. $\gamma_0 r_0$) is applicable when the ratio of drift tube radius b_0 to beam radius r_0 lies in the range 1.2 to 1.6 (that is, $1.2 \leq b_0/r_0 \leq 1.6$).

Also in that paper, on page 1958, second column, tenth line up from the end of the communication, the symbol should be r_0 (beam radius), and not Γ_0 .

S. V. YADAVALLI
G. E. Microwave Lab.
Palo Alto, Calif.

* Received by the IRE, February 19, 1959.
¹ S. V. Yadavalli, Proc. IRE, vol. 46, pp. 1957-1958; December, 1958.

Phase Relationships in Short Slot Hybrid Couplers*

For some years, many engineers developing microwave equipment have been using both the sidewall and the topwall short-slot hybrid couplers with complete satisfaction. We have, for example, been able to obtain economical and efficient packaging of balanced mixers with short-slot couplers.

For most applications of the short-slot coupler, it is a matter of no consequence whether the electric field in the auxiliary branch leads or lags the electric field of the directly connected branch by 90°. Moreover, it has been accepted generally¹ that in the case of the sidewall coupler the electric field in the auxiliary branch leads that in the directly connected branch by 90°, and that an interchanged lead and lag describes the situation for the topwall coupler.

In development work carried out at these laboratories in connection with a compartmented pillbox antenna, we noticed, and later verified both experimentally and analytically, that the commonly accepted phase relationships of both the sidewall and the topwall couplers are incorrect. Summarizing, we believe that the electric field in the auxiliary branch of the sidewall hybrid lags that in the main branch by 90°. Conversely, the electric field in the auxiliary branch of the topwall hybrid coupler leads that in the main branch by 90°.

With a little examination, the reason for this discrepancy becomes obvious. In the derivation contained in U. S. Patent No. 2739288, a voltage of the form

$$E = E_0 e^{+j\beta z}$$

is used to describe a voltage wave traveling down a waveguide. Attenuation has been neglected and β is the longitudinal phase

* Received by the IRE, January 30, 1959.
¹ H. J. Riblet, "The short-slot hybrid junction," Proc. IRE, vol. 40, pp. 180-184; February, 1952.

constant. We believe that it is more consistent with common usage to consider a voltage of the form

$$E = E_0 e^{-j\beta z}$$

for the wave traveling down the guide. The difference between the plus and minus signs in the equation gives rise to the error in the previously accepted phase relationship.

B. McCABE
E. J. SCHWESINGER
M. J. TRAUBE
Bell Telephone Labs., Inc.
Whippany, N. J.

Measurement of Minority Carrier Diffusion Length and Lifetime by Means of the Photovoltaic Effect*

Several authors¹⁻⁴ have noted that the photovoltaic effect may be used to measure the minority carrier lifetime near a $p-n$ junction. Gremmelmaier² has considered the relation of the short circuit current to the minority carrier diffusion distance when the ionizing radiation, such as gamma rays, is uniformly absorbed in the sample.

If light is filtered through a sufficient thickness of the same semiconductor as is being tested, the resultant light may be used to generate carriers almost uniformly near a $p-n$ junction. If an independent measurement of the minority carrier diffusion distance is available, the charge generation rate caused by a given filtered light source may be determined, and the value may be used in all subsequent measurements, even on samples which may have a lifetime so low that conventional measurements are difficult.

The writer has found that a simple apparatus, based on the photo-voltaic effect, may be used to determine minority carrier diffusion distances near silicon $p-n$ junctions by measurement of the short circuit photo current.

The light source in this apparatus consists of a 30 watt microscope light with a 40 mil thick silicon filter with lapped surfaces. The illuminated junction area is restricted by a 3/32 inch diameter aperture. The short circuit current is measured by a galvanometer (with 21 ohms internal impedance, 0.55

* Received by the IRE, January 8, 1959.
¹ J. J. Loferski and P. Rappaport, "Electron voltaic study of electron bombardment damage and its thresholds in Ge and Si," Phys. Rev., vol. 98, p. 1869; June, 1955.
² J. J. Loferski and P. Rappaport, "Electron irradiation induced lifetime changes in Ge as a function of conductivity," Bull. Amer. Phys. Soc., vol. 3, p. 141; March 27, 1958 (Abstract).
³ R. Gremmelmaier, "Irradiation of $p-n$ junctions with gamma rays: a method for measuring diffusion lengths," Proc. IRE, vol. 46, pp. 1045-1049; June, 1958.
⁴ L. S. Smirnov (Amer. Inst. Phys. translation), "Measurement of short lifetimes of charge carriers in germanium," Sov. Phys. Tech. Phys., vol. 2, pp. 2299-2301; November, 1957.

μa full scale deflection) with appropriate series and parallel resistances to provide galvanometer damping and sensitivity control.

Gallium or boron was diffused into 10-40 ohm-cm etched silicon wafers, at least 40 mils thick, to a junction depth of about 0.3 and 0.04 mil respectively. The short circuit photo current was measured as was the lifetime in the same samples, after removing the junction, by means of a photo-conductive decay measurement using a spark light source. In order to measure the photo current, contacts were made to the n and p regions with gallium. The area immediately under the contact to the high resistivity n region was sandblasted after diffusion. The excess boron was removed after the boron diffusion.

Using some of the spark lifetime measurements as calibrations for the photo-voltaic apparatus, the results are shown in Table I.

TABLE I

Short Circuit Current I_{sc} , μa	Spark Lifetime** μs	Lifetime Calculated from I_{sc} , μs	Diffusion Length, L_p , Calculated from I_{sc} , mils
1.09*	3.9	3.1	2.5
1.00*	3.6	3.4	2.3
1.05	3.9	3.7	2.4
3.72	36	48	8.6
3.48	34	42	8.0
3.10	21	34	7.1

* Calibration samples.

** Corrected for surface recombination.

The current arising from carrier generation in the diffused p layer and in the space charge region was neglected; the hole mobility was assumed equal to 394 cm²/volt-sec-ond in all cases.

The short circuit current is great enough with the arrangement used to obtain measurements on samples with lifetimes as low as 0.1 μsec. By using a thinner filter and more sensitive meter, the sensitivity could easily be increased.

The approximate uniformity of the carrier excitation was established by measuring the short circuit current in samples about 40 mils thick, first with the incident light exciting carriers through the thin diffused layer as in the above measurements, and second by shining the light through the 40 mil thick n region. From these measurements it was determined that the density of excited carriers decreased about 1 per cent per mil of sample thickness.

In order that the true short circuit current be measured, some restrictions must be placed upon the total resistance in the current measuring circuit.

Using the results derived by Cummerow,⁵ the photo current, for a semi-infinite sample with a thin p-type surface, may be written as

$$I = I_0 A_2 [\exp(eV/kT) - 1] - e g_0 A_1 L_p,$$

or if $eV/kT < 1$,

$$I \cong I_0 A_2 \frac{eV}{kT} - e g_0 A_1 L$$

where

- I = photo current
- I_0 = saturation current
- A_2 = total junction area
- A_1 = illuminated junction area

⁵ R. I. Cummerow, "Photovoltaic effect in p-n junctions," *Phys. Rev.*, vol. 95, pp. 16-21; July, 1954.

g_0 = charge generation rate per unit volume

V = applied junction voltage.

The first term should be small compared with the total current. Therefore, if the total external resistance (including meter resistance, and bulk contact resistance in the structure) is R , one needs

$$R \ll \frac{kT}{e} \frac{1}{I_0 A_2}.$$

Thus, for example, if $A_2 = 1$ cm², $I_0 = 10^{-6}$ amp/cm², it is required that $R \ll 25,000$ ohms. $I_0 = 10^{-6}$ amp/cm² is a very high value for the diffusion component of saturation current. Assuming a high value is an approximate method of including the space charge generated contribution to the current in this calculation.

A second restriction is that the external resistance be much less than the surface leakage resistance at the junction.

For silicon junctions, these requirements are easily met. It has been found desirable, however, to switch into the measuring circuit an additional resistance of about one thousand ohms in each measurement to check on the surface leakage.⁶ If the surface leakage is large, the introduction of the extra resistor will cause a significant drop in the measured current.

The writer wishes to thank R. Hysell for the spark lifetime data, L. Sivo for the preparation of the diffused junctions, and J. F. Elliott for his contributions to the conception of the experiment.

M. WALDNER
General Electric Co.
Syracuse, N. Y.

⁶ This suggestion was made and verified by T. R. Selig, private communication.

Contributors

William S. Ament (SM'55), for a photograph and biography, please see page 162 of the January, 1953 issue of PROCEEDINGS.



Leon A. Ames was born in East Waterford, Me., on December 13, 1907. He graduated from the Lowell Institute School, Cambridge, Mass., in 1931. After fourteen years in the power metering field associated with the Malden Electric Company, he joined the staff of the Radiation Laboratory, M.I.T., Cambridge, Mass., in 1943, where he was engaged in radar system research.

He joined the staff of the Air Force Cambridge Research Laboratory, Cambridge, in November, 1945. As a member of

the Communications and Relay Laboratory, he has done research in microwave relay, data transmission, and storage systems. On special assignment to the Lincoln Laboratory of M.I.T. Lexington, Mass., he did research on digital coding and storage. Following this, he returned to the Cambridge Research Laboratory to study tropospheric and ionospheric propagation.



L. A. AMES

Mr. Ames is chief of the field evaluation section of the Communication Sciences Laboratory.

Sterling R. Anderson (A'41-M'55) was born in Toledo, Ohio, on December 6, 1916. He received the degrees of B.S. in electrical engineering and B.S. in engineering mathematics from the University of Michigan, Ann Arbor, in 1940.



S. R. ANDERSON

Employed by the consulting radio engineer, Paul F. Godley, at Upper Montclair, N. J., from 1940 to 1942, he designed and adjusted directive antenna arrays for broadcasting stations.

Since 1942, he has been working on an-

tennas and associated equipment for air navigation radio aids at the United States Civil Aeronautics Administration Technical Development Center, Indianapolis, Ind. At the present time, Mr. Anderson is serving as Chief of the Transmitter and Antenna Branch of the Center.



Paul T. Astholz (M'45-SM'58) was born in Dayton, Ohio, on August 27, 1916. He received the B.S. degree in electrical engineering from Ohio State University, Columbus, in 1938 and joined a limited staff at the Aircraft Radio Laboratory, Wright Field, Dayton, Ohio, then a signal corps organization engaged in the research and development of navigation aids and communications for aircraft use. He advanced with this organization, which was transferred to the Air Force during World War II, and was placed in charge of development programs in automatic direction finders, instrument landing systems, airborne radars, and other navigational aids. Because of his interest in air traffic control, he transferred to the Technical Development Center of the Civil Aeronautics Administration at Indianapolis, Ind., in 1947. Here, he supervised development programs in ground radars, displays, and distance-measuring equipments. During the past eight years, he has directed the work of the Center on the development and evaluation of air traffic control procedures and equipments and systems, and introduced the use of simulation techniques as a study and evaluation tool in the CAA.



P. T. ASTHOLZ

He has played an active part in planning the research and development programs in air traffic control through his work with the various organizations charged with responsibilities in this area.

Mr. Astholz is a member of the Institute of Navigation.



David T. Barry (S'51-A'52-M'58) was born in Providence, R. I., on September 13, 1928. He received the B.S. degree in engineering from Brown University, Providence, R. I., in 1952 and did post-graduate studies at Pennsylvania State University at University Park and Brown University until 1957.



D. T. BARRY

In 1952 and 1953, he was a research assistant at the Ordnance Research Laboratory, Pennsylvania State. He was a member of the consulting staff of J. I. Thompson Co., Bellefonte, Pa. and an electronic engineer at Federal Products Corp., Providence, R. I.,

from 1953-1954. Since March 1954, he has been employed at the U. S. Naval Underwater Ordnance Station, Newport, R. I., where he has been in turn electronic engineer (instrumentation), Acting Head of the Field Operations Branch in Key West, Florida, project engineer on two underwater weapons. Mr. Barry currently holds a staff assignment as a member of the Advanced Studies Group.



Eugene H. Beach (M'46) was born in Highland, Mich., on October 9, 1918. He attended the University of Michigan, Ann Arbor, receiving the B.S. degree in 1941 in electrical engineering, the M.S. degree in 1947, in physics, and the Ph.D. degree in 1952, in nuclear physics.



E. H. BEACH

Except for periods when he was doing graduate work, he has been at the Naval Ordnance Laboratory in Silver Spring, Md., since 1941. He first worked on electrical systems used in mines and other underwater ordnance. During 1945, he was on active duty as a naval officer assigned to NOL. In his present position as Project Manager for Mine Firing Mechanisms, he is responsible for the design and development of all types of mine fuzing systems.

Dr. Beach is a member of the American Physical Society, Tau Beta Pi, Sigma Xi, Phi Kappa Phi, and Eta Kappa Nu.



Bradford R. Bean (A'56) was born in Massachusetts on September 14, 1927. He received the B.S. degree in meteorology from the University of New Hampshire, Durham, in 1949.



B. R. BEAN

Mr. Bean has worked for the U. S. Weather Bureau at the Durham, N. H. weather observatory from 1945 to 1949, and in the Weather Bureau's Central Analysis Unit in Washington, D. C. during 1949-1950.

Mr. Bean joined the staff of the Central Radio Propagation Laboratory in 1950. Since that time he has been engaged in research on the effect of the atmosphere upon the propagation of radio waves. From 1952 to the present, he has been leader of the radio-meteorology project and has represented the CRPL in several international meetings including the Symposium on Radio Wave Propagation held in Paris, France, September, 1956, the 12th General Assembly of URSI held in Boulder, Colo., August, 1957, and the International Radio Consultative Committee study group conference in

Geneva, Switzerland, in August, 1958. He was made Chief of the Radio Meteorology Section of the National Bureau of Standards Boulder Laboratory in 1958.

Mr. Bean is a professional member of the American Meteorological Society and is a member of the Scientific Research Society of America and the U.S.A. and Commission II of the International Scientific Union.



John S. Belrose (S'49-A'52-M'57) was born in Warner, Alberta, Can., on November 24, 1926. He attended the University of British Columbia, Vancouver, Can. and received the B.A.S. degree in electrical engineering in 1950, the M.A.S. degree in 1952, and the Ph.D. degree in 1958.



J. S. BELROSE

In 1951 he joined the Defence Research Telecommunications Establishment, Ottawa, Ontario, Can., where he worked for two years in the field of low frequency communications engineering. From 1953 to 1956, he studied at the University of Cambridge, England. His research studies, for which he received the Ph.D. degree, were concerned with the oblique reflection of low frequency radio waves from the ionosphere. He returned to the Defence Research Telecommunications Establishment in 1957, and since that time has continued his studies relating to the lowest ionosphere, particularly the disturbed ionosphere. He has been recently appointed Section Leader of the Ionospheric Physics Section.



Robert J. Bibby was born on April 6, 1928, in Shellbrook, Saskatchewan, Can. He was graduated from the University of Saskatchewan, Saskatoon, with the degree of B.S. in engineering physics in 1951.



R. J. BIBBY

During the years 1951 to 1957, he was with the Research Division of Ferranti Electric Ltd., Toronto, Can., where his work was concerned with circuit design and with data handling systems and electronic computers.

He joined the Electronics Laboratory of the Defence Research Board in 1957 where his work has been in transistor circuitry.

Mr. Bibby is a member of the Ontario Association of Professional Engineers.



Miles H. Bickelhaupt (M'55) was born on March 9, 1922 in Utica, N. Y. He received the B.S. degree in electrical engineer-

ing in 1948 at Cornell University, Ithaca, N. Y.

He was associated with private industry until February, 1951 when he joined the staff of Rome Air Development Center, Griffiss Air Force Base, N. Y., in the field of communications, and where he presently serves as chief of the transmission branch in the Directorate of Communications.

Mr. Bickelhaupt is an associate member of the AIEE.



M. BICKELHAUPT

❖

James R. Brown, Jr. was born in Philadelphia, Pa., on September 18, 1926. He received the B.S. degree in 1951 from Pennsylvania Military College, Chester, Pa.

He is now an electronic scientist in the transducer branch of the U. S. Naval Air Development Center, Johnsville, Pa., and a consultant to Ferro-electrics Associated, Oreland, Pa. He has received the Distinguished Service Award of the Hatboro Junior Chamber of Commerce, Pa., in 1956. He has applied for two patents as co-inventor.



J. R. BROWN, JR.

❖

R. Keith Brown (M'59) was born on July 19, 1913 in Canada. He received the Bachelor's degree in physics and mathematics from the University of British Columbia, Vancouver, B. C., Can., in 1941 and the Master's degree in 1948. From 1941 until 1948, he was a lecturer in the Department of Physics at the University of B. C. During this time he instructed RCAF radar mechanics recruits in the principles of radio and electronics (1941 to 1943) and later lectured in undergraduate physics (1943 to 1948).

In 1948 he joined the staff of B. C. Research Council as associate research physicist, where his work involved research and development directed towards industrial problems in the Province of British Columbia.

In 1951, Mr. Brown came to Ottawa and joined the staff of the Defence Research Board. From 1952 to 1958, Mr. Brown was a member of, and later leader of, the team which developed the Canadian lightweight Doppler navigation radar.

In 1951, Mr. Brown came to Ottawa and joined the staff of the Defence Research Board. From 1952 to 1958, Mr. Brown was a member of, and later leader of, the team which developed the Canadian lightweight Doppler navigation radar.

In 1951, Mr. Brown came to Ottawa and joined the staff of the Defence Research Board. From 1952 to 1958, Mr. Brown was a member of, and later leader of, the team which developed the Canadian lightweight Doppler navigation radar.



R. K. BROWN

William F. Cahill was born in Washington, D. C., on August 8, 1924. He received the B.E.E. degree in 1949, and the M.S. degree in mathematics in 1951, both from Catholic University, Washington, D. C.



W. F. CAHILL

He was a lecturer in mathematics at Catholic University from 1947 to 1951, and an associate in mathematics at George Washington University, Washington, D. C., from 1949 to 1953. He also worked as a mathematician at the National Bureau of Standards from 1951 to 1958. He is presently chief of the Mathematics Branch, Theoretical Division, of the National Aeronautics and Space Administration.

Mr. Cahill is a member of Sigma Xi and of the Association for Computing Machinery.

❖

Edward R. Carlson was born in Pittsburgh, Pa., in 1921. He received the Bachelor of Science degree in electrical engineering from the University of Pittsburgh, Pittsburgh, Pa. in 1948. He then joined the National Aeronautics and Space Administration, Cleveland, Ohio, where he has since been concerned with instrumentation of research propulsion test equipment and associated data systems.



E. R. CARLSON

❖

Charles J. Casselman was born in Sherman, N. Y., on February 15, 1918. He attended the Radio Materiel School of the U. S. Naval Research Laboratory, Washington, D. C., and the Extension Division of the University of California at San Diego.



C. J. CASSELMAN

From 1941 to 1945, he served with the U. S. Navy. For 1½ years, he was assigned to the staff of instructors in the Navy Radio Materiel School program at NRL, Utah State Agricultural College at Logan, and College of the Ozarks, Clarksville, Ark. From 1946 to 1947, he was a member of the Airborne Radio Division, NRL. Since 1947, he has been with the U. S. Navy Electronics Laboratory, San Diego, Calif., where he has been concerned with LF communications, automatic weather station development, and

From 1941 to 1945, he served with the U. S. Navy. For 1½ years, he was assigned to the staff of instructors in the Navy Radio Materiel School program at NRL, Utah State Agricultural College at Logan, and College of the Ozarks, Clarksville, Ark. From 1946 to 1947, he was a member of the Airborne Radio Division, NRL. Since 1947, he has been with the U. S. Navy Electronics Laboratory, San Diego, Calif., where he has been concerned with LF communications, automatic weather station development, and

VLF long-range navigation system development. At present, Mr. Casselman is an electronic scientist, head of the fixed station design group, navigation systems section.

❖

Channing Conger was born on November 9, 1924, in Caldwell, Idaho. He received the Bachelor of Science degree in electrical engineering from Iowa State College at Ames in 1948 and immediately joined what is now the National Aeronautics and Space Administration, Cleveland, Ohio. Mr. Conger is now engaged in research and design of instrumentation and data systems.



C. CONGER

❖

William A. Cumming (A'47-M'50-SM'57) was born in Detroit, Mich. on July 16, 1926. After receiving the B.S. degree from Queen's University, Kingston, Ont., Can., in 1947, he joined the staff of the Radio and Electrical Engineering Division of the National Research Council of Canada. Since that time he has been a member of the microwave section, working in the field of antenna design and development.



W. A. CUMMING

design and development.

Mr. Cumming is a member of the Professional Institute of the Public Service of Canada, and a member of the Association of Professional Engineers of the Province of Ontario.

❖

Stanley F. Danko (M'46) was born on January 6, 1916 in New York, N. Y. He received the B.S. degree in electrical engineering from Cooper Union in 1937, with graduate studies in UHF techniques at the Moore School of Electrical Engineering, University of Pennsylvania, and at Rutgers University. He was employed in the U. S. Army Signal Corps Inspection Laboratory in Philadelphia during 1940-1946, joining the U. S. Army Signal Research and Development Laboratory at Fort Monmouth, N. J. in 1946.



S. F. DANKO

He has been associated with the Labora-

He has been associated with the Labora-

tory's electronic component development, specializing in miniaturization techniques, including printed circuits and packaging of electronic assemblies. He has been primarily responsible for the establishment of the dip-soldered printed wiring process (auto-sembly) in use today, and has authored a number of papers on the subject. In 1957 he received a Department of the Army Award for his miniaturization contributions.

He is currently chief of the technical staff of the Electronic Components Research Department of the U. S. Army Signal Research and Development Laboratory.

Mr. Danko is a registered professional engineer in the state of Pennsylvania.

❖

Robert M. Davis, Jr. was born in Washington, D. C., on June 25, 1918. He worked as a mathematician with the U. S.



R. M. DAVIS, JR.

Coast and Geodetic Survey from 1943 to 1946, making studies in geomagnetism, including isomagnetic charts. He studied at the University of California, Berkeley, Calif., where he was a member of Phi Beta Kappa, and received the B.A. degree in mathematics from George Washington

University, Washington, D. C., in 1948. He also participated in the forecasting of magnetic and radio disturbances in the Sun-Earth relationships group at the Central Radio Propagation Laboratory, National Bureau of Standards.

He is studying at the University of Colorado, Boulder, Colo., in addition to his work at the NBS, where he is a physicist in the Ionospheric Communications Systems group of the CRPL. He has been concerned with the characteristics of radio propagation at very high frequencies and has made studies of several sources of interference to VHF transmission, such as F2 layer interference and Es interference.

Mr. Davis is a member of the American Geophysical Union and an associate member of the Research Society of America.

❖

Norman J. Doctor (M'58) was born in Brooklyn, N. Y., on October 5, 1929. He attended Purdue University, Lafayette, Ind., receiving the B.S. degree in physics in 1951.



N. J. DOCTOR

He has been working since 1955 in the fields

of printed circuits and electronic packaging and is now a research supervisor in the materials branch.

Mr. Doctor was a member of the DOFL group which won the 1957 Miniaturization Award presented by Miniature Precision Bearings, Incorporated, and is a member of Sigma Pi Sigma.

❖

L. W. Doremus was born in Essex Falls, N. J., on March 24, 1927. He received the B.S. degree in physics from the University of Miami, Miami, Fla., and plans to complete the requirements for the M.S. degree in electrical engineering this Spring at Stevens Institute of Technol-



L. W. DOREMUS

ogy, Hoboken, N. J. He is a project physicist in the Research Section at Picatinny Arsenal where he has been employed for the last six years. His duties at the Arsenal have been primarily concerned with the study of the physical and electrical characteristics of components, power sources, and energy conversion devices for special ammunition devices.

❖

Willie L. Doxey (M'53) was born in Montgomery, La., on February 21, 1912. He received the B.S. degree in physics and mathematics in 1934 from Northwestern State College of Louisiana at Natchitoches, and the M.S. degree in 1939 from Louisiana State University, Baton Rouge. He was employed in the Office of the Chief Signal Officer, Washington, D. C. during 1942-1943, directing research and development operations on frequency control devices and related equipments.



W. L. DOXEY

He was commissioned in the Signal Corps in 1942, attaining the rank of major in the Air Force in 1946. He entered the U. S. Army Signal Research and Development Laboratory, Fort Monmouth, N. J., in 1946 as a member of the technical staff of the Frequency Control Branch. His subsequent assignments were as chief of the Frequency Control Branch, and chief of the Power Sources Branch. He is currently deputy director of the Electronic Components Research Department.

Mr. Doxey is a member of the Armed Forces Management Association and the Armed Forces Communications Electronics Association.

❖

John H. Dunn was born in Rochester, N. Y., on July 17, 1921. He received the B.S. degree in electrical engineering at the

University of Notre Dame, Notre Dame, Ind., in 1943. Since 1943 he has been employed by the Naval Research Laboratory,



J. H. DUNN

Washington, D. C. From 1943 to 1947 he was engaged in research and development involving airborne equipment. Since 1947, Mr. Dunn has been engaged in research and development work in the field of precision automatic-tracking radar. In 1956 he was appointed head of the Tracking Branch of the Radar Division at NRL.

❖

Maynard Eicher was born in Washington, D. C., on September 30, 1916. He attended American University in Washington, D. C., receiving the B.A. degree in physics in 1937. In 1938 and 1939 he did graduate study at George Washington University, Washington, D. C.



M. EICHER

He was first employed at the National Bureau of Standards, and from 1938 to 1942 worked at the National Institutes of Health, Bethesda, Md., as a physicist. During the following two years, he was at the Naval Ordnance Laboratory doing experimental development of torpedo-firing mechanisms.

Since 1944, Mr. Eicher has been at the Naval Medical Research Institute, Bethesda, Md. Until 1947 he served there as an officer in the Navy and is now employed as a civilian electronic scientist. He is head of the Instrumentation Laboratory, being responsible for the design and development of electronic instruments for use in medical research.

❖

Clifton D. Ellyett was born in Christchurch, N. Z., on January 10, 1915. He received the M.S. degree from Canterbury University College, Christchurch, in 1937.



C. D. ELLYETT

He has been on the staff of Canterbury University College since 1938, and has served as senior lecturer in physics there since 1952. From 1942 to 1946, Dr. Ellyett supervised the processing of ionospheric data from Christchurch and from the Campbell, Raratonga, Kermadec, Fiji, and Pitcairn Islands for the Department of Scientific and Industrial Research in New Zealand. At the end of World War II, he undertook three years of research on the radar detec-

tion of meteors at Manchester University, Eng., and received the Ph.D. degree there in 1949. On returning to New Zealand in 1950, he established and has since directed a radio and radar field station near Christchurch, where his work has been associated mainly with radar-meteor studies.

He has represented the New Zealand Government at several important scientific meetings. In 1943, he spent three months in Australia coordinating New Zealand and Australian ionospheric work. In 1944, he was sent by the New Zealand Chiefs of Staff to Washington, D. C., to attend a Radio Propagation Conference convened by the Combined Communications Board. He was the sole representative from New Zealand at the Eighth General Assembly of the International Scientific Radio Union at Stockholm, Sweden, in 1948. In 1953, Dr. Ellyett was the New Zealand delegate to the Tenth General Assembly of the same group at Sydney, Australia. He served as a consultant to the Radio Propagation Physics Laboratory, National Bureau of Standards, Boulder, Colo., from 1957 to 1958, while on leave from the University of Canterbury.

Dr. Ellyett is a member of the New Zealand Radio Research Committee, the New Zealand Defense Science Advisory Committee, and the Electrotechnology Panel.



Robert B. Flint (S'50-A'53-M'58) was born January 7, 1925, at Eldon, Iowa. In 1952, he received the B.S. degree in electrical engineering from the Iowa State College at Ames.



R. B. FLINT

Since 1952, Mr. Flint has been employed by the Civil Aeronautics Administration, Technical Development Center, Indianapolis, Indiana, in development work and flight tests of electronic aids for air navigation.



John M. Formwalt was born on May 24, 1915, in Florence, Ala. He received the B.A. degree in mathematics in 1937 from Carson-Newman College, Jefferson City, Tenn., and the M.A. degree in physics in 1939 from Duke University, Durham, N. C.



J. M. FORMWALT

In 1939 and 1940 he taught physics at Presbyterian Junior College, Maxton, N. C. From 1940 to 1944, he was employed as a physicist by the Navy Department, Bureau of Ordnance, and the Eleventh Naval District in connection with the protection of ships from magnetic mines. In 1944, Mr. Formwalt was transferred to the U. S. Naval Torpedo Station, Newport, R. I., which later became the U. S. Naval Underwater Ordnance Station.

For a period of six months in 1946, he joined the Staff, Commander Joint Task Force One, on a loan basis as a member of the Bureau of Ships Instrumentation Group to assist in the determination of the effects of atomic bombs on Naval vessels. At the Underwater Ordnance Station he has been in turn physicist, project engineer, Head of the Instrument Section, Head of the Instrumentation Branch, Chief of the Test Division, Chief of the Controls Division, Chief of the Advanced Undersea Weapons Division. He holds a staff assignment as Head of the Advanced Studies Group.

Mr. Formwalt is a member of the American Physical Society.



Colin A. Franklin (S'54-M'57) was born in Hastings, N. Z., on December 9, 1927. He received the B.S. degree in physics in 1951



C. A. FRANKLIN

from Auckland University College, N. Z., and lectured in the Physics Department at Auckland, N. Z., in 1952. In 1953, he received the M.S. in physics at Auckland and in May, 1953 joined the New Zealand Scientific Defence Corps. In November, 1953, he registered in the Electrical Engineering Department of the Imperial College of Science and Technology, University of London, Eng., where he received the Ph.D. in 1957 for work concerning the application of transistors to pulse communications systems.

In August, 1957, he joined the Defence Research Board of Canada and since then has been employed in the Electronics Laboratory of the Defence Research Telecommunications Establishment at Ottawa on the design of a transistor Doppler navigator.



John H. Ganton was born in Toronto, Ont., Can., on October 13, 1934. He received the B.A.S. degree in engineering physics from the University of Toronto in 1956.



J. H. GANTON

After graduation he joined the Defence Research Board of Canada, Electronics Laboratory, and worked for two years on the design of airborne navigational computation. He is now working on the design of data handling equipment.

Mr. Ganton joined the RCAF reserve in 1952 and received his pilot's wings in 1955.



Charles Don Geilker was born on December 15, 1933, in Kingston, Mo. He received the A.B. degree in physics from William

Jewell College, Liberty, Mo., in 1955. An AEC Fellow in Radiological Physics at Vanderbilt University, Nashville, Tenn., and Oak Ridge National Laboratory during 1955-1956, he was awarded the M.A. degree in physics from Vanderbilt in 1957. For the past two years he has been a lecturer on the Training Staff, the Division of Radiological Health, Robert A. Taft Sanitary Engineering Center, U. S. Public Health Service, Cincinnati, Ohio. His work has involved particularly the development and modification of instrumentation for applied radiological physics.

Mr. Geilker is a member of the American Association of Physics Teachers, the American Association for the Advancement of Science, Sigma Xi, and the Health Physics Society.



C. D. GEILKER



Benjamin Glatt (A'51) was born in Alhambra, Calif. on January 27, 1926. He attended the University of California at Berkeley, where he received the B.S. degree in electrical engineering in 1950.



B. GLATT

After graduation, Mr. Glatt joined the staff of the U. S. Naval Ordnance Test Station in China Lake, Calif. and has worked there till the present time. As a member of the Aviation Ordnance and Test Departments, he has done work in the field of instrumentation for the Station's test ranges.



Francis P. Hanes (A'54) was born in Vicksburg, Miss., on January 16, 1920. He attended Mississippi State College, following



F. P. HANES

which he studied advanced radio engineering. He joined the Instrumentation Branch, U. S. Army Engineer Waterways Experiment Station, Vicksburg, Miss., in 1941. He was in charge of instrumentation on the sea-level canal study in the Canal Zone in 1946, returning to the USAEWES in 1947. During the period 1953-56 he conducted an R&D program on Hydrologic Data Telemetry in the Louisville District, Ky., Corps of Engineers. During this period he invented the Hanes precipitation telemetering device.

In 1956 Mr. Hanes rejoined the Instrumentation Branch, USAEWES, and is Chief

of the Dynamics Section wherein he is concerned with instrumentation-system design and analysis. He also serves as consultant for the Corps of Engineers on hydrologic data telemetry.

❖

Isadore Harris was born on April 4, 1927. He received the B.S. degree from Roosevelt College, Chicago, Ill., in 1951 and the Ph.D. degree in physics from Northwestern University, Evanston, Ill., in 1956.



I. HARRIS

He worked for the Naval Research Laboratory in Washington, D. C., from 1956 to 1958, and is presently a physicist with the Theoretical Division of the National Aeronautics and Space Administration, Washington, D. C.

Dr. Harris is a member of the American Physical Society.

❖

W. Lewis Hatton (S'49-A'51-M'57) was born at Trap Lake, British Columbia, Can., on January 22, 1923. He joined the Royal Canadian Air Force in 1941, and received his commission the same year.



W. L. HATTON

From 1941 to 1945, he was attached to the Royal Air Force serving as radar officer in England and the Middle East. From 1945 to 1951, he attended the University of British Columbia, Vancouver, Can., and received the B.A.S. degree in electrical engineering in 1950 and the M.A.S. degree in 1951. He then joined the Defence Research Telecommunications Establishment in Ottawa where he carried out research on low-frequency atmospheric noise and modulation methods.

From 1953 to 1955, Mr. Hatton was attached as a lecturer to the Royal Military College of Science, Shrivenham, Berkshire, Eng. In 1955, he re-joined the Defence Research Telecommunications Establishment. Mr. Hatton is now head of the communications research section of the communications wing of DRTE.

Mr. Hatton is a junior member of the Engineering Institute of Canada and a member of the Association of Professional Engineers of the Province of Ontario.

❖

Donald P. Heritage (A'46-M'55) was born in Vineland, N. J., on September 4, 1917. He received the B.S. degree in electrical engineering from Tri-State College, Angola, Ind., in 1940.

During 1940 and 1941, he was associated with Philco Corporation and RCA Manufacturing Company as a test engineer. From

1942 to 1943, he was employed by the Signal Corps Laboratories, Belmar, N. J., as a radio engineer in development and evaluation of radar components.



D. P. HERITAGE

In World War II, Mr. Heritage was a radar officer in the U. S. Navy, his last duty assignment being at the Naval Research Laboratory, Washington, D. C. From 1946 to 1947 he worked as a radio engineer at the Naval Research Laboratory in development of airborne communication systems and frequency measuring equipment. In 1947, he joined the U. S. Navy Electronics Laboratory, San Diego, Calif., where he has been concerned with the development of radio communications and navigation systems. Since July, 1951, he has been head of the radio branch at NEL.

❖

Adolph W. Herzog (M'56) was born in New York, N. Y., on February 23, 1925. He received the B.S. degree, after majoring in physics, from Queens College, Flushing, N. Y., in 1949. For four years he was employed, in the capacity of Customer's Engineer, by International Business Machines.



A. W. HERZOG

Mr. Herzog has been associated with Operational Flight Trainers at the U. S. Naval Training Device Center, Port Washington, N. Y., since 1953, and is presently employed there as an electronic scientist and as a member of the Computer Research Branch of the Research Department

❖

B. M. Horton (A'52-M'57) was born in Bartlett, Tex., on December 27, 1918. He received the B.A. degree at the University of Texas, Austin, in 1941 and the M.S. degree from the University of Maryland at College Park in 1949, both in physics. From 1942 to 1946, he served as radar officer in the U. S. Army Signal Corps in the U.S.A. and in Europe. In 1946, he joined the Naval Research Laboratory



B. M. HORTON

where he worked on high-altitude sliding contacts, low-noise sliding contacts, and detectors for high-intensity, high-energy radiation for use in some of the Eniwetok tests. In 1951, he joined the Ordnance Electronics Division of the National Bureau of Standards, which is now the Diamond Ord-

nance Fuze Laboratories, where he has been engaged in research and development in ranging and detecting systems. He is presently chief of the systems research laboratory at DOFL. He is also a member of the Army Physics Advisory Group and is a part-time instructor for the University of Maryland in the National Bureau of Standards Graduate School, Washington, D. C.

Mr. Horton is a member of the Philosophical Society of Washington.

❖

Dean D. Howard (S'47-A'50-M'55-SM'57) was born in Chatham, N. J., on January 2, 1927. He received the B.S. degree in electrical engineering from Purdue University, Lafayette, Ind., in 1949 and the M.S. degree in electrical engineering from the University of Maryland at College Park in 1952.



D. D. HOWARD

He was employed in the Radio Division, Communication Branch, of the Naval Research Laboratory, Washington, D. C., from 1949 to 1954 where he worked in the field of UHF communication systems and techniques.

Since 1954 he has been working with the Radar Division of NRL in the Tracking Branch engaged in basic research on target-noise theory. He is presently section head of the Special Systems I and the Radar Statistics Sections of the Tracking Branch.

Mr. Howard is a member of the NRL branch of the Scientific Research Society of America.

❖

John M. Ide (A'47-SM'50) was born in Mount Vernon, N. Y., on August 17, 1907. He received the B.A. degree from Pomona College in Claremont, Calif., in 1927 and the M.S. and D.Sc. degrees from Harvard University, Cambridge, Mass., in 1929 and 1931 respectively.



J. M. IDE

The Harvard degrees were in physics and communications engineering. He served as instructor in physics and communication engineering at Harvard from 1929 to 1936. While there he completed researches on magnetostriction devices and measurements of the velocity of sound in rocks and glasses. Results of this work were published in numerous scientific papers.

In 1936, he became staff geophysicist for the Shell Oil Company of Houston, Tex. He developed improvements to the design of gravity meters and other instrumentation for use in geophysical exploration. From 1941 to 1945, he served as Section Head in the Sound Division of the Naval Research

Laboratory, Washington, D. C. Here he was responsible for research and development on underwater loudspeakers, ultrasonic echoing equipments, low-frequency acoustic minesweeping devices, and underwater telephone and homing equipments for the use of small craft and commando swimmers.

In 1945 he became Technical Director of the U. S. Navy Underwater Sound Laboratory, New London, Conn. In this position he has been responsible for the past fourteen years for research and development on problems of undersea warfare. A recognized authority in the naval sciences, concentrating on subsurface warfare, Dr. Ide is in charge of a scientific staff of about 250 well-qualified scientists and engineers. In May, 1958, he was given an award by the National Civil Service League, one of ten such awards given annually within the federal service. He was also cited for "outstanding achievements in research, development, and organization" and "aiding materially in improving sonar and other underwater devices which are now essential to the successful operation of atomic powered submarines."

Dr. Ide is a member of Phi Beta Kappa, Tau Beta Pi, and the American Institute of Electrical Engineers, and a Fellow of the Acoustical Society of America. He is a member of the Board of Trustees of Mitchell College in New London, Conn.

❖

Robert Jastrow was born in New York, N. Y., on September 7, 1925. He received the B.A. degree from Columbia College, New York, in 1944, and the M.A. and Ph.D. degrees in 1945 and 1948, respectively, both from Columbia University, New York, N. Y.



R. JASTROW

He was a lecturer in physics at Columbia University from 1944 to 1947, and an instructor in physics at Cooper Union, New York, from 1947 to 1948. He was a Postdoctoral Fellow at Leiden University, Leiden, The Netherlands, from 1948 to 1949. From 1949 to 1950, and in 1953, he was a member of the Institute for Advanced Study at Princeton, N. J. He was a research associate with the University of California from 1950 to 1953, and an assistant professor of physics at Yale University, New Haven, Conn., from 1953 to 1954. He also became a lecturer in physics at the University of Maryland, College Park, in 1954. From 1954 to 1958 he was a consultant with the U. S. Naval Research Laboratory, Washington, D. C., and since 1958, has been a consultant with the Los Alamos Scientific Laboratory, the Convair Division of the General Dynamics Corporation, and a member of the Space Science Board Committee 7 (Ionospheres of Earth and Planets). He is presently chief of the Theoretical Division of the National Aeronautics and Space Administration.

Dr. Jastrow is a member of the American Physical Society, the Physical Society of

Great Britain, the American Geophysical Union, and the American Astronomical Society.

❖

Arno M. King (A'50-M'55) was born in Cleveland, Ohio, on December 12, 1920. He attended Washburn University, Topeka, Kan., in 1939, then transferred to Bucknell University, Lewisburg, Pa., where he received the B.S. degree in 1943. Following graduation he was employed by the Naval Research Laboratory, Washington, D. C., where he has spent all of his professional career working on problems associated with tracking radar. He is currently serving as head of the Terminal Equipment Section of the Tracking Branch, Radar Division.



A. M. KING

Mr. King is a member of Tau Beta Pi, Pi Mu Epsilon, Sigma Pi Sigma, and the Scientific Research Society of America.

❖

Myron W. Klein was born in Rochester, N. Y., on June 17, 1919. He studied at the University of Rochester, from which he received the A.B. degree in physics in 1943.



M. W. KLEIN

Entering the Armed Forces in June, 1943, he took basic training at Fort Riley, Kan., and then studied at the University of Nebraska at Lincoln under the Army Student Training Program. Later, he taught mathematics in the Ordnance Automotive School at Fort Crook, Neb., and then was assigned to the Engineer Board (now the U. S. Army Engineer Research and Development Laboratories, Fort Belvoir, Va.) where he worked on the development of the first Sniperscope and Metascope. Following his release from the Army, in February, 1946, he remained at the Laboratories in a civilian capacity and is presently employed as a Section Chief in the Warfare Vision Branch. He received the Legion of Merit for his work on the development of the Metascope, Type F. Since entering civil service, he has taken courses at Catholic University, Washington, D. C.

Mr. Klein is a member of the Optical Society of America

❖

Robert B. Knowles (S'43-A'45-M'51) was born on April 30, 1921, in Des Moines, Iowa. He received the B.S. degree from Iowa State College, Ames, in 1943, and came to work at the Naval Ordnance Laboratory, Silver Spring, Md., the same year.

During the war, he was on active duty in the Navy and assigned to the Laboratory. His work is on the design and development of electronic and electrical components for underwater weapons and missiles. He is now the chief of the Electrical Systems Materials Division in the Underwater Electrical Engineering Department.

Mr. Knowles is a member of AIEE, the Acoustical Society of America, and Eta Kappa Nu.



R. B. KNOWLES

❖

Harry P. Kramer was born in Chicago, Ill., on March 17, 1909. He received the B.S. degree from St. Mary's College, St. Mary's, Kan., and the M.S. degree from Creighton University, Omaha, Neb.



H. P. KRAMER

He did additional graduate work at the University of Chicago, and in the Graduate School of Civil Engineering at the Illinois Institute of Technology in Chicago.

He was sanitary chemist for the Sanitary District of Chicago and later sanitary engineer of the city of Chicago before joining the staff of the United States Public Health Service's Robert A. Taft Sanitary Engineering Center in Cincinnati, Ohio in 1949. He has served as chief of the training program since 1952, and is also assistant clinical professor of industrial health at the University of Cincinnati.

Mr. Kramer is a member of the Sewage Works Association, the American Chemical Society, the American Public Health Association, and consultant to the American Public Works Training Committee. He is Chairman of the New Standards Methods Subcommittee on Preparation of Standard Samples of the AWWA, and a member of the Task Group on the Effects of Synthetic Detergents in Water Supplies.

❖

Raymond E. Lacy (SM'46) was born on March 17, 1916, in Camden City, N. J. He received the B.S. degree in electrical engineering from Drexel Institute of Technology, Philadelphia, Pa., in 1938, and the M.S. degree in electrical engineering from New York University, N. Y., in 1940. He has intermittently continued graduate study from 1940 to the present at Brooklyn Polytechnic Institute, La



R. E. LACY

Salle College, Columbia University, Har-

vard University, and New York University.

From 1933 to 1938 he worked on engineering jobs at the Potomac Electric Power Co., A. C. Gilbert Co., Philco Radio & Television Co., and Chubbuck & Patrick Consulting Engineers. From 1938 to 1940 he was a member of the electrical engineering faculty of New York University. In 1940 he joined the U. S. Army Signal Engineering Laboratory at Fort Monmouth, N. J., as a radio engineer. Initially he designed many portions of early communication and test equipments, including the first FM transmitters, the first time-division multiplex to operate in the U.S.A., and later directed many large electronic research and development programs which, among other advancements, first applied microwaves and PCM to communications, and evolved the surface-wave-transmission-line principle, and various advancements in propagation, such as obstacle gain. He has contributed numerous publications on such work. He is presently an electronic scientist directing communication research in the U. S. Army Signal Research and Development Laboratory, Fort Monmouth, N. J.

Mr. Lacy is a member of Sigma Xi, Eta Kappa Nu, and Tau Beta Pi, and has received various meritorious civil service awards, and was the recipient of a Harvard Fellowship in 1940.



James C. Laurence was born in Terrell, Tex., on September 17, 1909. He received the A.B. and A.M. degrees in physics at the University of Akron, Ohio, in 1932 and 1933, respectively.



J. C. LAURENCE

After several teaching assignments he joined the National Aeronautics and Space Administration, Lewis Research Center, Cleveland, Ohio in 1944. Since joining this agency he has been engaged in research in aerodynamic turbulence and noise research.

Mr. Laurence is a member of the Acoustical Society of America.



Gilbert Lieberman (A'55) was born in New York, N. Y., on December 30, 1922. He received the B.A. degree in mathematics from New York University in 1948 and the M.A. in mathematical analysis from Columbia University, New York, in 1949.



G. LIEBERMAN

From 1950 to 1955, he was employed at the Naval Research Laboratory in Washington, D.C., performing research and data analysis on

the processing of signals in noise backgrounds. He transferred to the Naval Ordnance Laboratory in Silver Spring, Md., in 1955, and has been engaged since then in applications of statistical methods in signal processing.

At present he is a mathematician in the acoustics division of the physics research department.

Mr. Lieberman is a member of the Institute of Mathematical Statistics and of the American Statistical Association.



Edward J. Martin was born in St. Albans, N. Y., on January 12, 1932. He received the A.B. degree in mathematics from Fordham University, New York, N. Y., in 1954, and is currently completing work toward the M.S. degree in mathematics and physics at Northeastern University, Boston, Mass.



E. J. MARTIN

Upon entry into the U. S. Air Force in December 1954, he was assigned to the Air Force Cambridge Research Center, Bedford, Mass., where he was engaged in statistical studies of tropospheric radio propagation. In December, 1956, he joined the Communication Sciences Laboratory of that Center as a civilian staff member and continued analytical studies of tropospheric and ionospheric radio propagation and communications.

Recently, as chief of the Analysis Section of the Communication Sciences Laboratory, Mr. Martin has been involved in studies of lunar and satellite relay communications.



C. A. McKerrow was born at North Bay, Can., in 1919. He attended the Marconi Radio School in Toronto, Can. During the early part of World War II, he was employed by the Canadian Marconi Company as a marine radio operator. In 1941, he was employed by the department of transport of the Radio Monitoring Division. The period from 1945 to 1946 was spent in the eastern Arctic on the east coast of Baffin Island taking measurements of the ionosphere and other geophysical data.



C. A. MCKERROW

He joined the Radio Propagation Laboratory in 1947, which has now become the Defence Research Telecommunications Establishment, and continued research in ionospheric measurements. Later his attentions were directed to research engineering in low-frequency communications systems. For the

past two years, Mr. McKerrow has been engaged in the measurement of low-frequency atmospheric noise in Canada.



James P. McNaul (M'57) was born in Madison, Wis., on August 20, 1933. He received the B.S. degree in electrical engineering from the University of Wisconsin in 1956. He worked as a field engineer for Square D Company, Milwaukee, Wis., before being commissioned in the Signal Corps in May, 1956.



J. P. McNAUL

Lt. McNaul's first assignment was at the U. S. Army Signal Research and Development Laboratory, Fort Monmouth, N. J., as assistant project officer of Project MONMOUTH, a communications system and frequency compatibility study. He served in this capacity until May, 1958, when he joined the technical staff of the Electronic Components Research Department.

He is a member of the Armed Forces Communications and Electronics Association and Kappa Eta Kappa, and an associate member of the AIEE. He is presently a member of the Administrative Committee of the IRE Professional Group on Radio Frequency Interference and is serving as treasurer.



A. G. McNish was born in Washington, D. C., on January 18, 1903. Just after graduation from high school, during the summer of 1920, he was appointed laboratory apprentice at the National Bureau of Standards. He continued his education by attending evening classes at George Washington University, where he received the A.B. degree in 1924 and the M.S. degree in 1931.



A. G. McNISH

In 1925 he was appointed physics teacher in the District of Columbia high schools. In 1930 he joined the staff of the Carnegie Institution's Department of Terrestrial Magnetism. While with the Institution he conducted original research in terrestrial magnetism, atmospheric electricity, and related branches of geophysics, and published numerous papers and several reference articles.

Following the outbreak of World War II in Europe, he became involved in the defense effort, preparing a proposal for degaussing of ships in November 1939. He helped initiate the Navy's degaussing program and continued to be active in the defense program for the next six years, working on many different projects. For this work he received the Presidential Certificate of Merit, the Office of Scientific Research and Development

Certificate of Appreciation, and the Bureau of Ordnance Certificate of Appreciation.

He returned to the National Bureau of Standards in 1946 as Chief of the Basic Ionospheric Research Section of the Central Radio Propagation Laboratory. In 1949 he was made Assistant Chief of the Laboratory. In 1951 he went on temporary detail for 18 months to initiate and organize a research and development program for land-mine fuzes. When the Laboratory moved from Washington to Boulder, Colo., in 1954, he remained in Washington and was made Consultant to the Director to undertake studies on the physical constants and improvement of accuracy in standards of measurement.



Erwin H. Meyn was born in Erie, Pa., on September 27, 1926. He received the B.S. degree in electrical engineering from the Pennsylvania State University, University Park, in 1950. From 1950 to 1952 he was associated with General Chemical Division, and Hammernill Paper Company in their Plant Engineering Departments. From 1952 to 1954 he worked in the Electrical Design Department of Marathon Electric Corporation.



E. H. MEYN

From 1954 to 1956 he was associated with the Engineering Department of Reliance Electric. In 1956 he joined the National Aeronautics and Space Administration, Lewis Research Center, Cleveland, Ohio, in its Instrument and Computing Research Division, designing and developing instruments for measurement and computation of aeronautical research data.

Mr. Meyn is a professional engineer in the state of Ohio and an associate member of the AIEE.



Jonathan Mitchell (S'52-A'54) was born on December 28, 1929, in Ottawa, Can. He received the B.S. degree in engineering physics from Queen's University, Kingston, Ont., in 1952 and the D.I.C. from the Imperial College of Science and Technology, London, Eng., in 1953.



J. MITCHELL

After studying in London and Manchester for two years on an Athlone Fellowship, Mr. Mitchell went to work for the Defence Research Board of Canada, where he designed transistor circuits for a Doppler radar navigational system. In 1957, Mr. Mitchell went to work for Link Aviation, Inc., in Binghamton, N.Y., where he is supervising the development of transistor circuits and systems for flight simulators.

Norman F. Moody was born on December 22, 1915 in Great Britain. He held senior research positions in the fields of television and radar until the end of the war, when he came to Canada as head of the electronics branch at the Atomic Energy Project, Chalk River, Ont. Afterwards he joined the Ministry of Supply in England where he made measurements on the first U.K. A-bomb explosion. He then returned to Canada to take a position with the Defence Research Board Electronics Laboratory as head of the Transistor Section.



N. F. MOODY

Mr. Moody has published some sixteen papers in these fields and has taken out about thirty-five patents. He is an associate member of the IEE of Great Britain and has won the Kelvin Award for papers on radar presented to that society. He has recently accepted a post as professor of electrical engineering at the University of Saskatchewan, Saskatoon, Can.



John S. Nader was born in Farrell, Pa., on November 16, 1921. He received the B.S.E.E. degree from the University of Cincinnati, Ohio, in 1944. In 1945 he was project engineer for servo control units with Sperti, Inc., Norwood, Ohio, and in 1946 he was development engineer for physical therapy equipment with R. R. Gannon Company, Cincinnati. From 1946 to 1949 he was an instructor in engineering physics and completed part of his graduate work at the University of Cincinnati which awarded him the M.S. degree in physics in 1953.



J. S. NADER

In 1949 he joined the U. S. Public Health Service as a physicist at the Sanitary Engineering Center in Cincinnati. He was engaged in instrumentation and measurements of radioactivity in the Water Pollution Program until 1955. Since 1955 he has been engaged in the development of sampling and analytical instrumentation for the Air Pollution Engineering Research Program at the Center.

Mr. Nader is a member of the American Association of Physics Teachers, the Air Pollution Control Association, and the American Society of Heating and Air-Conditioning Engineers' Technical Advisory Committee on Odors.



James R. Nall was born in Ford City, Pa., on May 19, 1926. He attended the University of Kansas City, Mo., in 1944, and

then entered the Navy where he was assigned to the Medical Corps. In 1946 he entered George Washington University, Washington, D. C., receiving the B.S. degree in chemistry in 1952. While conducting undergraduate study he worked in the Department of Pharmacology at the University. He joined the staff of the Electron Tube Laboratory at the National Bureau of Standards in 1952. In 1956 he transferred to the Diamond Ordnance Fuze Laboratories where he is currently employed in semiconductor device research.



J. R. NALL

At the present time Mr. Nall is project leader in device development which is concerned with all devices applicable to ordnance electronic systems as well as micro-electronic systems.

Mr. Nall was a member of the DOFL group which received the 1957 Miniaturization Award presented by Miniature Precision Bearings, Incorporated. He is also a member of Alpha Chi Sigma and the Electrochemical Society.



Joseph J. Naresky (M'52) was born in Plymouth, Pa., on December 27, 1923. He received the B.A. in physics, summa cum laude, from Syracuse University, Syracuse, N. Y., in 1954 and is currently working, part-time, toward the degree in electrical engineering at the same school.



J. J. NARESKY

Mr. Naresky has been working in military electronics since 1941, in either a military or civilian capacity. He is currently chief of the Applied Physics Branch of the General Engineering Laboratory at Rome Air Development Center, Griffiss Air Force Base. The Applied Physics Branch is responsible for study and application of advanced techniques in the field of physics for the improvement of USAF ground electronic equipment.

This includes studies of such diverse areas as: component and equipment reliability; microwave techniques, transistorization, miniaturization, and utilization of new or little-used physical phenomena in the development of new electronic products.



Warren D. Nupp (A'45) was born on June 25, 1910, in Swissvale, Pa. He received the B.S. degree from the Carnegie Institute of Technology, Pittsburgh, Pa., in 1933 and the M.S. degree from the University of Pennsylvania, Philadelphia, in 1957, both in electrical engineering.

After work with electric power in the Pittsburgh area, he joined the Naval Aircraft Factory in Philadelphia in 1939 to work with the electrical accessories of naval aircraft. The association with the Navy's Bureau of Aeronautics, which started then, has continued to the present time. He is now with the U. S. Naval Air Development Center located in Johnsville, Pa., which he joined at



W. D. NUPP

its inception in 1947.

❖

M. J. Parker (M'46-SM'58) was born in Plymouth, Iowa, on November 6, 1920. He received the B.S. degree from the University of California at Berkeley in 1942 and the M.S. degree from the University of Maryland, College Park, Md., in 1950, both in electrical engineering. Since then he has completed most of the work toward a second Master's degree in business administration.



M. J. PARKER

He is now Project Manager for HASP (the naval weather rocket) at the Naval Ordnance Laboratory, Silver Spring, Md. He has been with NOL since 1942 when he was engaged in the design and development of underwater ordnance, including mine and depth charge mechanisms. His later work was mainly in connection with fuzing systems for surface-to-air projectiles and with miss-distance measuring systems. He has represented NOL as technical liaison in the United Kingdom, and has been engaged in guided missile fuze and systems development at NOL, Corona, Calif.

❖

Thomas A. Prugh (A'46-M'52-SM'56) was born in Dayton, Ohio, on June 11, 1920. He attended the University of Cincinnati, receiving the E.E. degree in 1942.



T. A. PRUGH

Following a short period with Lear, Incorporated, he joined the Signal Corps in 1943 and was subsequently assigned to a military electronic laboratory. After leaving the service in 1946, he continued working for the Department of Defense. In 1951 he transferred to the Ordnance Development Division of the National Bureau of Standards. This Division later became the Diamond Ordnance Fuze Laboratories, where he is currently em-

ployed. He has been working since 1952 on the application of transistors to ordnance electronic systems. He is now chief of the Micro-Systems Branch, which is concerned with the circuit and system aspects of microelectronics.

Mr. Prugh was a member of the DOFL group which received the 1957 Miniaturization Award presented by Miniature Precision Bearings, Incorporated. He is also a member of Tau Beta Pi, Sigma Xi, and Eta Kappa Nu.

❖

T. F. Rogers (M'50-SM'57) was born in Providence, R. I., on August 11, 1923. He received the B.S. degree from Providence College in 1945 and the M.A. degree from Boston University in 1949—both in physics. During the last year of the war he was engaged as a research associate in radar countermeasures work at the Radio Research Laboratory, Harvard University, Cambridge, Mass. In 1946



T. F. ROGERS

he worked on television circuit design at the Bell and Howell Company, Chicago, Ill.

He joined the Air Force Cambridge Research Center, Bedford, Mass., in 1946 where his primary work was in automatic radar data transmission. Assigned to M.I.T. as a staff member in 1950, he assisted in organizing one of the original groups at Lincoln Laboratory, Lexington, Mass., and continued there as an associate group leader for two years. He returned to AFCRC to form an Electromagnetic Propagation Laboratory of which he was assistant head until early in 1956. He was then appointed head of the AFCRC Communication Sciences Laboratory. He left government service in March, 1959, and is now assistant to the chief of the Radio Physics Division of the Lincoln Laboratory of M.I.T. His publications reflect his research work on electronic memory devices, ultrasonics, magnetism, and radiowave propagation and communication.

Mr. Rogers is a member of the American Physical Society, the Physical Society (England), the American Geophysical Union, and Commission II, URSI.

❖

Joseph L. Ryerson (A'55-M'55) was born on October 20, 1918, in Goshen, N. Y. He received the B.S. degree at Thomas S. Clarkson College of Technology, Potsdam, N. Y., in 1941, and the M.S. degree at Syracuse University, Syracuse, N. Y., both in electrical engineering.

Upon his graduation in 1941, he was employed as a development engineer by the Ward Leonard Electric Co. of Mt. Vernon, N. Y. In 1946, he was employed by the Associated College of Upper New York as a professor of physics and later transferred to Evansville College, Evansville, Ind., as a

professor of physics and electrical engineering. Here, in addition to teaching, he became a registered professional engineer and entered the consulting field. He was employed as an electronics engineer by the Rome Air Development Center, Griffiss Air Force Base, N. Y., in June, 1951 and at that time was assigned the problem of investigating automatic landing. He is now chief of the Advanced



J. L. RYERSON

Development Laboratory in the Directorate of Communications.

Mr. Ryerson is a member of Sigma Phi Sigma.

❖

Oscar A. Sandoz (SM'57) was born on October 30, 1921, in Nova Scotia, Can. He obtained the B.S. degree in physics in 1941, a diploma in engineering in 1942, and the Master's degree in physics in 1947, all at Dalhousie University, Halifax. He entered the University on a Regional Scholarship and was a student instructor throughout his under- and post-graduate years. His post-graduate work was in



O. A. SANDOZ

the field of electroacoustics.

During the period 1942-43 he was a senior research assistant with the National Research Council engaged on naval problems. He then joined the Royal Canadian Navy and served in the Torpedo School and later HMC Naval Research Establishment, dealing with mining, antitorpedo, and anti-submarine work. He was discharged with the rank of Lieutenant in 1946 and returned to the University.

When the Naval Research Establishment became a DRB laboratory in 1947, he joined the staff, and was employed there until 1956 as a scientific officer with the underwater physics group. As a senior staff member he was responsible for many development aspects of the underwater acoustical program. He was then assigned to the Radio Physics Laboratory of the Defence Research Telecommunications Establishment, Ottawa, as leader of the Ionospheric Propagation and Arctic Predictions Section.

In August, 1958, Mr. Sandoz was appointed assistant to the Chief Superintendent of the DRTE. Prior to this appointment, he was leader of the Applied Propagation Section of the Communications Wing.

❖

Wayne B. Scanlon was born in Youngstown, Ohio on March 21, 1916. He attended Wayne University, Detroit, Mich., receiving the B.S. and M.S. degrees in physics in 1938 and 1939, respectively. In 1948 he received

the Ph.D. degree from Purdue University, Lafayette, Ind.

He has been a research physicist at the Naval Ordnance Laboratory, Silver Spring, Md., since 1948. His first work was in connection with infrared radiation detectors. He has since specialized in solid state research and is now studying the chemical, electrical, and optical properties of polar semiconductors. This research applies to the fields of photoconductivity and thermal electric generators.

Dr. Scanlon is a Fellow in the American Physical Society and a member of Sigma Xi and the Washington Academy of Sciences.



W. B. SCANLON



Francis B. Silsbee was born at Lawrence, Mass., in 1889. He received the B.S. and M.S. degrees in electrical engineering from Massachusetts Institute of Technology in 1910 and 1911, and his Doctorate in physics, from Harvard, in 1915.

He joined the staff of the National Bureau of Standards in 1911, became Chief of the Electrical Instruments Section in 1931, and a Division Chief in 1946; he is now Chief of the Electricity and Electronics Division.

His scientific work has covered a wide range. Most of his publications deal with electrical measuring apparatus, particularly methods for testing current transformers and voltage transformers. One of his early papers on current transformers forms the basis of the current transformer testing set developed by the Leeds & Northrup Company, to which they have attached his name. He has also published papers on power factor and reactive power in polyphase circuits and on static electricity.

During World War I, Dr. Silsbee was leader of a group which carried on research and development on spark plugs, magnetos, and other components of the ignition systems of aircraft engines. The results of this work were published in a number of reports of the NACA. In 1916, he suggested a theoretical relation between the values of critical currents and of critical magnetic fields in the then newly discovered phenomenon of superconductivity at very low temperatures. Later work at Leiden confirmed this relation, which has come to be known as the "Silsbee Hypothesis."

During World War II, he directed research looking toward the mitigation of the lightning hazard to nonmetallic aircraft. The results of this work have also been published by the NACA. Later, he was in administrative charge of the later phases of the development of a special bomb director.



F. B. SILSBEЕ

More recently, he has published papers on the basic electrical standards and systems of units. He has been active in the standardization work of the AIEE, serving on its committees on Standards, on Instruments and Measurements (Chairman, 1941-1943), and on Research, and on numerous ASA sectional committees, including Definitions, Transformers, Code for Electricity Meters (Chairman), and on the U. S. National Committee of the International Electrotechnical Commission.

Dr. Silsbee is a fellow of the American Physical Society, of the AIEE, and of the American Association for the Advancement of Science. He is a member of the Philosophical Society of Washington (President, 1936), and of the Washington Academy of Sciences (President, 1950).



Claude C. Sims was born in Tampa, Fla., on May 9, 1924. He enlisted in the Navy in 1941, and served in the Pacific throughout World War II. He attended Tampa University, Tampa, Fla., and Florida State University, Tallahassee, where he received the B.S. degree in physics in 1950. After an additional two years in the Navy he joined the Navy Underwater Sound Reference Laboratory,



C. C. SIMS

Orlando, Fla., in 1952. For the past two years he has been in the Transducer Design Section, where he is now the head of the Research Branch.

Mr. Sims is a member of the Acoustical Society of America.



Ernest K. Smith (A'46-M'48-SM'55) was born in Peking, China, on May 31, 1922. He received the B.A. degree in physics from Swarthmore College, Swarthmore, Pa., in 1944. After graduation from college, he entered the U. S. Army, and from 1944 to 1945 he was with the Signal Corps Radio Propagation Unit. After being discharged, he joined the Mutual Broadcasting System as assistant radio engineer. He was promoted to chief of the Plans and Allocations Division of MBS in 1946 and retained the position until 1949. He enrolled at Cornell University, Ithaca, N. Y., in 1949, and received the M.S. degree in 1951. While at Cornell, he was a research assistant in the Cornell University Ionosphere Project. He received the Ph.D. degree from Cornell in 1956.

Dr. Smith has been on the staff of the Central Radio Propagation Laboratory,



E. K. SMITH

National Bureau of Standards, at Boulder, Colo., since 1951. He was on a leave of absence to Cornell University from 1952 to 1954. While with the CRPL, he has been project leader of the sporadic-E effects projects. He joined the Ionospheric Research Section in 1955 and has been project leader for the IGY Sporadic-E, World Maps, F2MUF, and Polar Plots projects. Since 1958, he has been Chief of the Ionospheric Research Section.

Dr. Smith is a member of Sigma Xi, the Scientific Research Society of America, and Commission III, International Scientific Radio Union, and is currently the International Vice-Chairman of CCIR Study Group VI.



E. E. Stevens was born on May 29, 1915, at Lexington, Mass. In 1931 he graduated from high school and in 1932 from the Maritime Business College, both in Oxford, Nova Scotia, Can.

During the period 1932-1938, he was employed by the Bank of Nova Scotia. He obtained an amateur radio license in 1934 and a commercial operator's certificate in 1938. During 1939, he was employed as a radio operator for the Canadian Marconi Company, and in 1940 joined the Department of Transport. After a varied experience, including supervision of ionospheric stations in the Canadian North, he joined the Defence Research Telecommunications Establishment. Since then he has been associated with their program of ionospheric measurements and high-latitude frequency predictions.



E. E. STEVENS



James L. Stewart was born in 1918, in Szechwan, China, and became a U. S. citizen in 1949. He obtained the B.A. degree in 1938, and the M.A. degree in 1940 from the University of Saskatchewan, Saskatoon, Can., working in slow neutrons. He then received the Ph.D. degree in 1943 from The Johns Hopkins University Baltimore, Md., with a dissertation in ultrasonics in gases which led to the discovery of rotational relaxation in hydrogen.

He served with CARDE, Valcartier, Quebec, Can., from 1943 to 1945 working on recoil ballistics, and was an assistant professor of physics at Rutgers University, New Brunswick, N. J., from 1946 to 1951. He joined the Navy Electronics Laboratory in 1951, specializing in the application of correlation analysis and techniques to underwater acoustics.



J. L. STEWART

Dr. Stewart is a Fellow of the Acoustical Society of America and a member of Sigma Xi



Robert J. Stirton was born in Berkeley, Calif., on June 30, 1923. He attended the University of California from 1942 to 1950 with an interruption for military service from 1943 to 1946. During that time he attended the State University of Iowa at Iowa City for four terms of electrical engineering. He received the B.A. in physics from the University of California in 1948 and continued there for two years of graduate study and research in physics.



R. J. STIRTON

From 1950 to 1956, he held a position in the Research Department of the Naval Ordnance Test Station as an exterior ballistician and is currently working as a consultant physicist in the Test Department of the Naval Ordnance Test Station.

Mr. Stirton is a member of the American Physical Society, Sigma Xi, and Phi Beta Kappa.



R. S. Thain (M'55-SM'58) was born in Nelson, British Columbia, Can. He served for two and a half years with the RCAF in World War II in radar work, being stationed in England for two years of this service. He returned to university after the war and graduated from the University of British Columbia, Vancouver, Can., in 1949, receiving the B.A. degree in physics.



R. S. THAIN

During the summer of 1948, he was employed by the Defence Research Board as a summer student in Ottawa, working in the field of radar meteorology. Following University graduation in 1949, he rejoined the Defence Research Board and was employed in radio propagation research at the Radio Physics Laboratory, concentrating his efforts mainly in the field of low-frequency radio communications. In 1953 he was made deputy head of the ionospheric communications research section, and in 1955 was transferred to Defence Research Board Headquarters as a Staff Officer in the directorate of physical research (electronics). He returned to the Radio Physics Laboratory late in 1957 for a brief period, and was appointed in April, 1958, as staff officer in the directorate of plans at Defence Research Headquarters.

Mr. Thain is a member of the Canadian Association of Physicists, and is a director of

the Professional Institute of the Public Service of Canada.



Gordon D. Thayer was born in Glen Ridge, N. J., on October 24, 1931. He attended Cornell University, Ithaca, N. Y., from 1949 to 1951, Newark College of Engineering, N. J., from 1954 to 1955, and the University of Colorado at Boulder, where he received the B.S. degree in engineering physics in 1957.



G. D. THAYER

Mr. Thayer joined the Boulder Laboratories of the National Bureau of Standards in 1957 and is presently with the radio-meteorology group. His projects with NBS include tropospheric analysis, mathematical analysis and data reduction (including IBM 650 programming), and research on problems of tropospheric refraction of radio signals.

Prior to joining NBS, Mr. Thayer was employed as a chemistry laboratory aide at the Cathode Ray Tube Division of Allen B. DuMont Labs, Allwood, N. J. in 1952, an instrument calibrator and checker at Weston Electric Instruments Corp., Newark, N. J. in 1952, and while in the U. S. Army Signal Corps, he was at White Sands Proving Ground working on missile guidance engineering analysis and ionospheric research analysis. He has also worked at the Evans Signal Lab, Belmar, N. J., as an engineering aide on reflection of radar signals analysis.

Mr. Thayer is an associate member of the Scientific Research Society of America, a member of Tau Beta Pi, Sigma Pi, and Sigma Pi Sigma.



Philip Thompson was born in England on April 20, 1926. He completed a science Tripos and received the M.A. degree at Cambridge University in 1946.



P. M. THOMPSON

After a short period at Automatic Telephone and Electric in Liverpool, he returned to Cambridge to join the Radio Physics Section of the Cavendish Laboratory. He spent two years with Salford Electrical Instruments before he sailed for Canada in 1950 to join the Radio Physics Laboratory of the Defence Research Board in Ottawa. In 1953, he was transferred to the D. R. B. Electronics Laboratory where he joined the Transistor Section.

Since then he has worked on various phases of transistor circuit development and has been responsible for several papers on the subject. Recently he has been the chief

circuit designer for the transistor navigational radar project.

Mr. Thompson is a graduate member of the British IEE and the British IRE, and is also a member of the Canadian Association of Physicists.



M. Lowman Tibbals (S'35-A'37-VA'39) was born in Seattle, Wash., on June 19, 1912. He received the B.S. degree in electrical engineering in 1935 from the University of Washington, Seattle. He did part-time graduate study at the University of Washington in 1936.



M. L. TIBBALS

He has been an engineer with the Colonial Radio Corporation (now Sylvania), Buffalo, N. Y., from 1937 to 1940, with Heintz and Kaufman, Limited, San Francisco, Calif. from 1940 to 1941, and with the C.A.A. radio engineering section, Washington, D. C., from 1941 to 1944. He worked as electronic scientist, with the U. S. Naval Electronics Laboratory, San Diego, Calif., from 1944 to 1952. In addition, he has been head of the navigation systems section at the U. S. Naval Electronics Laboratory since 1952.

Mr. Tibbals is a registered professional engineer in the state of California.



Robert W. Turner (S'37-A'39-M'55) was born in Littleton, N. H., on May 20, 1915. He received the B.S. degree in electrical engineering from Northeastern University, Boston, Mass., in 1937, and the M.S. degree in communication engineering from Harvard University, Cambridge, Mass., in 1940.



R. W. TURNER

He was employed from 1936 to 1939, by the General Radio Company as a member of the calibration laboratory. He joined the Federal Telegraph Company in 1940 as a receiver design engineer. From 1943 to 1946, he was an officer in the U. S. Navy and in 1946 joined the U. S. Navy Underwater Sound Laboratory, in New London, Conn., as a development supervisor in the Submarine Antenna Branch. Since 1951, he has been the Branch Head of this group.

Mr. Turner is a member of Tau Beta Pi.



Harold C. Urschel, Jr. was born in Toledo, Ohio, on February 17, 1930. He received the B.A. degree from Princeton University,

Princeton, N. J., in 1951, where he graduated with high honors from the Department of Biology. He was awarded the M.D. degree



H. C. URSCHEL, JR.

from Harvard Medical School, Boston, Mass., in 1955. Following graduation he was appointed to the surgical service of the Massachusetts General Hospital in Boston.

In 1957 Dr. Urschel entered the Navy and was assigned to the Naval Medical Research Institute, Bethesda, Md. As the Chief of Experimental Surgery, he is involved in problems of cardiovascular research and extracorporeal circulation.



Joseph H. Vogelmann (M'46-SM'49-F'59) was born in New York City, N. Y., on August 18, 1920. He received the B.S. degree



J. H. VOGELMAN

in 1940 from the City College of New York, and received the M.S. and Ph.D. degrees in 1948 and 1957 from the Polytechnic Institute of Brooklyn, N. Y., both in electrical engineering.

In 1945, after several years at the Signal Corps Radar Laboratory at Ft. Hancock and Belmar, N. J., he joined the staff of the Watson Laboratories in Red Bank, N. J. where he served as chief of the Development Branch until 1951, responsible for research and development of test equipment and microwave components and techniques. In 1951, he moved to the Rome Air Development Center, Griffiss Air Force Base, N. Y. From 1951 to 1953, he was chief scientist in the general engineering laboratory and consultant on UHF and SHF theory and techniques to the U. S. Air Force. From 1953 to 1956, he was chief of the electronic warfare laboratory, directing all research and development in ground based electronic warfare for the USAF. Since 1956, he has been technical director of the Communications Directorate with responsibility for the Air Force research and development effort in ground based and ground to air communications.

Dr. Vogelmann is a member of Sigma Xi, the AIEE, and the Armed Forces Communications and Electronics Association.



John D. Wallace was born in Fall River, Mass., on January 31, 1920. He received the B.A. degree in physics at Brown University, Providence, R. I., and has done graduate study in physics at Boston University, Boston Mass., and in physics and mathematics at Brown University. He has also studied in the fields of electronics at Ohio University, Athens, Ohio, and ceramics at

Pennsylvania State University, University Park, Pa.

In 1942, he was an instructor in physics at Amherst College, Amherst, Mass. From 1942 to 1946, he was Officer in Charge of Special Weapons on Project CUTE. The following three years, he was with the U. S. Navy Underwater Sound Laboratory, New London, Conn. Since 1949, he has been with the U. S.



J. D. WALLACE

Naval Air Development Center in Johnsville, Pa., where he has been head of the transducer section, head of the special methods and analysis branch, and is now assistant superintendent for research and development. He is president of Ferroelectrics Associated, Oreland, Pa. He is associate physicist in the division of cardiology at Philadelphia General Hospital and consulting physicist at the Germantown Dispensary and Hospital, both in Philadelphia, Pa.

Mr. Wallace is associate editor of the *U. S. Navy Journal of Underwater Acoustics*. He has received the Distinguished Service Award and the Arthur S. Fleming Scientific Award of the U. S. Junior Chamber of Commerce in 1956 and 1957, respectively. Presently, he has four patents to his credit and has applied for thirteen more. His fields of activity have included ordnance, submarines, anti-submarines, and sonar.

He is a member of the Acoustical Society of America, the American Physical Society, and the American Institute of Physics.



Eldon S. Warren was born in Digby, Nova Scotia, Can., on August 7, 1922. He received the M.Sc. degree from Dalhousie University, Halifax, Nova Scotia, in 1951, and the Ph.D. degree from Yale University, New Haven, Conn., in 1955.



E. S. WARREN

Since 1954 he has been at the Defence Research Telecommunications Establishment, engaged in ionospheric propagation studies.

Dr. Warren is a member of the Canadian Association of Physicists.



Everett C. Westerfield (M'56) was born at Beda, Ky., on April 3, 1901. He received the B.A. degree in physics in 1928, the M.A. degree in mathematics in 1930 and the Ph.D. degree in physics in 1940, all from the University of Colorado, Boulder.

He was employed as a process control engineer in charge of physical testing of aircraft material and parts from 1941 to 1945,

at Douglas Aircraft Company in Long Beach, Calif. From 1945 to 1946, he worked as an associate physicist at the University of California Division of War Research in San Diego, Calif., studying the scattering of underwater sound. Since 1946 he has been employed as a civilian scientist at the U. S. Naval Electronics Laboratory in San Diego, doing research and theoretical studies in sonar reverberation,



E. C. WESTERFIELD

communication and detection problems.

Dr. Westerfield is a member of the American Physical Society and the American Mathematical Society.



Philip Yaffee (A'43-M'46-SM'54) was born in Chelsea, Mass., on December 28, 1917. He received the B.S. degree with honor from Northeastern University, Boston, Mass., in 1942, and his M.S. degree from the University of Maryland, College Park, in 1952, both in electrical engineering.



P. YAFFEE

Since 1942, he has been on the engineering staff of the Naval Ordnance Laboratory, Silver Spring, Md., where he has been engaged in the development of instrumentation related to ordnance and missile systems. In 1953, he became active in work on methods for the measurement of missile and projectile firing errors. His major effort in this field has been spent on the AN/USQ-11 miss distance measuring system. He is now a senior project engineer in the Fire Control and Guidance division of the Laboratory.

Mr. Yaffee is a member of Tau Beta Pi.



Raymond A. Yocke was born on January 25, 1928 in Delaware, Ohio. He received the B.S. degree in physics at John Carroll University, Cleveland, Ohio, in 1958.



R. A. YOCKE

Mr. Yocke was commissioned in the United States Army in 1946 and served on active duty for a two-year period. He joined the National Aeronautics and Space Administration in 1954. During this time he has been engaged in the design and development of instruments and data reduction equipment for use in flight propulsion research facilities.

Scanning the Transactions

Electronic standards and measurements. We are compelled to make note of the jumbo December issue of the IRE TRANSACTIONS ON INSTRUMENTATION, which is devoted to the 38 papers presented last summer at the Conference on Electronic Standards and Measurements in Boulder, Colo. It is probable that of the 40-odd conferences in which the IRE participated last year, none will have any greater or broader influence on the future of the electronics field. The rate of progress of every branch of radio-electronics is dependent on improving the accuracy and extending the range of measurements of important physical phenomena. The above conference was of special importance because it focused much needed attention on the whole spectrum of electrical, radio and microwave measurement problems. These problems have now become acute because the electronics industry has expanded so rapidly in the past decade that its needs in this vital area have outstripped the development of standards and related measurement techniques. One of the most interesting aspects of current standardization work is the marked trend toward the development of primary standards based on universal physical constants rather than on arbitrary quantities. Length and time can now be expressed in terms of certain physical constants (*i.e.*, wavelength of light and microwave spectra of atomic and molecular systems) with greater precision than either can be referred to the primary standards. No doubt these and possibly other physical constants will eventually be adopted as the basis for primary standards as soon as the accuracy of measurement techniques is sufficiently improved. When that day comes we will not only be able to define basic units of measure with greater accuracy but we will also have universal standards that are immediately and freely available to all scientists and to all laboratories.

$1+1=1$. Ten years ago the foregoing combination of symbols would have been considered as nonsense by all right-thinking engineers, save for an enlightened few engaged in advanced switching circuit design. The rediscovery of Boolean algebra and its widespread application during the past decade to digital computer design has made such seemingly unorthodox expressions commonplace and has transformed them into meaningful concepts which have been of inestimable value in the analysis and simplification of computer logic circuits. If George Boole's special form of algebra is still a complete mystery to some IRE members we heartily recommend that they obtain the most recent issue of the IRE Student Quarterly (at only \$1.00 per copy or \$3.00 per year) for an exceptionally fine exposition of the basic elements of this important subject. (W. G. Schmidt, Jr., "Boolean algebra and the digital computer," IRE STUDENT QUARTERLY, February, 1959.)

A new type of inertial guidance device has been proposed which promises to overcome an important limitation of conventional inertial systems. It has been shown that an accelerometer, which is sensitive to both gravitation and acceleration, cannot be made to distinguish between these two effects. This has several important consequences for inertial navigation, which relies on accelerometer measurements to define the motion of the craft. The new inertial device mentioned above overcomes this difficulty by measuring the difference in gravitational field strength between two points within the craft, subtracting out the effects of acceleration in the process. From this information, the distance and direction to a known heavenly body can be accurately determined. Moreover, a new "vertical" can be established for each heavenly

body encountered on the trip, instead of requiring the instrument to maintain a preset reference alignment for the entire journey. This new instrument working in concert with present inertial devices could result in an entirely new approach to the guidance of orbital and space vehicles. (J. C. Crowley, *et al.*, "Some properties of the gravitational field and their possible application to space navigation," IRE TRANS. ON SPACE ELECTRONICS AND TELEMETRY, March, 1959.)

The diode is beginning to disappear from some types of computer circuits, judging by several very recent papers. Heretofore, the diode has been an indispensable element for providing isolation between various stages in a computer. However, the development of magnetic core logic elements, which are inherently self-isolating, is resulting in a noticeable trend toward the demise of the diode as a computer component. One paper on this subject appeared in the PROCEEDINGS only last January; two more were presented at the Philadelphia Solid-State Circuits Conference in February; another has now appeared in the PGEC TRANSACTIONS—all pointing to this important simplification in computer circuitry. (N. S. Prywes, "Diodeless magnetic shift registers utilizing transfluxors," IRE TRANS. ON ELECTRONIC COMPUTERS, December, 1958.)

Video tape recording techniques and problems were given a thorough going over last Fall at the Annual Broadcast Symposium in Washington, D. C. The primary use of video tape to date has been in providing a one- to three-hour time delay in broadcasting East Coast television shows to other time zones around the country. However, it is interesting to note that local stations, as well as networks, are finding a growing use for video tape recorders. One station reported it broadcast three hours of live programs during the day, necessitating an eight-man crew, using film the rest of the daytime. Meanwhile, in the evening another crew did virtually nothing during network time till a news program came in. To take full advantage of the night crew, it was decided to tape the live daytime shows the previous night, thus reducing the daytime staff. Although the initial cost of the video tape recorder and associated equipment was \$50,000, the savings in operating costs amounted to about \$45,000 a year. It appears that the video tape recorder is providing a long step forward toward the economical operation of the local television station. (H. W. Wessenberg, "The economics of video tape recorder operation," IRE TRANS. ON BROADCASTING, February, 1959.)

Much of the tedium of filter design has been eliminated by previously published tabulations of element values, pole locations, polynomial coefficients, and other pertinent design data for several conventional types of filters. The transient responses of some of the same types of filters have been computed and plotted in this paper. The curves presented are the normalized impulse and step responses of the first ten orders of the low-pass and high-pass Butterworth (maximally-flat) and Tchebycheff (equal-ripple) types, with pass-band ripple of 0.5, 1, and 2 db for the latter, and the low-pass Bessel (maximally-flat time delay) type. The method of calculation of the response is described in detail, and some interesting properties of the pole loci are explained. Incidentally, the calculations in this and several other papers in the same issue were performed with a digital computer, showing once again how important this new tool has become to network theorists. (W. Henderson and W. H. Kautz, "Transient responses of conventional filters," IRE TRANS. ON CIRCUIT THEORY, December, 1958.)

Books

Industrial Electronics Handbook, edited by William D. Cockrell

Published (1958) by McGraw-Hill Book Co., Inc., 330 W. 42 St., N. Y. 36, N. Y. 1355 pages+24 index pages+ xv pages. Illus. 6½×9½. \$22.50.

There may be few fields in industry where the systems concept is so easily manifested as in industrial electronics. The specialist is confronted with the necessity of integrating the design of his devices into an elaborate system of operations to be able to meet these requirements. He must have a thorough knowledge, not only of his own specialty and its fundamentals, but also of the other fields involved. In view of the extremely broad range of potential applications of electronic techniques and devices in industry, it is not surprising that reference works covering this area are becoming more and more voluminous. This recent one, edited by W. D. Cockrell covers nearly 1400 pages and was produced by about one hundred contributing authors. Such a large number of contributors is advantageous sometimes, but it certainly places a tremendous burden on the editor of such handbooks. As a whole, the one presently reviewed is surprisingly well organized.

It is gratifying to find in the introductory chapters, dealing with fundamentals, such necessary emphasis on the use of mathematical tools, running from algebra to operational calculus and computers. Of course, the reader must remember that handbooks are intended primarily as guideposts and reminders, showing directions of where to find what. From this point of view, this reviewer would like to see the bibliographical references at the end of each chapter somewhat extended in the surely forthcoming next editions. Subsequent sections cover control elements including circuit components, electron and solid state devices, magnetic amplifiers, rotating machinery, and mechanical, hydraulic and pneumatic devices used in control systems. Although emphasis in the book is on instrumentation and control, other phases of the industrial electronics art are not neglected. There is a well written chapter on high frequency heating, both induction and dielectric. Other sections cover in sufficient detail most of the elements and methods encountered in the broad field of industrial electronics. The mere listing of the section headings cannot do justice to the great amounts of material presented. They include fundamentals, control elements, constant potential power supplies, control circuits, circuit applications, instruments and computers, mechanical design of equipment, users requirements, letters patents in the United States, and technical information sources.

The individual, specializing in one narrow field, might be disappointed sometimes to see a rather terse treatment of his own specialty, but he will be well compensated when seeking and finding information on subjects somewhat more removed from his own specialty, yet still important to the industrial electronics art as a whole. The chap-

ters on control circuits and circuit applications are excellent. They contain much that the electronics engineer must know when entering the field of industrial applications. The chapter on the mechanical design of electronic equipment should be required reading for many. Perhaps it would help to reduce the inexcusable number of cases where often excellently designed circuitry fails simply on account of poor mechanical design.

There are no ready made recipes for special designs as such, to be copied, and the reader who looks for them will be disappointed. But he will find enough material pertaining to principles to lead his way toward more detailed information elsewhere, if he needs it. Most important, there is enough material to stimulate him to take a closer look at the fields adjoining his own specialty. This is as it should be, and the importance of this viewpoint cannot be overly emphasized in the special case of industrial electronics. Intended primarily for control and application engineers, "Industrial Electronics Handbook" will prove useful to anyone who wishes to familiarize himself with the rapidly expanding field of industrial electronics.

The few annoying misprints and misspellings of authors names in text and index will surely be corrected in the next edition.

EUGENE MITTELMANN
Consulting Engineer
Chicago 6, Ill.

Logical Design of Electric Circuits, by Rene A. Higonnet and Rene A. Grea

Published (1958) by McGraw-Hill Book Co., Inc., 330 West 42 St., N. Y. 36, N. Y. 183 pages+4 index pages+30 appendix pages+ix pages. Illus. 6½×9½. \$10.00.

This is a somewhat abbreviated translation of a book which was published in France in 1955. The authors are experienced designers of relay circuits, which constitute the principal subject matter of the book.

The chapters on Boolean algebra, two-terminal contact networks, geometrical representation of switching functions, and symmetry types of functions contain a wealth of interesting information, tables, and diagrams. The presentation is clear and elementary; it provides a generally admirable introduction to the subject but, unfortunately, gives no more than an introduction. Much worthwhile material is missing. Map techniques are dismissed in one brief paragraph, there is no mention of prime implicants (although the Harvard Minimizing Chart is discussed), and although the discussion of symmetry types is augmented by an up-to-date table of optimal two-terminal circuits for the 402 distinct types of functions of four variables, the most useful auxiliary tables to facilitate finding the representative of a given function are not provided.

The table of 402 circuits was prepared by E. F. Moore, and presents for each function the circuits using the fewest contacts and

the fewest springs known to date. In all but 38 cases the circuits have been proved minimal in contacts; the 38 are marked, and they present an interesting challenge. Although the treatment of sequential circuits is also clear and simple, it includes no discussion of the modern viewpoint associated with the names of Huffman, More, and Mealy. (A reference to Huffman's work which appeared in the French edition has been deleted in translation.)

The discussion of electronic logical elements is confined to 13 pages. No problems are included and no bibliography is provided; in fact the reviewer has found but a single citation of another book, and no references to papers.

Because of its clarity and simplicity, this book will be of interest to persons having curiosity about the subject but little background in it. It cannot be recommended to professionals because of important omissions mentioned above.

J. P. RUNYON
Bell Telephone Labs.
Murray Hill, N. J.

Oszillatoren mit Schwingkristallen, by Werner Herzog

Published (1958) by Springer-Verlag, Berlin/Göttingen/Heidelberg, Germany. 308 pages+5 index pages+4 bibliography pages+xi pages. Illus. 6×9. DM. 45. (in German).

The consideration of oscillator circuits in the same manner as filter circuits, described by Herzog in previous papers, e.g., "The Relation Between Oscillators and Filters,"¹ gave impetus to the book "Oszillatoren mit Schwingkristallen" ("Oscillators With Crystals"). This book is of interest particularly from the standpoint of circuit theory.

After a few short introductory remarks regarding crystal oscillators, the general theory of oscillators using electron tubes and general four terminal network considerations are covered (Chapters 2 and 3). The following chapter considers the frequency stability of oscillators. Nyquist's criterion of frequency stability is well presented by Dr. H. Lueg in Chapter 5. A great variety of oscillator types, e.g., Pierce, Miller, Heegner, Hartley, Mencham, and others, by converting filter circuits into oscillator circuits are considered in Chapter 6. The analysis of the bridge type oscillator using hybrid coils is novel. The treatment of transistor oscillators using modern four terminal network theory in Chapter 7 is excellent and of great value. The possibility of change in the resonance frequency and the correlating change in the quality of the circuit are discussed in Chapter 8. The concluding chapter on crystal and atomic clocks is well presented, but it is of limited value due to its brevity. The section on the feedback theory of oscillators would gain much by including some theory on servomechanism.

¹ Arch. elekt. Übertragung, vol. 1, pp. 47-58; 1947.

Although no practical considerations are given to the design or evaluation of oscillators, this book is of great theoretical value. This reviewer misses a survey of the parameters of crystals used in the various frequency ranges, particularly for the VHF and UHF frequency ranges using overtone crystals which would make the oscillator design more practicable. Crystals and their environments should be considered, as the limitation of the frequency stability is dependent on them. The references could be organized better by alphabetization or by a name index. The inclusion of further references, e.g., "Vacuum-Tube Oscillators," by W. A. Edson, would be of value.

This book can be recommended as presenting a broad background in the field of crystal oscillators and a detailed treatment of many oscillator circuits. The large variety of oscillator circuits presented certainly furnishes valuable suggestions for oscillator designers.

RUDOLF BECHMANN
U. S. Army Sig. Res. & Dev. Labs.
Fort Monmouth, N. J.

A History of Technology, Vol. V (The Late Nineteenth Century), edited by Charles Singer, E. J. Holmyard, A. R. Hall, and Trevor I. Williams

Published (1958) by Oxford University Press, 417 Fifth Ave., N. Y. 16, N. Y. 841 pages + xxxviii pages + 40 index pages. 415 Figs. 10 X 7 1/2. \$26.90.

The first four volumes of this "History of Technology" were published over a four year span from 1954 to 1958. They deal with the period from the Stone Age to the Industrial Revolution—up to about 1850.

This volume deals with the late nineteenth century, from about 1850 to 1900, and is the final volume of the series. Published in England, the work of some forty contributors and editors, it contains 34 chapters describing, in general outline under seven classifications, the processes and methods which had begun to develop in the technical and economic beginnings of the complicated industrial life of the twentieth century. This book therefore deals with many inventions made (principally in England but some in Germany, America and Russia) prior to the era of modern physics, but during a period when many branches of manufacture and production were beginning to be affected by scientific discoveries.

The subject matter is handled with a minimum of specialized technical terminology and is well illustrated with numerous figures and plates. Following the chapter texts there are extensive bibliographical references. The book should be very useful as a reference source for nineteenth century foundation material in the industrial fields.

The seven parts under which the material is presented are: I. "Primary Production" (food, steel and other metals, and petroleum); II. "Prime Movers" (steam engines and internal combustion engines); III. "The Rise of the Electrical Industry" (generation, distribution, and utilization); IV. "The Chemical Industry" (heavy chemicals, dyestuffs, explosives, and fine chemicals); V. "Transport" (railway engineering, ship-building, aeronautics, road-vehicles, cartography, and aids to navigation); VI.

"Civil Engineering" (building materials and techniques, bridges and tunnels, hydraulic engineering, and water supply); and VII. "Manufacture" (textiles, metals, machine-tools, ceramics, glass, printing, photography, and rubber).

The communication field is touched upon as one aspect of the utilization of electricity. References to the beginnings of the telegraph and telephone are followed by a page and a half of description of the early theoretical and experimental "wireless" activities, beginning with such workers as Clerk Maxwell and Heinrich Hertz, and ending with Marconi's successful transmission of wireless telegraph signals across the Atlantic on December 12, 1901.

Part VIII is entitled, "The Threshold of the Twentieth Century" and undertakes to deal with the educational, organizational, and social aspects of technology. Here one may read of the struggles which took place in England, during this period, in its efforts to improve technical education in order to compete with the superior technicians on the European continent. In industrial management the American form of corporate organization was demonstrating its vitality. American methods of mass production were beginning to make consumer goods cheaper and more universally available. Improved medical practices were leading to longer lives and increased population. Applications of technology were beginning to relieve human drudgery. Technology was bringing new power to machines, to the middle-class employer and to labor.

LAURENS E. WHITTEMORE
Short Hills, N. J.

Man's World of Sound, by John R. Pierce and Edward E. David, Jr.

Published (1958) by Doubleday & Co., Inc., 575 Madison Ave., N. Y. 22, N. Y. 258 pages + 9 index pages + 4 bibliography pages + 2 appendix pages + xiii pages. Illus. 5 1/2 X 8 1/2. \$5.00.

A former Editor of the IRE and his co-author, both of Bell Telephone Laboratories, have written a remarkable book on "the ability to communicate thought and emotion by means of language." The intensity of their interest communicates well to the reader that "the flexibility of speech and the analytical ability of hearing should excite far more wonder than the telephone, the radio and the phonograph, which can merely recreate sounds, or the most complex electronic musical instruments, which are inflexible indeed compared with the human voice."

There is a modern verbal treatment of the classical acoustical topics of sound waves, resonators, the vocal mechanism, the physical nature of speech sounds, the hearing mechanism, and theories and experiments in hearing. The authors present their "view of the phenomena of speech and hearing," occasionally taking issue with widely held viewpoints to the contrary. Information theory is discussed from a speech communication standpoint, with interesting observations upon the relative speed of thought, speech, and hearing processes.

Although the emphasis is modern, numerous references are made to historical and

even ancient happenings of great acoustical significance. There is excellent integration of acoustical knowledge gained from sources which are necessarily diverse, both in the scale of time and in the spectrum of different mental disciplines (e.g. the studies of physics, engineering, language, physiology, neurology, psychology).

The book is highly readable but cannot be easily scanned. It informs and challenges, with a special appeal to both amateur and professional scientists and engineers, for more experimentation as well as learning.

DANIEL W. MARTIN
The Baldwin Piano Co.
Cincinnati, Ohio

Electronic Avigation Engineering, by Peter C. Sandretto

Published (1958) by the International Telephone and Telegraph Corp., 67 Broad St., N. Y. 4, N. Y. 755 pages + ix pages. Illus. 6 X 9. \$9.50.

The scope of this book is indicated in the author's quotation of Webster's definition of the word "avigation," which is "the science or art of conducting aircraft in flight from one point to another." Technical coverage is restricted, for the most part, to devices that are in actual use or that have been given recognition by aviation policy making bodies, although some discussion of the trends of new developments is included.

The book is divided into four major sections which discuss devices involved in 1) the enroute long-distance zone, 2) the enroute short-distance zone, 3) the approach and landing zone, and 4) the airport zone. Each section includes its own introductory notes.

The first section includes chapters on airborne direction finders and radiophones (nondirectional ground stations); four course low-frequency radio range and markers; consol; some low-frequency developments; high-frequency direction finding from ground stations; Loran; electronic pilotage equipment (airborne radar); and electronic aids to dead reckoning.

The second section covers VOR, DME, other navigational aids, and Tacan. The third section covers airport surveillance radar, fixed beam low approach systems, radar low approach systems, and landing altimetry. The last section deals with airport surface detection equipment.

As the author states in his prefacing remarks, the book deals with electronic devices applied to transport aviation, which embraces operation of all types of aircraft, civil and military, engaged in moving from point to point. As such it certainly is written for a sizeable segment of the electronics profession. This reviewer feels that this book, which assumes the aspects of a "handbook," is "must" reading for electronics specialists in military aviation, the FAA, commercial aviation, and for those in industry associated with aviation electronics. The author has achieved a very excellent balance of the historical and tutorial with a judicious but frugal use of mathematics which is a joy to behold. References are carefully selected and adequate.

Although the book covers a considerable range of subjects, the coverage of each is by no means superficial. This reviewer, how-

ever, feels that information relating to instrument display and how the pilot utilizes the information provided in flight could have been covered more adequately.

Anyone reading this book will gain a clear insight to the growth and present status of the aviation art. The book is a fitting sequel to the author's previous publication entitled "Aeronautical Radio Engineering," (1942).

PAUL G. WULFSBERG
Collins Radio Co.
Cedar Rapids, Iowa

RECENT BOOKS

- Burnstein, Herman, *Fundamentals of High Fidelity*. John F. Rider, Inc., 116 W. 14 St., N. Y. 11, N. Y. \$2.95.
- Erikson, William H. and Nelson H. Bryant, *Electrical Engineering: Theory and Practice, 2nd Edition*. J. Wiley and Sons, Inc., 440 Fourth Ave., N. Y. 16, N. Y. \$8.00.

- A college textbook on circuits, machines, and electronics for non-electrical engineers.
- Hartley, H. A., *Audio Design Handbook*. Gernsback Library, Inc., 154 W. 14 St., N. Y. 11, N. Y. Paperbound \$2.90; hardcover \$5.00
- Hix, C. Frank, Jr., and Robert P. Alley, *Physical Laws and Effects*. J. Wiley and Sons, Inc., 440 Fourth Ave., N. Y. 16, N. Y. \$7.95.
- Langer, Rudolph E., Editor, *Symposium on Numerical Approximation*. The University of Wisconsin Press, 811 State St., Madison, Wis. \$4.50. Proceedings of a symposium conducted by the Mathematics Research Center of the United States Army at the University of Wisconsin, Madison, Wisc., April 21-23, 1958.
- Schure, Alexander, Editor, *Video Amplifiers*. John F. Rider, Inc., 116 W. 14 St., N. Y. 11, N. Y. \$1.80.
- Shrader, Robert, *Electronic Communication*.

- McGraw-Hill Book Co., Inc., 330 W. 42 St., N. Y. 36, N. Y. \$13.00. Electricity, electronics, and radio fundamentals required to pass all commercial and amateur radio licenses.
- Reliable Electric Connections, Third EIA Conference on Reliable Electric Connections, Dallas, Texas, December 2-4, 1958*. Engineering Publishers, P.O. Box 1151, N. Y. 1, N. Y. \$7.75.
- Research Highlights of the National Bureau of Standards, Annual Report, 1958*. U. S. Dept. of Commerce. For sale by the Superintendent of Documents, U. S. Govt. Printing Office, Washington 25, D. C. \$0.45.
- Rhodes, Donald R., *Introduction to Monopulse*. McGraw-Hill Book Co., 330 W. 42 St., N. Y. 36, N. Y. \$6.00.
- White, D. C. and H. H. Woodson, *Electromechanical Energy Conversion*. J. Wiley and Sons, Inc., 440 Fourth Ave., N. Y. 16, N. Y. \$12.50.

Abstracts of IRE Transactions

The following issues of TRANSACTIONS have recently been published, and are now available from the Institute of Radio Engineers, Inc., 1 East 79th Street, New York 21, N. Y. at the following prices. The contents of each issue and, where available, abstracts of technical papers are given below.

Sponsoring Group	Publication	Group Members	IRE Members	Non-Members*
Aeronautical & Navigational Electronics	ANE-5, No. 4	\$1.30	\$1.95	\$3.90
Audio	AU-7, No. 1	0.55	0.85	1.65
Broadcasting	BC-13	1.00	1.50	3.00
Circuit Theory	CT-5, No. 4	2.50	3.75	7.50
Education	E-2, No. 1	1.20	1.80	3.60
Electronic Computers	EC-7, No. 4	1.00	1.50	3.00
Instrumentation	I-7, Nos. 3 & 4	1.00	1.50	3.00
Space Electronics and Telemetry	SET-5, No. 1	1.80	2.20	5.40

* Libraries and colleges may purchase copies at IRE Member rates.

Aeronautical and Navigational Electronics

VOL. ANE-5, NO. 4,
DECEMBER, 1958

ECCANE Chairman—James A. Houston (p. 180)

Report on ECCANE—James A. Houston (p. 181)

Indication of the Vertical from Moving Bases—W. Wrigley, F. E. Houston and H. R. Whitman (p. 182)

A comprehensive review is made of the problem of indicating the vertical from moving

bases and methods for its solution, both theoretical and practical. This is based on research work carried out at the Instrumentation Laboratory of The Massachusetts Institute of Technology. The concept of Schuler tuning for precise vertical indication is shown to be a particular application of aided-tracking feedback systems in which proper use of the effect of vehicle acceleration can be made to cancel the dynamic tracking error that otherwise would exist. The 84-minute natural period resulting requires accurate instrumentation. A discussion of equipment suitable for this work concludes the paper.

Performance Profiles of Doppler Navigation Systems—Walter F. Fried (p. 194)

A method of determining position error of Doppler navigation systems is discussed that is based on the component errors and on statistical considerations. Sample calculations for four typical systems are presented. The historical behavioral performance of the three major components of a Doppler navigation system (the velocity sensor, the computer, and the heading reference) are analyzed and presented graphically. The weight trends of the components of Doppler navigation systems and of the complete systems are described. Possibilities of certain combinations of self-contained and ground-referenced navigation systems are discussed.

Airborne Direction Finding—The Theory of Navigation Errors—Clinton J. Ancker, Jr. (p. 199)

The situation where *n* stationary, accurately located, direction-finding stations fix the position of an unknown emitter by intersecting angular bearings has been previously investigated by Stansfield for the case where the angular bearings have small, normally distributed errors. The most probable location of the emitter and the surrounding confidence regions was determined.

Stansfield's solution is here extended to include the case of airborne direction finding, which introduces uncertainties in the locations of the direction-finding stations. These uncertainties arise from navigation errors and increase the uncertainty of the fix, which is reflected in an increase in the variance of the solution. The increase in uncertainty is derived for celestial, dead reckoning, and direction-finding (Loran, Shoran, etc.) navigational procedures. Small, normally distributed navigation errors are assumed. In addition, a model in which celestial position fixes are taken between direction-finding cuts is investigated. The various procedures are compared.

An Extension to the Theory of the Perform-

ance of Airborne Moving-Target Indicators—Harry Arkowitz (p. 210)

The performance of an airborne moving-target indication (MTI) system, designed to detect moving ground targets, is shown to depend upon the video autocorrelation function of the ground return, both with and without a target. The pulse-to-pulse video autocorrelation function is derived, and from it are obtained formulas for MTI cancellation and moving-target enhancement. Results are given only for a square-law detector.

Factors Influencing Target Detectability on CRT Screens—J. W. Ogland (p. 215)

Some factors that influence radar target detectability on a cathode-ray tube at high ambient light levels, such as occur in fighter airplanes, have been investigated. The signals considered are those that are intensity-modulated, as in B scopes and PPI. The problem is considered to consist of three parts that should be investigated separately: 1) the ambient brightness at which the indicator shall produce detectable signals, 2) the target spot brightness or intensity required for detection in this surround, and 3) the conversion of electrical signals into visual signals in the CRT. In this paper only 2) and 3) are discussed. It is found that with an intensity-modulated display, target detectability is determined by the spot intensity rather than by the spot brightness. The brightness of a conventional CRT shows an early saturation, but the intensity increases uniformly. Neither brightness nor intensity is a linear function of the radar parameters.

Some Aspects of the Thermal Design of Electronic Equipment Operating at 300–500°C Environmental Temperature—James P. Welsh (p. 220)

Aspects of electronic part-cooling techniques applicable to 300–500°C equipments are presented. Contrary to some beliefs, the minimization of thermal resistance at these temperatures is as important as at lower temperatures. The significant shifts in the natural modes of heat transfer which occur with high-temperature electronic parts are outlined, together with some recommended methods of cooling high-temperature parts and assemblies.

Correction to "Transac® C-1100: Transistorized Computers for Airborne and Mobile Systems"—Gerhard L. Hollander (p. 224)

Correspondence (p. 225)
Abstracts (p. 226)
PGANE News (p. 228)
Contributors (p. 230)
Suggestions to Authors (p. 232)
Annual Index 1959 (Follows page 232)

Audio

VOL. AU-7, No. 1, JANUARY-FEBRUARY, 1959

The Editor's Corner—Marvin Camras (p. 1)
Chapter News (p. 2)
Announcements (p. 3)
Performance of Enclosures for Low-Resonance High-Compliance Loudspeakers—James F. Novak (p. 5)

This paper analyzes the low-frequency behavior of the vented loudspeaker enclosure and the pressure-tight closed box. Inherent interrelationships of the speaker-amplifier system Q , efficiency and response balance are discussed. A method is described for design of small enclosures with very low-resonance high-compliance loudspeakers.

The author concludes that the vented enclosure can have greater acoustic output for a given amount of distortion, lower total harmonic, intermodulation, and transient distortion than a completely closed box of similar size.

A Sliding Class-A Audio Output System—Joseph A. Worcester (p. 14)

A single-ended transistor output stage is described in which the low battery drain and high-power output characteristics of push-pull Class B are achieved by a sliding bias produced by the signal.

The Delta Sound System for Television Receivers—Robert B. Donne (p. 16)

The delta sound system for television receivers incorporates the three features of AM compression, high-level audio output FM detection, and fundamental AM cancellation in a circuit having cost advantages.

Design information is presented for the system to enable the circuit designer to specify component parts and to predict the performance. AM compression in the range of 12 to 24 db and a peak-to-peak audio output voltage of about 60 volts with cancellation of the fundamental component of the undesired AM are obtainable with readily available low-cost tubes and circuit components.

Correspondence (p. 22)
Contributors (p. 23)

Broadcasting

VOL. BC-13 (Follows BTS-12), FEBRUARY, 1959

Fifth Annual Broadcast Symposium, IRE Professional Group on Broadcast Transmission Systems, Washington, D. C., September 26–27, 1958

A Preset Switching System—R. W. Rodgers (p. 1)

Automatic Control of Video Tape Equipment at NBC, Burbank—R. W. Byloff (p. 5)

CBS New York Video Tape Facilities—K. B. Benson and P. E. Fish (p. 14)

Panel Discussion of Video Tape Operational Experiences—R. M. Morris, chairman (p. 22)

Video Tape Production Problems—K. B. Benson (p. 24)

The Economics of Video Tape Recorder Operation—H. W. Wessenberg (p. 25)

Comments on Video Tape Cueing Methods—R. J. Bowley (p. 26)

Audience Participation Period (p. 30) PGBC Directory (p. 34)

Circuit Theory

VOL. CT-5, No. 4, DECEMBER, 1958

Recent Developments in Filter Theory—Vitol Belevitch (p. 236)

This paper is a review of the most important developments in filter theory during the post-war years. The 57 items contained in the bibliography appended to this paper have been selected either for their intrinsic novelty or because they contain reviews of earlier work with historical references. The integration of the various contributions has been attempted in this paper and several original ideas and results are included. No attempt has been made to comment on the practical value of insertion loss design vs image parameter design. As appears from the review, recent extensions of both methods have brought them closer to each other, and their cross fertilization has contributed to progress towards a unified filter theory.

On the Realizability of Ladder Filters—J. Meinguet and V. Belevitch (p. 253)

For practical design, it is useful to be able to determine in advance, from design parameters alone, the condition under which mutual

inductances will not appear in ladder networks designed on the image-parameter or on the insertion loss basis. The present paper starts with some comments on Fujisawa's criterion and extends it to a class of image-parameter band-pass filters. It expresses Fujisawa's criterion in terms of the design parameters for the symmetrical, low-pass filter having Tchebycheff behavior in both the attenuation and transmission ranges and for some degenerate cases. Various approximations yield simple results of great practical importance.

Synthesis of Band-Pass Ladder Network—Hitoshi Watanabe (p. 256)

The synthesis of frequency-asymmetrical band-pass filters which do not depend on the reactance-transformation from low-pass filters presents an important problem from the theoretical and practical points of view. In this paper, sufficient conditions are derived for realizing a nondissipative general "band-pass" filter by means of a reactive ladder network which excludes mutual inductance, except for an ideal transformer connected in cascade to one end thereof, and for realizing reactive "band-pass" ladders using a minimum number of inductors (or capacitors).

Moreover, the relation between negative values of reactive elements of the filter and its transmission characteristic is investigated. The method of synthesis for these kinds of ladders is presented in detail and some experimental results are included.

A Modification of Brune's Method for Narrow-Band Filters—R. F. Baum (p. 264)

If a given driving point immittance $Z(p)$ represents a narrow-band filter it can be developed into a ladder network without closely coupled coils by a modification of Brune's method.

The method is applicable to filters with phase or amplitude characteristic which are not symmetric with respect to the band center frequency.

$Z(p)$ is represented by the ratio of two polynomials in p with complex coefficients. A fictitious low-pass equivalent network is developed, which contains frequency-independent imaginary components. This network becomes realizable after a low-pass, band-pass transformation is made. It then represents a narrow-band approximation to the exact solution.

Synthesis of Three-Terminal Networks with Two Kinds of Elements—Hiroshi Ozaki (p. 267)

The present paper is concerned with the synthesis of LC, RC, and RL three-terminal networks without mutual inductance. It shows that the immittance matrices which satisfy the following sufficient conditions may be realized as networks of these kinds: (sufficient condition for RC case)

- 1) Admittance (or impedance) matrix satisfies the realizability conditions as a four-terminal RC network.
- 2) Numerator of y_{12} (z_{12}) is a polynomial with nonnegative coefficients whose zeros are restricted to the left half of the complex frequency plane including boundary, where the denominator is assumed to be a polynomial with nonnegative coefficients.
- 3) $(y_{11} - y_{12}) : (y_{22} - y_{12}) = 1 : n [(z_{22} - z_{12}) : (z_{11} - z_{12}) = 1 : n]$.

The theory may be applicable to the two important problems in network synthesis; that is, to the synthesis of filter circuits as three-terminal reactance networks and to the realization of RC transfer functions as three-terminal RC networks without mutual inductance.

Furthermore, for the case of a symmetrical circuit, the theory offers the theoretical method of transforming from a symmetrical lattice to an unbalanced form.

Four-Port Constant R Network Required by a Two-Way Single Amplifier Repeater—Charles A. Desoer (p. 276)

Two-way single amplifier repeaters have some definite advantages for long transmission systems. Such repeaters may be built according to various configurations. This paper concerns itself with only one of the possible configurations, namely, the one usually referred to as the four-filter configuration, and which is shown in Fig. 2.

First, the necessary and sufficient conditions for the physical realizability of the associated networks are derived. Second, the approximation technique is described that leads to physical networks having a scattering matrix very close to that required. The network realization proceeds following well-known methods. Finally, a particular design is outlined briefly.

On the Design of Filters by Synthesis—R. Saal and E. Ulbrich (p. 284)

Following an introduction to the methods and significance of filter design, this paper deals with definitions and filter synthesis according to prescribed attenuation requirements. It shows how the characteristic function for cases of practical importance can be calculated, particularly for the special cases 1) where this function is a power term in the frequency variable (maximally flat), 2) where it is a Tchebycheff polynomial function, and 3) where the effective attenuation in both the pass band and the stop band is of Tchebycheff character, and, also for 4) the general case of Tchebycheff behavior in the pass band with prescribed attenuation poles. The way, and the separate steps, of the determination of the circuit-element values for the realization of reactance ladder-network filters without mutual inductances are indicated, and the various necessary computation formulas are described.

For the first two special cases mentioned above, computation formulas applicable in general are given. New formulas are presented for the transformation of a low-pass into a band-pass configuration with a minimum number of inductances, and into a band-pass with the stray capacitance being taken into consideration. These are illustrated by several numerical examples.

Next, the paper deals with the design of filters, taking into consideration additional requirements involving the group delay. Possibilities of solving the problem are described. An actual example, together with the measured results, illustrates the excellent agreement between the theoretical determination and the behavior of the filter constructed accordingly.

The Appendix contains a selected section of a table which gives the normalized values of circuit elements for low-pass filters with the effective attenuation in the pass band and the stop band behaving in the Tchebycheff sense, for the degrees $n = 6, 7, 8, 9, 10, \text{ and } 11$. Graphs are also appended for easy estimation of the degree needed for a filter design and, thereby, the complexity of the resulting structure.

Note on Calculation of Ladder Coefficients for Symmetrical and Inverse Impedance Filters on a Digital Computer—J. K. Skwirzynski and J. Zdunek (p. 328)

The authors have been engaged for the last two years on a wide project of tabulation of ladder coefficients and subsidiary data for an important class of filters, designed on the basis of insertion loss theory and having Tchebycheff behavior in both the pass and the suppression bands. All calculations were done on the English Electric Digital Computer DEUCE (Digital Electronic Universal Computing Engine) by A. K. Rhodes, who was responsible for programming, computation, and tabulation of the results. The tables have now been ready for some time and publication of a volume "Tables of Ladder Coefficients and Related Data for Symmetrical and Inverse Impedance Filters,"

embodying these as well as a detailed instruction for their use is expected to take place early in 1959. The aim of this article is to describe generally the contents of these tables and also to relate certain interesting features of the programs used, as well as difficulties encountered during computation.

Transient Responses of Conventional Filters—K. W. Henderson and W. H. Kautz (p. 333)

Much of the tedium of filter design has been eliminated by previously published tabulations of element values, pole locations, polynomial coefficients, and similar data for several conventional types of filters. The transient responses of some of the same types of filters have been computed and plotted in this paper.

The curves presented are the normalized impulse and step responses of the first ten orders of the low-pass and high-pass Butterworth (maximally-flat) and Tchebycheff (equal-ripple) types, with band-pass ripple of 0.5, 1, and 2 db for the latter, and the low-pass Bessel (maximally-flat time delay) type.

The method of calculation of the responses is described in detail, and some interesting properties of the pole loci are explained. The calculations were performed with a digital computer.

An Iterative Method for the Direct Hurwitz-Factorization of a Polynomial—F. Jessie MacWilliams (p. 347)

If $N(p)$ is an even polynomial in p with real coefficients, and further, if any zeros of $N(p)$ on the imaginary axis occur to an even multiplicity, $N(p)$ can be factored in several ways into the form

$$N(p) = Q(p) \cdot Q(-p).$$

The particular $Q(p)$ which has zeros in the left half-plane only, (including, perhaps, the imaginary axis), is of importance in the theory of vibrating systems, and occurs frequently in theoretical methods of network synthesis, as for example, in Darlington's insertion loss method. It is now possible, owing to the work of the German mathematician, Bauer, to factor out $Q(p)$ directly from $N(p)$, by a method which is well suited to use in digital computers.

The first section of this article is essentially a translation of Bauer's original paper; the second summarizes the experience gained by writing a computer program based on his work.

Computation of Elliptic Functions of Rational Fractions of a Quarterperiod—H. J. Orchard (p. 352)

The evaluation of Jacobian elliptic functions with arguments which are rational fractions of a quarterperiod is often necessary in network design. Conventional methods, using theta functions, are quite satisfactory for this purpose, but, in the more interesting range of the modulus k near unity, they do not represent the best that can be done. A slight modification of the theta functions, together with a change of parameter made possible by the special nature of the rational-fraction form of argument, produces a computing scheme which, in most practical instances, converges more quickly than the classical one. The scheme is illustrated by a numerical example.

Numerical Determination of Cascaded LC Network Elements from Return Loss Coefficients—Daniel C. Fielder (p. 356)

The data for describing the behavior of a lossless network and for synthesizing a network which exhibits that behavior can be presented in many ways. One of the lesser used descriptions is the Taylor series expansion in s of the return loss about a transmission zero. As is well known, a return loss is the natural logarithm of the reciprocal of the reflection coefficient as measured between a resistance termination and the remainder of the network. While it is realized that the return loss expansion is not so immediately useful a network

function as, for example, the input impedance or the reflection coefficient, the analytical aspects are very interesting in themselves and may well find application in future work.

If the low-pass LC ladder network of n elements starting with a series L is considered, the first $(2n-1)$ coefficients of a given return loss expansion about the transmission zero at infinity contain all the necessary information for finding numerical values of the ladder elements. It can be shown that the first coefficient depends on the first ladder element, the third coefficient depends on the first and second elements, etc. Formulas for finding up to four elements from the return loss expansion are available. However, a recursion form for extending the range of these formulas is not immediately evident from these available formulas.

Two general equations, one for the series L 's and one for the shunt C 's are presented. The equations depend only on a knowledge of the Taylor coefficients for the particular type of ladder network under consideration. The method of finding the L 's and C 's is a straightforward algebraic approach and is novel only in that the elimination of redundant information leads to simple expressions for the L 's and C 's. Application of the equations leads to a recursion method for alternately finding an L , the succeeding C , the next L , etc. Accumulated results from one equation are used in finding the next equation.

Synthesis of Driving Point Impedance with Geometric Symmetry—R. F. Baum (p. 359)

If a driving point impedance shows geometric symmetry with respect to a center frequency ω_0 , it can be developed in Brune's sense by means of constant resistance lattices and without the use of coupled coils.

Note on Maximally Flat Delay Networks—I. E. Wood (p. 363)

A general solution in closed form is given for the reflection coefficient of a network with transfer function equal to the reciprocal of an n th degree Bessel polynomial. The result is applicable to a Darlington-type synthesis of a maximally flat delay network with unequal resistance terminations.

Reviews of Current Literature Abstracts of Articles on Circuit Theory (p. 365)

Correspondence (p. 367)
The 1959 Solid-State Circuits Conference (p. 377)

Education

VOL. E-2, No. 1, MARCH, 1959

Looking Forward—The Editor (p. 1)

The Time is Now—Eric A. Walker (p. 2)

To supply our economy with the professional manpower it must have to meet the political, economic, and military challenges of the Soviet Union and to harvest for the benefit of all people the tremendous potential of the scientific age, America must provide an education that will permit all young people to realize their full potential for mental and intellectual development. These are challenges to man's capacity to know and to understand and to his capacity for disciplined and responsible intelligence. To meet these challenges, America must 1) devise a diverse, multilevel collegiate structure aimed at the development of individual talents and abilities rather than at the inculcation of a fixed body of knowledge; 2) conduct basic research into the educative process to find ways of improving our instruction while increasing its efficiency; 3) re-examine our methods of teaching to find means of emphasizing the development of creativity; 4) establish systematic methods of identifying our talented youth early in their school programs; 5) effect a change in general attitude toward scholarship in our country; and 6) find an equitable and

satisfactory means of financing education in proportion to its value for America.

A Senior-Year Semiconductor and Electron Device Course—A. G. Milnes (p. 6)

Energy gap, minority carrier mobility, diffusion and lifetime, and electron emission and space charge concepts present difficulties for many students. In the course described here, considerable emphasis is placed on student experiments to create sound understanding of these ideas.

College Recruiting—A Portfolio of Expectations—W. T. Hudson, W. M. Hoyt, D. M. Cook, F. L. Carson, L. R. Hillyard, and D. T. Faticato (p. 10)

Each year America's industries and colleges join forces in an action aimed at converting the engineering student into an engineer employee. Principals in this operation are the company recruiter and the college placement director. Through the years each member of this unique alliance has come to learn the strength and weaknesses of his partner; each has profited and suffered through the other; each has developed his own ideas on how the other should operate.

IRE TRANSACTIONS ON EDUCATION recently asked a number of executives on company recruiting staffs and college placement offices for their opinions on the current state of college recruiting, and for suggestions on improving the operation in future years. This two-part article presents excerpts from some of the replies received.

Teaching Machines—B. F. Skinner (p. 14)
Role of the Laboratory in Engineering Education—Ronald L. McFarlan (p. 23)
Contributors (p. 28)

Electronic Computers

VOL. EC-7, NO. 4, DECEMBER, 1958

Editorial—J. P. Nash (p. 261)

A Magnetic Core Parallel Adder—Mao-Chao Chen (p. 262)

A logical design using magnetic core elements which does not have the usual carry time limitations is described. The synthesis uses a truth-table technique.

Significant Digit Computer Arithmetic—N. Metropolis and R. L. Ashenurst (p. 265)

The usual floating point arithmetic makes error analysis difficult. This paper describes an alternative system which offers a means of analyzing floating point calculations more effectively and which also possesses certain advantages from an equipment standpoint.

Minimal "Sum of Products Sums" Expressions of Boolean Functions—Shreeram Abhyanker (p. 268)

The problem of economical synthesis of circuits for digital computers leads to the problem of finding Boolean expressions of minimal length equivalent to a Boolean expression f . Previous authors restricted themselves to "sum of products" expressions; dualizing this gives "products of sums." The next more efficient step is to find minimal "sums of products of sums" expressions. In this paper, the basic concepts are formulated in Part I, and general theorems are given in Part II. In Part III, all the distinct minimal "sum of products of sums" expressions are obtained in case the cell complex of f consists of two isolated points, for the case of three isolated points partial results have been obtained which will be published in a later communication.

A Method for Synthesizing the Waveform Generated by a Character, Printed in Magnetic Ink, in Passing Beneath a Magnetic Reading Head—I. Flores and F. Ragonese (p. 277)

When a character printed in magnetic ink passes beneath a magnetic reading head it generates a waveform. A method is described here

to determine this waveform from the geometry of the printed character. The character shape is divided into elementary vertical units and the height of these units is then tabulated. A single set of experimental data is obtained in the laboratory by passing a magnetically printed bar beneath the same reading head which will be used to read the characters. Formulas are derived in this paper for combining the geometrical data obtained from the character with the experimental data obtained from the laboratory run with the head. A method is then described for programing a high-speed digital computer to derive the waveforms. The discussion explains why these waveforms are often superior to what might be obtained in an actual laboratory run of the printed characters.

On the Minimum Logical Complexity Required for a General Purpose Digital Computer—S. P. Frankel (p. 282)

A definition is provided for the term "general purpose computer" (gpc) which is compatible with usage and is analogous to, but distinct from, Turing's definition of a "universal computer." A gpc is presented in functional and logical design which seems to approximate the minimum complexity consistent with this definition. No specific definition of complexity nor a bound on required complexity is presented.

Iterative Combinational Switching Networks—General Design Considerations—E. J. McCluskey, Jr. (p. 285)

An iterative network is a combinational switching circuit which consists of a series of identical "cells" or sub-networks; for example, the stages of a parallel binary adder. A formal design method for iterative networks is presented. This is similar to the flow table technique for designing sequential circuits.

Some Properties of Boolean Equations—N. Rouche (p. 291)

Solubility conditions for a set of Boolean equations are established, first with respect to one variable, then with respect to all variables. By consideration of relations between minimal terms, a simple matrix form is deduced for Boolean equations. Using finite group theory and the properties of the characteristic equation of the matrix, a classification is introduced for Boolean mappings and their iterations, to which corresponds a classification of sequential machines.

Analysis of Sequential Machines II—D. D. Aufenkamp (p. 299)

Mealy's model of a sequential machine is assumed and a relation of "compatibility" of states is introduced to further the analysis of such machines. In the event that input restrictions exist it is often possible to effect combinations of states under this relation in addition to those permitted under equivalence of states, a relation previously studied. Compatibility of states is analyzed by an iterative technique, rigorously established, which makes it possible to determine readily connection matrices of simpler "compatible" machines.

Theoretical Consideration of Computing Errors of a Slow Type Electronic Analog Computer—T. Miura and M. Nagata (p. 306)

When solving differential equations using analog computers, there are two causes of errors: one results from the integrators and the other results from the coefficient setting elements. The former is independent of the order of the differential equation if the integrating time constants are all equal, while the latter is dependent on the setup procedure. The amount of the error from the coefficient setting elements is comparatively small in the operating frequency range from $\omega \leq \text{rad/sec}$. Therefore, theoretical analysis considering errors due to the integrators only is quite sufficient. However, higher accuracy is required in $\omega = 10$ to 100 rad/sec (for example in flight simulators)

and consequently the error caused by the coefficient setting elements has become impossible to neglect.

The authors have succeeded in formulating a generalized and practical equation developed from conventional formulas.

BIDEC—A Binary-to-Decimal of Decimal-to-Binary Converter—John F. Couleur (p. 313)

Simple, high-speed devices to convert binary, binary coded octal, or Gray code numbers to binary coded decimal numbers or vice versa is described. Circuitry required is four shift register stages per decimal digit plus one 30-diode network per decimal digit. In simple form the conversion requires two operations per binary bit but is theoretically capable of working at one operation per bit.

Diodeless Magnetic Shift Resistors Utilizing Transfluxors—Noah S. Prywes (p. 316)

Shift registers using magnetic cores have been conventionally designed with diodes to provide isolation between stages. It is possible to use transfluxors in a two-core-per-bit fashion so as to provide isolation between stages by magnetic means rather than with diodes. A design procedure for doing so is discussed.

Correction to "Analytical Design of Resistor-Coupled Transistor Logical Circuits"—M. W. Marcovitz and E. Seif (p. 324)

Correspondence (p. 325)

Contributors (p. 326)

PGEC News (p. 327)

SENESW, Science Education Subcommittee Newsletter (p. 328)

Annual Index 1958 (Follows p. 330)

Instrumentation

VOL. I-7, NOS. 3 & 4, DECEMBER, 1958

Conference on Electronic Standards and Measurements, August 13-15, 1958, Boulder, Colorado.

Abstracts (p. 128)

Papers from Boulder Conference on Electronic Standards and Measurements

Standards and Physical Constants (Abstract only)—Robert D. Huntoon (p. 128)

At the present time at the National Bureau of Standards, measurements at optimum magnitudes may be made in terms of our primary standards with probable errors of about 3 in 10^9 (1 kilogram mass), 3 in 10^8 (1 meter length), 1 in 10^8 (1 second time or more), and 3 in 10^7 (triple point temperature). Since all of these, except temperature, are presently defined in terms of prototype standards (even time in a certain sense), selected for their ready measurability and assumed stability, they may be subject to long-time changes. It is desirable, therefore, when the state of measurability permits, to define the primary standards in terms of physical constants.

Length and time can now be expressed in terms of certain physical constants with greater precision than either can be referred to the primary standards. This is accomplished through the measurement of the wavelength of light and microwave spectra of atomic and molecular systems. Great degradation in precision arises in measuring many of the other physical constants, particularly important elemental ones as h , e , and m . This comes about partly because electrical standards with probable errors of 6 in 10^8 (ampere), 4 in 10^6 (ohm), 7 in 10^6 (volt), and the transfer constant g (3 in 10^9) are involved in the measurement chain. Their degradation is partly due to the fact that nonoptimum magnitudes are used in their determination and partly because the experimental arrangement involved does not permit the correspondingly precise measurement.

The National Bureau of Standards has

measured a number of important physical constants during recent years, among which are the speed of light (c), the Faraday (F), the gyromagnetic ratio of the proton (γ), the ratio of charge to mass for elemental particles (e/m), and the second radiation constant (c_2). Precise determinations of the electric standards, the acceleration of gravity (g), and extensive inter-comparisons of optical wavelengths have also been completed recently.

Future plans call for measurement of several of these, particularly c , γ , F , and the transfer constant g . In these measurements, it is hoped to achieve a probable error in the accuracy approaching 1 in 10^7 in both c and g with appreciable improvement in the accuracy of γ and F . The recent and proposed determinations will be discussed.

Editor's Note: For a fuller account see: R. D. Huntoon and A. G. McNish, "Present status of research on the physical constants at the (U. S.) National Bureau of Standards," *Nuovo Cim.*, suppl. no. 1, vol. 6, ser. 10, pp. 146-184; 1957.

Standards and Measurements for Electronics—A. V. Astin (p. 134) No abstract.

Present Status of Precise Information on the Universal Physical Constants. Has the Time Arrived for Their Adoption to Replace Our Present Arbitrary Conventional Standards?—Jesse W. M. DuMond (p. 136)

Our present sources of precise information on the universal physical constants are reviewed, and the weak and strong points are discussed. Four primary unknowns α , e , N , and $\Lambda = \lambda_j/\lambda_s$, were chosen as the object of a 1955 least-squares adjustment. This adjustment was based on seven kinds of experimental input data whose precision of measurement was the order of a few tens of parts per million (ppm) and on eight auxiliary constants whose values are so much more precisely known as to place them outside the category of variables subject to least-squares adjustment. This led to least-squares adjusted output values for α , e , N , and Λ , with error measures of order ± 5 , ± 19 , ± 27 , ± 14 ppm, respectively. By combining these four values with the auxiliary constants, a long list of constants and conversion factors of physics and chemistry was calculated.

An earlier (1952) adjustment had thrown strong suspicion of systematic error upon some of the sources of information which were out of line with the rest and which had subsequently been shown, for good and sufficient physical reasons, to have been incorrect. In the 1955 adjustment these therefore were not used. Subsequent to 1955 errors have also been discovered in some of the theoretically calculated constants, an important example being μ_e/μ_0 , the electron magnetic moment anomaly. Experiments now in progress promise more reliable re-evaluations of some of the data avoiding the sources of error which vitiated the earlier results. Thus, in respect to the following six sources of information, it is hoped soon to have available new and much more reliable input data for a new and much improved least squares adjustment:

1) A redetermination of the electrochemical Faraday using the iodine coulometer by Dr. MacInnes at the Rockefeller Institute for Medical Research. Iodine is essentially an isotopically pure substance, and it is hoped that some of the uncertainty regarding the Faraday because of possible variations in isotopic abundance (and hence in mean atomic weight) occasioned by the electrolysis itself may thus be cleared up.

2) Similar work at the United States NBS in Washington, D. C. using other types of voltmeter.

3) Work now in progress on $\Lambda = \lambda_j/\lambda_s$, the conversion constant from x units to milliangstroms, earlier results on which (by Tyrén in Sweden) recently have been shown to be un-

reliable because of neglect of the Lamb shift.

4) Theoretical work on μ_e/μ_0 , the anomalous ratio of the electron magnetic moment to the Bohr magneton, which has recently shown the coefficient of the fourth-order correction term to be in need of a very substantial correction, a correction requiring a change of about 14 ppm.

5) Work of increased precision by Bergstrand and by Froome on the velocity of light.

6) Recent work by Hofstadter, at Stanford, to determine, by high-energy electron scattering, the finite extension of the proton dipole moment which may permit evaluation of an important, hitherto undetermined, correction term to the formula relating the hyperfine structure shift in hydrogen (measured by Prodel and Kush to 0.2 pp) to the fine structure constant, α , the velocity of light, c , and the Rydberg constant, R_∞ .

Emphasis will be given to the importance of using the method of least-squares if a unique and consistent set of values of the constants is to be obtained with a minimum degree of arbitrariness in the selection of input data. The value of least-squares methods for spotting data inconsistent with the majority (and, therefore, suspected of systematic error) is also pointed out. The more general applicability and significance of the least-squares procedure (which by no means is limited to data with Gaussian error distributions only, but permits inclusion of any distributions whatever having finite second moments) will be explained, a result well known among the students of error statistics as the Gauss, Laplace, Markoff theorem. The need for careful attention to correlation coefficients and the use of generalized quadratic formula for computing propagated errors will be emphasized.

On the basis of the precision at present attainable in our knowledge of the universal constants expressed in terms of our present arbitrary conventional cgs units as compared with the precision with which those arbitrary conventional cgs units themselves can be defined, it is probably still premature to adopt any far-reaching system of units based on the universal constants. It is also premature to do this now because there is still a great embarrassment of choice as to how to set it up—we do not know enough fundamental theory to select wisely a set of universal constants as a basis for a system of units. The adoption of spectroscopically defined wavelengths as standards of length and atomic clocks as standards of time, because of their higher reproducibility and convenience rather than because of any deep fundamental theoretical significance they possess, seems to the present writer about as far as it is safe to go at the present time in the direction of revision of units and standards.

A Free Precession Determination of the Proton Gyromagnetic Ratio—P. L. Bender and R. L. Driscoll (p. 176)

A determination of the proton gyromagnetic ratio has been made in a magnetic field of about 12 gauss. A water sample was polarized in a strong magnetic field and then fired through a pneumatic tube into a precision solenoid. After arrival the protons were re-oriented by about 90° with respect to the solenoid field by a RF pulse and the subsequent free precession signal induced in a pick-up coil was observed. Short-term fluctuations in the earth's field and in the solenoid current gave a scatter of about one part per million in individual frequency measurements. A preliminary value of $\gamma_p = (2.67513 \pm 0.00002) \times 10^4$ gauss $^{-1}$ sec $^{-1}$ has been found for the gyromagnetic ratio of protons in water, uncorrected for diamagnetism. Radiation damping of the free precession signals is strong under certain conditions and can cause small frequency shifts.

The Absolute Determination of g (Abstract only)—H. Preston-Thomas (p. 129)

The acceleration due to gravity (g) enters into the definition of the unit of current as a force and into certain fixed points on the International Temperature Scale as a pressure. The experimental realization of these quantities involves, therefore, a determination of the absolute value of g and this should be known to the order of a part in a million if it is not to constitute one of the limiting factors in the accuracy of the experiment.

Although it is theoretically possible to determine g in any one of a number of ways, in practice it is done by observations on either compound pendulums or bodies in free fall. Until about 1950 the former of these two methods was invariably used, since the measuring and timing techniques available did not have the necessary accuracy when applied to bodies in free fall. Unfortunately, although the main sources of error in pendulum experiments are known (flexure of the support, flexure of the pendulum, deformation and indeterminate shape of the knife edge), estimating their magnitude is a matter of extreme difficulty and it is by no means certain that residual systematic errors of the order of a few parts per million do not remain even in those pendulum experiments which are generally allowed to have been the most carefully carried out and the most reliable.

During the last few years the general availability of timing systems with the necessary (a few parts in 10^6) short-term accuracy has resulted in the setting up of free fall experiments in some half a dozen countries. These involve the photographing of a free falling bar at known times, the point illumination of a free falling bar coated with a photographic emulsion at known times, or the photoelectric timing between fixed points of a sphere that is initially thrown upwards and then falls. Although very few results have been published so far, the approximate limits of accuracy of these experiments can be estimated; it is about two parts per million for the first two cases and about half a part per million for the third, the increased accuracy being due to the initial throwing upwards of the sphere.

Although the particular sources of error that affect the pendulums are not present in the free fall experiments, a number of other experimental difficulties arise. The optical paths must be horizontal to within a few minutes of arc; any rotation of the rules about a horizontal axis must be limited to a few minutes; temperatures must be known and stable, often to within about 0.02°C over a path length of two meters and under conditions that make thermostating difficult; positions must be known to a few tenths of a micron or less, also under adverse conditions; the drops must be monitored for microseisms, which affect free fall measurements to a considerably greater extent than they do the pendulum ones, and so on.

In sum, these experiments constitute a formidable metrological problem and as such take a number of years to set up and complete. However, it is expected that at the next international conference on geodesy and geophysics at Helsinki in 1960 the results of about ten determinations of g from various laboratories will allow the present Potsdam system, known to be in error by between 8 and 20 parts per million, to be replaced by one that is accurate enough for present metrological requirements.

Atomic Beam Sources and the Standard of Length—K. G. Kessler, R. L. Barger, and W. G. Schweitzer (p. 181)

Three devices which utilize the wavelength of optical transitions in an atomic beam of mercury as a standard of length are described: 1) a beam of atoms excited by electron impact; 2) a beam in absorption which utilizes a Michelson interferometer to produce a narrow line in emission; and 3) a beam in absorption utilized as a narrow band detector. The half-width of

the 2537 Å line, observed in both emission and absorption, is 0.002 cm^{-1} and is equivalent to a coherence limit of 500 cm. The comparable half-widths for the 5462 Å mercury and the 6058 Å krypton lines produced in cooled discharge tubes are 0.018 cm^{-1} and 0.011 cm^{-1} .

Ultra-Precise Quartz Crystal Frequency Standards—A. W. Warner (p. 185)

The Bell Telephone Laboratories have been active for the past eleven years in the development of high-precision, high-frequency quartz crystal oscillators to meet the requirements of the Bell System and the armed services. At the present time a concerted effort is being made to improve the frequency stability of a fixed station oscillator to 1 part in 10^{10} . The goal is an oscillator whose frequency averaged over a 1-second period will not differ by more than 1 part in 10^{10} from that averaged over any other second within a period of one month. Two phases of the work will be described: 1) use of temperatures near absolute zero to increase Q and frequency stability; and 2) refinement of the best crystal unit designs for use at room temperature.

In addition, methods of improvement in frequency stability under conditions of shock, vibration, and acceleration will be described.

Use of Internal Friction Measurements in Determining the Causes of Frequency Instabilities in Mechanically Vibrating Frequency Standards—Warren P. Mason (p. 189)

Internal friction measurements for quartz, carried out at low temperatures, have shown the presence of two types of relaxations. One of these is connected with the impurity content, and the height of the attenuation peak is affected peak is affected by the amount of aluminum present. This relaxation is affected by x radiation.

Measurements have been made at room temperature for high stress ranges and three stress dependent regions have been found. The first of these appears to be due to the break-away of dislocations from pinning points and the loss follows a Granato-Lücke plot. The next region appears to be due to Frank-Read sources, and an evaluation can be made of the dislocation distribution from the internal friction measurements. The final region appears to be due to the unrestricted production of dislocation loops by the Frank-Read mechanism, and it ends up in fracture of X cut crystals for stresses of from 7 to 9×10^8 dynes/cm².

Crystals strained by considerable amounts drop in frequency initially and show a slow increase with time. This effect appears to be due to the production of dislocations which become pinned by impurity atoms or vacancies. Since such dislocations can be generated by the fabrication process, this effect appears to be the cause of long time aging of the elastic properties.

It is pointed out that germanium or silicon crystals may be more stable frequency standards than quartz, since no dislocations can be generated up to the fracture stress at room temperature. Low internal friction and low temperature coefficients are available at low temperatures.

Cesium Beam Frequency Standard Development in Canada (Abstract only)—S. N. Kalra, R. Bailey, and H. Daams (p. 129)

A few years ago a standard of frequency based on Essen ring crystals was established in Canada. It has been found that some of these crystals have a very high order of stability over both short and long periods of time. A periodic check of frequency against an atomic beam resonator would be quite adequate. It was felt that an absolute standard should be as simple and direct as possible. It was therefore decided to build a resonator type of standard first.

Starting with the assumption that the principles of this type of a standard are well under-

stood, the basic design of the system at the National Research Council, Ottawa, is described.

The atomic beam apparatus is designed as a research apparatus with a fair amount of flexibility built in. This necessarily complicates the design and many of the features of this model will, of course, be absent from the unit designed for routine operation.

The over-all length of the apparatus from source to detector is about 210 cm and the distance between the oscillatory fields is 80 cm in the Ramsay technique of obtaining resonance. The detector presently in use is a hot wire type of ionizer with a collector for ions, the current being measured with a vibrating reed electrometer. Almost every control point is adjustable.

A good crystal oscillator at 5 mc (nominal) is used to obtain the exciting field. The frequency is synthesized using vacuum tube circuits and a final crystal diode tripler. The crystal oscillator is offset in frequency to obtain the exact frequency of resonance. The measurement of frequency is accomplished by using another multiplier chain starting with the 100-kc standard reference crystal oscillator and going up to 900 mc. A signal at a frequency of 10×900 mc is mixed with the signal exciting the resonance in a balanced microwave mixer. The beat frequency is mixed with the second harmonic of 100 mc to give a second beat frequency of about 7 mc. This can be measured by conventional techniques to an accuracy of about 1 cps using an integrating time of one second, thus giving an accuracy of measurement of $1:9.2 \times 10^9$. The phase stability of the synthesizing chain has been measured to be good enough to give this accuracy of measurement. The width of the resonance line is about 290 cps.

International intercomparison is accomplished by phase measurement of the MSF transmission at 60 kc from Rugby, England, and by precision time signal measurements of WWV and others.

The Ammonia Maser as an Atomic Frequency and Time Standard—R. C. Mockler, J. Barnes, R. Beehler, H. Salazar, and L. Fey (p. 201)

The ammonia maser shows some promise as a frequency and time standard. It has distinct advantages over atomic beam apparatus in frequency stability, signal-to-noise ratio, and physical size. In addition, it is an oscillator and as such the measurement of its frequency is simpler than for other nonoscillating devices. The maser has, however, the basic shortcoming that the output frequency is not easily reproduced.

Either a precise recipe for construction in the case of single beam masers or precise symmetry in double beam apparatus appear essential to reproducibility. The effects on the maser frequency resulting from cavity pulling, non-uniform radiation of the beam inside the cavity, hyperfine structure of the spectral line, variations in beam flux, thermal effects, and magnetic fields, are either avoidable or can be controlled with an adequate degree of precision to fix the frequency to within an estimated 3×10^{-10} .

Our measured signal to noise ratio is 10^4 . The frequency stability for periods of a few minutes exceeds 1×10^{-11} . A preliminary frequency for the $J=3, K=3$ ammonia transition is: $23,870,129,007 \pm 10$ cps (precision) ± 100 cps (accuracy).

Results of the Comparison: Atomichron®—British Cesium Beam Standard—A. O. McCoubrey (p. 203)

During March of this year, two Atomichrons were compared to the cesium atomic beam frequency standard at the National Physical Laboratory, Teddington, England. The tests included systematic measurement of

frequency differences following repetitive misalignment of the Atomichrons and realignment according to a procedure established for field use. In addition, tests were carried out to establish the origin of the small frequency difference which was found. Analysis of the data indicated the Atomichrons to be within one part in 10^{10} of each other and about 3.5 parts in 10^{10} high with respect to the NPL standard. The standard deviation associated with the measurements was about ± 5 parts in 10^{11} . The methods used in making the measurements, the results of the comparison, and the probable sources of the residual frequency difference will be discussed.

Recent Long Distance Frequency Comparisons—J. A. Pierce (p. 207)

Since March 1, 1958, a number of comparisons have been made between the frequency of an Atomichron® at Cruft Laboratory, and the frequencies of other Atomichrons in Boulder, Colo. and Washington, D. C. Comparisons have also been made with the cesium resonator at the National Physical Laboratory at Teddington, England.

A limited number of daily direct comparisons with Boulder had a standard deviation of 1.4 parts in 10^{11} . This value may be determined by the uncertainties of the phase of the low-frequency transmission at 60 kc, or by changes in the Atomichrons.

Comparisons with Washington and Teddington using a VLF carrier had a standard deviation of about $2/10^{10}$ that was apparently controlled by vagaries of the crystal oscillator at the transmitter. The vagaries of high-frequency WWV transmission, even under the best conditions, limited the standard deviation at about $4/10^{10}$.

It appears that the average frequencies of different Atomichrons, aligned and operated in different ways, may be in very close agreement or may differ by as much as 3 parts in 10^{10} .

Introductory Remarks to Session on Direct Current and Low Frequency Standards—Francis B. Silsbee (p. 211) No Abstract.

The Derivation of Resistance, Inductance and Capacitance from the NPL Primary Standard of Mutual Inductance—G. H. Rayner (p. 212)

The determination of the unit of resistance at the National Physical Laboratory depends upon a primary standard of mutual inductance and frequency. The measurement is carried out on the Campbell bridge; this has recently been modified to make its operation simpler and to improve its accuracy. The calibration of inductors is usually obtained from the primary standard without the introduction of any other electrical quantities, although in some cases it is convenient to measure inductance in terms of capacitance and frequency or resistance. The measurement of capacitance is made with a bridge of the Carey Foster type and is obtained in terms of mutual inductance divided by the product of two resistances. In addition subsidiary standards are required for determining the phase-angles of resistors, inductors, and capacitors.

The Construction and Characteristics of Standard Cells—George D. Vincent (p. 221)

The development, construction, maintenance, and use of standard cells, are described. The characteristics of cadmium cells are considered, with special reference to those which lead to best adaptation to particular requirements. Recent improvements and needs for further improvements or new developments are discussed.

AC-DC Transfer Instruments for Current and Voltage Measurements—Francis L. Her-mach (p. 235)

AC-DC transfer instruments provide the vital link between the fundamental dc standards and the measurement of alternating voltage and current. Electrodynamic, electrostatic,

and electrothermic transfer instruments can be used for measurements of better than 0.1 per cent accuracy at power and audio frequencies. The paper outlines their characteristics, the limitations on transfer accuracy, the methods by which these limitations are overcome, and the useful range of voltage, current, and frequency. In their best ranges accuracies of 0.01 per cent can be achieved with all three types of instruments.

An Interim Report on the Inductronic Electrodynamometer for the Precise Measurement of Voltage, Current, Power, and Energy—R. F. Estoppey (p. 241)

The Inductronic Electrodynamometer was developed for use as a product resolver for the multiplication of two electrical inputs to produce a direct current output accurately related to the product of the two input currents. This unit has now been extended to include the measurement of volts, amperes, and watts to within 0.1 per cent over a frequency span of dc to 2500 cycles. The instrument consists of an electro-dynamometer mechanism and a permanent magnet moving coil mechanism having a common shaft and connected so that the restoring torque for the dynamometer is supplied by the permanent magnet mechanism. Torque is automatically balanced by a sensing means and a feedback amplifier which produces a current in the dc element to balance exactly the dynamometer torque.

Energy is measured by connecting the output current of the Inductronic Wattmeter to a precision integrator. The integrator measures the time integral of the millivolt drop of this current through a standard resistance by opposing this voltage with the voltage produced by a tachometer generator driven by a motor controlled in turn by the amplified difference in millivolt seconds. The revolutions of the generator are a measure of millivolt seconds and a register geared to the motor is used to make the count, which is equivalent to watt hours.

The Precise Measurement of Small Capacitances—A. M. Thompson (p. 245)

The National Standards Laboratory has developed a capacitance standard whose value in electrostatic units can be computed accurately from a measured dimension. By extending and refining known principles it has been found possible to use this computable capacitor for calibrating working standards in a wide range of values. Recent determinations of the velocity of light are considered to be accurate to 1 to 10⁶; the values of these capacitors are thus known precisely in practical units.

This paper reviews the concept of direct capacitance and shows how one of the direct capacitances of a 3-terminal capacitor may be perfectly definite. Such definite direct capacitances are strictly additive, and if the solid dielectric supports used in their construction are suitably positioned the direct capacitance may be practically loss free. Attention is given to the general design features of closely coupled transformer ratio arms which provide voltages of accurately known ratio. Practical transformer bridges suitable for the measurement of small direct capacitances are described and their accuracy analyzed.

New Apparatus at the National Bureau of Standards for Absolute Capacitance Measurement—M. C. McGregor, J. F. Hersh, R. D. Cutkosky, F. K. Harris, and F. R. Kotter (p. 253)

Techniques have been developed which enable a small calculable capacitor of about 1 pf to be used as a standard for the measurement of capacitance to an accuracy of at least 1 in 10⁶. Measurements are made at audio frequencies and are based on the use of 3-terminal capacitors and bridges with transformer ratio arms.

Transformer ratio arms using interleaved windings on a high permeability core give

ratios accurate to at least 1 in 10⁶. Ratios accurate to 1 in 10⁷ may be obtained by adding a third winding and an electromagnetic shield.

Small direct capacitances may be measured with transformer bridges without appreciable error from the ground capacitances. With a bridge supply of 100 volts at a frequency of 1592 cps and a narrow-band vacuum tube amplifier as the detector, a sensitivity of 0.1 μ pf may be obtained.

Separate coaxial cables are used for the two leads to a capacitor and are threaded through a high permeability core to eliminate errors due to their finite impedance, thus increasing the impedance of the ground loop and suppressing unwanted circulating currents.

Small 3-terminal capacitance probes are suitable for the measurement of very small mechanical displacements.

High-Frequency Standards of the Electronic Calibration Center ANBSBL—M. C. Selby (p. 262)

Equipment to calibrate high quality reference standards of power, voltage, attenuation, and field strength at 30 kc to 300 mc and higher frequencies has been developed for use in the Electronic Calibration Center of the National Bureau of Standards Boulder Laboratories, Radio Standards Division. The major objective is maximum absolute accuracy consistent with the practical expedient of rendering rapid calibration services to government, armed services, and private scientific laboratories. Principles involved and special features of this equipment are discussed. This description, which covers only a fraction of the equipment planned for the Center, gives a general idea of the type of services anticipated at the National Bureau of Standards for an integrated national RF standardization program, with continuous improvements in the quality and accuracy of primary and reference standards in general.

High-Frequency Impedance Standards at the National Bureau of Standards—R. C. Powell, R. M. Jickling, and A. E. Hess (p. 270)

The impedance standards and techniques used at the National Bureau of Standards in the frequency range from 30 kc to 300 mc are described. This paper covers the primary standard and how it is obtained, the comparison methods used in calibrating the working standards and instruments, and finally the equipment used in the new Electronic Calibration Center to make measurements for other laboratories. A survey of the field showing the instruments and standards involved, the limitations on standards presently available, in both range and accuracy, and problems involved in accurate calibrations are included.

Calibration of Signal Generator Output Voltage in the Range of 100 to 1000 Megacycles—A. L. Hedrich, G. U. Sorger, B. O. Weinschel, and S. J. Raff (p. 275)

Considerations in the calibration of signal generators are discussed with the advantages and disadvantages of the various methods as applied to high frequencies. Methods involving the direct and indirect measurement of output voltage are discussed. A system using an intermediate frequency substitution method as reduced to practice under Air Force Contract No. AF 33 (60C) 25238 is described in detail.

The unknown signal is converted to an intermediate frequency in a linear mixer and then compared to a standard signal, the level of which has been set by a piston attenuator. A precision means is provided to indicate equality of the two voltages. The amount of change in the setting of the piston attenuator to maintain equality of the voltages is thus a measure of the ratio of the unknown at two different levels.

Design and Development of a Standard White Noise Generator and Noise Indicating Instrument—H. Zucker, V. Baskin, S. I. Cohn, I. Lerner, and A. Rosenblum (p. 279)

The design and development of a standard

white noise generator in the frequency range from 0 to 1000 mc are presented. The basis of the generator is Nyquist's Law which relates the noise output of a resistor to its temperature. The noise generator consists of a low reflection termination heated in a coaxial furnace. The basis of the design of the low reflection termination is presented together with experimental impedance measurements. The operating temperature of the generator is currently 1300°C. The techniques used in the development of a resistive material to withstand high temperatures are presented together with the temperature characteristics of the termination. The thermal emission effects which occur at the operating temperature of the generator are analyzed, and possible causes for deviation of the generator from a true standard are considered.

The design of an instrument to measure the linearity of the generator's noise power output as a function of its temperature and the design of a noise indicating instrument to measure the noise output of generators at spot frequencies in the frequency range from 10 to 500 mc are given.

A Dry, Static Calorimeter for RF Power Measurement—P. A. Hudson and C. M. Allred (p. 292)

A new calorimetric type standard RF wattmeter has been developed at the National Bureau of Standards. Its dynamic range extends from 20 milliwatts to 12 watts and overlaps the range of another standard, a 100 μ v to 100 mw thermistor bridge. The frequency range is from dc to 300 mc and approximately 40 minutes is required between readings.

The wattmeter is a transfer standard between accurately known values of dc power and the RF powers to be measured. A complete description of the instrument is given. Analysis of errors indicates a maximum uncertainty of $\pm(0.5$ per cent + 2 mw) in the measured RF power. In comparison measurements with other independent methods, agreements of $\times 0.5$ per cent or better were obtained. This accuracy represents an improvement of one order of magnitude over the best presently available commercial instruments designed for the above power and frequency ranges.

Principles and Application of Radio Interference Measurements—W. E. Pakala and R. M. Showers (p. 297)

Methods for measurement of radio interference from noncommunications equipment have been revised for adoption as an American standard. The improvements have been accomplished as a result of more than 15 years of development in instrumentation and techniques. In instrumentation more rigid specifications are placed on bandwidth characteristics, detector time constants, and methods of calibration, among others. In techniques, more attention is devoted to the effects of environment, including use of impedance stabilization networks and shielded enclosures. Current work directed toward future improvements in standardization is outlined.

Recent Developments in the Field of Microwave Power Measurements at the National Bureau of Standards—Glenn F. Engen (p. 304)

A number of significant advances in the field of microwave power measurement have been recently realized at the Boulder Laboratories of the National Bureau of Standards. These include an improved model of the X-band microwave microcalorimeter, featuring improved accuracy, stability, resolution, and flexibility, which was recently placed in operation, and an improved method of measuring bolometer mount efficiency, using directional couplers, based on the impedance method of Kerns. A preliminary implementation of this second method in waveguide has given agreement of better than 0.5 per cent with measurements by the microcalorimetric method which

in turn have an absolute accuracy within 0.2 per cent.

The problems involved in the intercomparison of bolometer mounts have also been studied, leading to the development of an improved method.

A New Estimating Method of Equivalence Error in the Microwave Microcalorimeter—S. Omori and K. Sakurai (p. 307)

A microwave microcalorimeter has been established at the Electrotechnical Laboratory as a microwave low-level power standard in Japan. The most important problem in this work is the accuracy of this calorimeter as the standard of microwave power. To assure this accuracy, a precise analysis of the instrument has been made. The authors have introduced a new experimental estimating method of RF-dc equivalence error and have assigned a more accurate upper limit of the error. Thus "the effective efficiency" of the standard bolometer mount has been evaluated within the accuracy of 0.5 per cent.

The Theory of Low-Temperature Bolometer Detectors Applied to the Measurement of Low-Level RF Power—D. L. Bix and N. Fuschillo (p. 310)

The sensitivity of bolometer power measurement is ultimately limited by noise, either of the Johnson or temperature fluctuation type. With present bolometers operating at room temperature, accurate measurement is limited to RF powers above one microwatt. Operation of thermally sensitive superconducting bolometer elements at liquid helium temperature has the advantage of considerably lowering the noise level, reducing the specific heat, and permitting the utilization of extremely high temperature coefficients of resistance thus extending the measurement capabilities to much lower limits. Theoretical requirements are discussed for optimum sensitivity of low temperature thermal detectors for the measurement of RF powers down to 10^{-11} watt.

Reproducible Gas Discharge Noise Sources as Possible Microwave Noise Standards—Keith W. Olson (p. 315)

In microwave frequency ranges, the necessity exists for a standard source of noise power at levels of 10–20 db above 290°K. A fundamental property of a given gas discharge is that in the positive column of the discharge a specific effective electron temperature exists. When considered over a particular bandwidth, this fact implies a well-defined noise-output power. Physical variables which could cause fluctuations in this output power can be controlled. S-band noise tubes presently are being manufactured in large quantities to a tube-in-mount specification of 15.25 ± 0.2 db. The short-time stability of tube and test system is ± 0.02 db. The long-time stability of the tubes during life is ± 0.05 db maximum. The state of the art of radiometer type measurements has advanced so that absolute excess noise ratio measurements now can be made to ± 0.1 db or better. In view of this reproducibility and stability of gas-discharge noise tubes, consideration now may be given to the use of a gas discharge tube as a standard of noise power at microwave frequencies.

Recently Developed Microwave Impedance Standards and Methods of Measurement—R. W. Beatty and D. M. Kearns (p. 319)

Recent developments by the Radio Standards Division of the Boulder Laboratories of the National Bureau of Standards have included microwave impedance standards and measurement methods of greatly increased accuracy.

The concept and analysis of the half-round inductive obstacle in rectangular waveguide as an impedance standard by D. M. Kearns has formed the basis for a series of highly accurate fixed impedance standards.

Advanced techniques for measuring VSWR

have permitted measurements of these standards which agree with the calculated values within 0.05 per cent, for VSWR's up to 1.5.

In the measuring techniques described, magnified and squared VSWR responses are used as well as directional couplers with auxiliary tuners.

Precision Microwave Phase Shift Measurements—Milton Magid (p. 321)

A basic method for making precision microwave phase-shift measurements is to compare electrical phase shift with the precise mechanical displacement of a short circuit. Serious errors are possible in the use of short circuits as phase standards unless care is taken to isolate the short circuit from the component under test and to eliminate the error in the short circuit itself.

A system for accomplishing phase shift measurements of high accuracy is treated in this paper. The determination of phase shift is made in a microwave bridge circuit where a basic comparison is effected in a novel manner between electrical phase shift and mechanical displacement of a short circuit. The setup is self-checking since the maximum errors may readily be determined after final assembly. The system may be used in any type of transmission line where satisfactory components are available.

The experimental application of the system to the calibration of two types of 1 inch × ½ inch waveguide phase shifters is treated. Over-all system accuracies of 0.3° or better (better than 0.1 per cent of 360°) are readily attained. A comparison of the calibrations of these phase shifters in the present system and another less accurate system involving the use of a short circuit is given.

A Precision Calibrator for Microwave Demodulators—Joseph Gindsberg (p. 332)

A portable precision generator of modulated signals has been designed for calibrating microwave discriminators and detectors. Fed from an external CW source, a ferrite modulator generated odd- and even-order signal components which are compared in a power bridge, thereby calibrating the modulation level in terms of a precision variable attenuator and directional coupler. Pure phase or amplitude modulation can be selected by adjusting a phase shifter. The three X-band calibrators already built have no measurable calibration disparities and yield linear phase modulation up to 0.6 radian with excellent AM suppression. Calibration charts are given, relating modulation index (or factor), carrier-to-sideband power ratio, rms frequency deviation, and attenuator dial setting.

Evolution of a Standards Laboratory—Charles E. White (p. 339)

The need for standards laboratories in industry is constantly expressed by the National Bureau of Standards. This need must be transformed into an expression of desire on the part of management. The expanse of the field of measurement in guided missile research and development is matched by the requirements of extreme accuracy of test instrumentation. AVCO Research and Advanced Development Division is meeting contract obligations and responsibilities by formation of Standards activities in the electrical, electronic, electro-mechanical, and mechanical measurements fields. The procedures involved in planning, originating, processing, and maintaining such standards are outlined.

The Organization, Administration and Operation of an Industrial Standards Laboratory—James N. Whitaker (p. 345)

Operational requirements and modern manufacturing techniques require close tolerances which can only be realized by the careful reference of all measuring equipment to a common basic standard. These requirements have resulted in the establishment of commercial

standardizing or measurement laboratories, tailored to the needs of the company of which they are a part.

The need for accurate measurements was recognized by the Hughes Aircraft Company, and calibration laboratories were set up early in its existence. As the company expanded, a formal primary standards laboratory was established to maintain the primary standards for all measurements required by the various divisions of the company and to act as the central clearing house for all contacts with the National Bureau of Standards.

The Primary Standards Laboratories of the Hughes Aircraft Company are divided into four groups, permitting the highest degree of specialization consistent with common points of reference to basic standards elements. The organization is such that only these basic elements are normally submitted to the NBS for periodic certifications, all other units being certified internally by comparison with NBS certified elements.

The responsibilities of the Primary Standards Laboratories with respect to the other divisions of the company are set forth in the text. Complete records are maintained for all instruments certified, providing for a valuable pool of information regarding the dependability of instruments.

A Centralized Facility for Electrical and Microwave Calibrations in a Large Company—Lloyd B. Wilson (p. 348)

A Centralized facility for electrical and microwave calibrations at the Sperry Gyroscope Company is provided by its Electrical Measurements Laboratory. In addition to performing calibrations, this department does equipment design and evaluation, sets up calibration programs, and provides state of the art information, consultation, and assistance in any work involving calibrations and physical standards. Standards are classified as either primary reference standards, secondary reference standards, or working standards, depending on their position in the Company hierarchy of standards and calibration levels. A large number of standards is maintained for fundamental parameters over the range from dc to 40,000 mc. Special attention is given to the development and use of precision comparison equipment for making calibrations against the various standards.

Instrument Calibration Program Within the Department of the Navy—BuOrd-BuShips Calibration—Lt. F. L. Roach and M. L. Scroggs (p. 357)

In order to cope with the measurement requirements of the complex modern weapon systems and both shipborne and airborne electronic equipment, the various bureaus concerned in the Department of the Navy have vigorously inaugurated calibration programs.

The two programs presently in operation are Joint BuOrd—BuShips Calibration Program and BuAer Calibration Program.

Both programs are organized in echelons of accuracy with the National Bureau of Standards as the highest accuracy level.

Laboratories are established based on indicated workload and equipped according to measurement requirements of customers.

The calibration programs provide the Navy with the high degree of quality control in measurements required and demanded in this rapidly expanding missile and space age.

Physical and Electrical Standardization Program for the AEC—H. C. Biggs (p. 364)

This paper traces the development of the functional standards program for the AEC Atomic Weapons System. It describes the needs for such a general program and the factors which influenced decisions for both general organization and choices of equipment. The paper emphasizes the technical and administrative guidance provided by NBS personnel.

It touches briefly on audits, intercomparisons with other similar laboratories, and technical publications for the purpose of establishing and maintaining a broad coordinated standards program.

Classification and Nomenclature of Standards of Measurement—A. G. McNish (p. 371)

Standards may be classified and arranged in five categories according to the nature of the standards—"prototype standards," "derived standards," "calibration standards," "standard instruments," and "standard materials." Within each category the classification is by orders which show, for standards of a given kind, the way in which they are intercompared. In a central national laboratory these orders would be: "national standards," "national reference standards," and "working standards." For standardizing laboratories which depend either directly or indirectly on a national laboratory for the accuracy of their standards, the orders are: "laboratory reference standards," "laboratory working standards," and an additional order, "interlaboratory standards," for standards used to effect intercomparisons in case it is undesirable to transport the reference standards.

While the classification indicates to some extent the relative accuracy of the various standards, no simple classification system will exactly specify the accuracy of the various standards. This is best done by a number of each standard giving its limits of reliability.

The results of devices performed by a laboratory should be expressed in terms of the national standards, and the stated accuracy of the calibration should include allowance for uncertainties linking the particular calibrations back to the national standards.

PGI News (p. 378)

Annual Index (Follows p. 379)

Space Electronics and Telemetry

VOL. SET-5, NO. 1, MARCH, 1959

Data Collection Problems in Second Generation Ballistic Missiles—Loren Clifford Sackett (p. 1)

This paper deals with the problem of airborne data collection from advanced ballistic missiles, sometimes called second generation ballistic missiles. Because of improvements in warhead technology, these advanced missiles have a much smaller warhead compartment than previous missiles of the same class. The airborne data collection system has on many previous R&D missiles normally occupied part or all of the warhead space. It becomes obvious therefore, that with a smaller and lighter warhead, more severe weight and space limitations on the airborne instrumentation will tend to exist than heretofore.

Analysis of data requirements reveals that the reduction in the quantity and the complexity of measurements for solid propellant systems over that required on liquid propellant systems is not as great as one might easily assume.

Data collection problems for the guidance and control system involve the transmission at a high repetition rate of pulse or digital data of extreme accuracy. Problem areas exist with analog to digital and digital to digital converters for use in monitoring and totalizing angular shaft motions associated with the guidance systems. Problems are also present with the encoding of serial pulse trains for transmission over existing telemetry systems. Proposed means for handling some of the pulse data collection problems are presented.

A fundamental principle for use in the design of an airborne data acquisition system can be stated as follows: "The nature and amount of the data to be collected should be a major determining factor of the over all system design." Because of the large amount of pulse and digital data expected to be encountered, a telemetry system specially tailored to handle this data is required. Additional systems to handle the high-frequency analog data are also necessary.

Studies show that elimination of the commonly used 50 to 100 watt airborne RF power amplifiers system may be possible by taking advantage of the recent improvements in the sensitivity of telemetry ground stations. Large saving in weight and space is made possible by the elimination of commonly used airborne RF power amplifiers. This is illustrated by graphical charts.

A summary is presented of the advantages and problems of recording high speed pulse or digital data on the ground with wideband video tape recorders.

An Automatic Data Handling System at High Speed and High Capacity—R. L. Sink, C. E. Pettingall, and B. N. Posthill (p. 8)

The high-speed, high-capacity automatic data handling system has been in the process of design and development during the preceding eighteen months. This paper is devoted to the salient system characteristics which now exist in the final system design. Certain of the operating characteristics which have been established as a result of several months of operation are reported. Emphasis is placed on the systems design features intended to make the over-all system appropriate for a high order of automatic operation.

Instrumentation and Range Safety System for Vandenberg Air Force Base—R. E. Hadady (p. 14)

This paper is a survey of the Instrumentation and Range Safety System (IRSS) now under design for installation at Vandenberg Air Force Base. The Vandenberg Air Force Base will be used for launching intercontinental and intermediate range ballistic missiles, such as the Atlas, Titan, and Thor. These missiles, their launching pads and other associated information, are identified throughout the rest of the paper by the Air Force designation number. The number 65 identifies the Atlas; 68, the Titan; and 75, the Thor.

The Vandenberg Air Force Base has been established in order to train Air Force crews in the launching of ballistic-type missiles and to establish an initial operating capability.

Research and development will be provided only to the extent necessary to achieve the program. The range area is limited to the Vandenberg Air Force Base and the immediate launching facility, and does not include the area normally referred to as "downrange."

The Instrumentation System's basic function is to provide data substantiating the proper operation and reliability of the missile. The Range Safety System's basic function is to protect the life and property on and near Vandenberg Air Force Base.

Some Observations on the Challenging Aspects of the Telemetry Field—L. Eugene Root (p. 28)

Basic Theory of Locked Oscillators in Tracking FM Signals—Glenn W. Preston (p. 30)

Frequency Domain Applications to Data Processing—Marcel A. Martin (p. 33)

Filtering a set of $(2N + 1)$ sampled data consists in taking a weighted average of these data. The operation can be compared to that of a linear transducer, characterized by its frequency response. The only filters discussed in this paper are those for which the weights are either symmetric or skew symmetric about the center weight. With these types of filters, it is possible to design low low-pass filters to determine means and trends, low-pass filters to smooth data, sampling filters for frequency analysis, differentiators, integrators, etc. It is shown that least square polynomial fitting and Fourier analysis of sampled data are particular cases of filtering. Power of filtering techniques and precautions to take in their use for data processing are discussed.

A Digital System for Position Determination—Dan C. Ross (p. 42)

A prime requirement in improving the nation's air traffic control system is the provision of frequent, accurate, and identified position information on all aircraft in the system. This presentation will describe a system which satisfies these requirements for all aircraft aloft on a single pair of radio frequency channels without intra-system interference. Flight tests of the radio communication portions of the system have been successfully completed.

Some Properties of the Gravitation Field and Their Possible Application to Space Navigation—J. C. Crowley, S. S. Kolodkin, and A. M. Schneider (p. 47)

This paper describes the principle of operation of a new inertial instrument and the elements of a technique for making measurements using this instrument. This device measures certain spatial properties of the gravitational field, from which the direction of the vertical and altitude with respect to a nearby heavenly body can be obtained. An elementary treatment of the mathematics establishes the theoretical foundations for the instrument and also provides an indication of the effects of various classes of disturbing inputs. Some of the problems associated with the reduction to practice of the basic transducer are considered, and gross feasibility is established.

Proceedings of the 1958 National Symposium on Telemetry (p. 54)

Abstracts and References

Compiled by the Radio Research Organization of the Department of Scientific and Industrial Research, London, England, and Published by Arrangement with that Department and the *Electronic and Radio Engineer*, incorporating *Wireless Engineer*, London, England

NOTE: The Institute of Radio Engineers does not have available copies of the publications mentioned in these pages, nor does it have reprints of the articles abstracted. Correspondence regarding these articles and requests for their procurement should be addressed to the individual publications, not to the IRE.

Acoustics and Audio Frequencies.....	1030
Antennas and Transmission Lines.....	1031
Automatic Computers.....	1032
Circuits and Circuit Elements.....	1032
General Physics.....	1034
Geophysical and Terrestrial Phenomena.....	1035
Location and Aids to Navigation.....	1036
Materials and Subsidiary Techniques.....	1036
Mathematics.....	1039
Measurements and Test Gear.....	1039
Other Applications of Radio and Electronics.....	1040
Propagation of Waves.....	1040
Reception.....	1041
Stations and Communication Systems.....	1042
Subsidiary Apparatus.....	1042
Television and Phototelegraphy.....	1042
Transmission.....	1043
Tubes and Thermionics.....	1043
Miscellaneous.....	1044

The number in heavy type at the upper left of each Abstract is its Universal Decimal Classification number. The number in heavy type at the top right is the serial number of the Abstract. DC numbers marked with a dagger (†) must be regarded as provisional.

ACOUSTICS AND AUDIO FREQUENCIES

534.1 1039

Variation of the Natural Frequencies of Membranes and Resonators with Additional Loads—Yu. N. Dnestrovskii. (*Akust. Z.*, vol. 4, pp. 244–252; July/September, 1958.)

534.2 1040

The Sound Field Generated by a Point Source in a Layer Lying on a Half-Space—Yu. L. Gazarian. (*Akust. Z.*, vol. 4, pp. 233–238; July/September, 1958.) An investigation of the frequency dependence of the "lateral-wave" field.

534.2 1041

The Effective Dynamic Parameters for Sound Propagation in Inhomogeneous Media—L. M. Khaikovich and L. A. Khalfin. (*Akust. Z.*, vol. 4, pp. 275–281; July/September, 1958.) Equations are derived for calculating the "effective" parameters of a two-component medium in which the inhomogeneities are distributed in the form of a regular lattice.

534.2 1042

Scattering of Sound Waves in Irregular Waveguides—A. D. Lapin. (*Akust. Z.*, vol. 4, pp. 267–274; July/September, 1958.) A method is proposed for calculating the scattering in waveguides for the case in which the scattered field is not small in comparison with the incident field. The case of small fluctuations in the parameters of the medium inside the waveguide and the case of wall roughness in the waveguide are examined. The scattered field is found in the form of superpositions of the normal waves which would occur in the waveguide if the irregularities were absent.

534.2-14:534.6 1043

Reverberation Tank Method for the Study of Sound Absorption in the Sea—V. P. Glotov.

The Index to the Abstracts and References published in the PROC. IRE from February, 1958 through January, 1959 is published by the PROC. IRE, May, 1959, Part II. It is also published by *Electronic and Radio Engineer*, incorporating *Wireless Engineer*, and included in the March, 1959 issue of that journal. Included with the Index is a selected list of journals scanned for abstracting with publishers' addresses.

(*Akust. Z.*, vol. 4, pp. 239–243; July/September, 1958.) The tank has a capacity of 0.5 m³ and was designed for ocean measurements at depths not exceeding 50 m. Results of measurements at 20 m depth and frequencies 15–80 kc are given.

534.23 1044

On the Acoustic Field [caused] by a Vibrating Source Arbitrarily Distributed on a Ribbon Plate: Part I—N. Kawai. (*J. Phys. Soc. Japan*, vol. 13, pp. 1374–1384; November, 1958.) Two-dimensional rigorous solutions of the field in air are given for sources of infinitely thin and infinitely thin and infinitely long ribbons.

534.232 1045

Calculations for a Piston-Type Piezoelectric Radiator when Internal Losses are Neglected—A. A. Anan'eva. (*Akust. Z.*, vol. 4, pp. 223–232; July/September, 1958.) Analysis of flat piston-type quartz and BaTiO₃ transducers operating in water. The importance of choosing materials with a high piezoelectric factor, a high density and large modulus of elasticity is stressed.

534.241 1046

The Sound Echo Reflected from a Sphere under Pulse Conditions—M. Federici. (*Ricerca sci.*, vol. 28, pp. 1659–1667; August, 1958.)

534.26-8-14 1047

Contribution to the Study of the Scattering of an Ultrasonic Plane Wave by Surfaces with a Periodic Structure—H. Blons. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 50–52; July 7, 1958.) Results of experiments made in water at 625–2500 kc on reflections from a periodic surface show good agreement with diffraction theory.

534.26-8-14:548.73 1048

The Diffraction of Ultrasonic Waves by Multi Layer Arrays of Rods—P. de Villers. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 52–54; July 7, 1958.) Experiments were made in water at a wavelength of 0.57 mm and with rod diameter 0.3 mm, in order to confirm Bragg's laws for the diffraction of X-rays by crystal lattices.

534.522.1 1049

Diffraction of Light by Ultrasonic Waves—Oblique Incidences—S. Parthasarathy and C. B. Tipnis. (*Nature (London)*, vol. 182, pp. 1083–1084; October 18, 1958.) Experimental confirmation of a closed-form expression for diffraction intensity.

534.522.1 1050

Investigation of Progressive Ultrasonic Waves by Light Refraction—M. A. Breazele

and E. A. Hiedemann. (*J. Acoust. Soc. Am.*, vol. 30, pp. 751–756; August, 1958.) Adaptations of the method used by Loeber and Hiedemann (1941 of 1956) for the study of stationary waves are described, for the detection of wave-form distortion due to finite-amplitude effects in liquids.

534.522.1 1051

Study of the Intensity Distribution of the Light Diffracted by Ultrasonic Waves—R. B. Miller and E. A. Hiedemann. (*J. Acoust. Soc. Am.*, vol. 30, pp. 1042–1046; November, 1958.) Differences between experimental results and theory are discussed.

534.7 1052

Signal Detection as a Function of Frequency Ensemble: Part I—F. A. Vaniar. (*J. Acoust. Soc. Am.*, vol. 30, pp. 1020–1024; November, 1958.) Results of an experimental investigation of the detection of multiple-tone signals in noise are compared with predictions based on mathematical models.

534.75 1053

Intensity and Duration of Noise Exposure and Temporary Threshold Shifts—W. Spieth and W. J. Trittipoe. (*J. Acoust. Soc. Am.*, vol. 30, pp. 710–713; August, 1958.) See also 682 of March.

534.75 1054

Residual Effects of Low Noise Levels on the Temporary Threshold Shift—W. J. Trittipoe. (*J. Acoust. Soc. Am.*, vol. 30, pp. 1017–1019; November, 1958.)

534.76 1055

On the Mechanism of Binaural Fusion—E. E. David, Jr., N. Guttman and W. A. van Bergeijk. (*J. Acoust. Soc. Am.*, vol. 30, pp. 801–802; August, 1958.) The relation between interaural intensity difference and interaural time difference is obtained by a pulse test method. The interpretation of test results is discussed.

534.76 1056

An Artificial Stereophonic Effect Obtained from a Single Audio Signal—M. R. Schroeder. (*J. Audio Eng. Soc.*, vol. 6, pp. 74–79; April, 1958. Discussion.) Report of experimental investigations of the stereophonic effect achieved by feeding a signal to both ears of a listener once directly in phase, and a second time, with a delay of 50–150 msec, in antiphase.

534.78:39 1057

Phonetic Pattern Recognition Vocoder for Narrow-Band Speech Transmission—H. Dudley. (*J. Acoust. Soc. Am.*, vol. 30, pp. 733–739;

- August, 1958.) A phonetic-pattern recognition circuit (1061 below) is combined with a synthesizer to form a phonetic vocoder.
- 534.781 1058
Modern Instruments and Methods for Acoustic Studies of Speech—G. Fant. (*Acta Polytech. Scand.*, No. 246, vol. 1, p. 81; 1958.)
- 534.782 1059
Segmentation Techniques in Speech Synthesis—G. E. Peterson, W. S. Wang and E. Sivertsen. (*J. Acoust. Soc. Am.*, vol. 30, pp. 739-742; August, 1958.) In the method described, discrete segments of recorded utterances are joined together to produce continuous speech. Techniques for obtaining the segments are discussed.
- 534.782 1060
Segment Inventory for Speech Synthesis—W. S. Wang and G. E. Peterson. (*J. Acoust. Soc. Am.*, vol. 30, pp. 743-746; August, 1958.) See 1059 above.
- 534.784:621.395 1061
Automatic Recognition of Phonetic Patterns in Speech—H. Dudley and S. Balashek. (*J. Acoust. Soc. Am.*, vol. 30, pp. 721-732; August, 1958.) The performance of a ten-word digit recognizer for voice dialing of telephone numbers is discussed. A model of a phonetic pattern recognizer based on a detailed analysis of the power frequency spectrum is described; this provides almost perfect digit recognition when set for a particular voice.
- 534.79 1062
Advantages of the Discriminability Criterion for a Loudness Scale—W. R. Garner. (*J. Acoust. Soc. Am.*, vol. 30, pp. 1005-1012; November, 1958.)
- 534.846 1063
Acoustic Criteria of Outstanding Old and New Concert Halls—F. Winckel. (*Frequenz*, vol. 12, pp. 50-59; February, 1958.) Comparison and discussion of architectural and acoustic features of pre and post-war auditoria and concert halls.
- 534.85/.86:534.76 1064
Stereophonic Sound with Two Tracks, Three Channels by means of a Phantom Circuit (2PH3)—P. W. Klipsch. (*J. Audio Eng. Soc.*, vol. 6, pp. 118-123; April, 1958.) Practical details are given for combining the signals of each channel and using the resultant signal to drive a third central loudspeaker.
- 534.85/.86:534.76 1065
Stereo-reverberation—R. Vermeulen. (*J. Audio Eng. Soc.*, vol. 6, pp. 124-130; April, 1958.) The effect can be achieved using a single-channel system feeding several loudspeakers, distributed around an auditorium, with delayed signals repeated many times at decreasing levels.
- 534.88 1066
Correlation with Similar Uniform Collinear Arrays—M. J. Jacobson. (*J. Acoust. Soc. Am.*, vol. 30, pp. 1030-1034; November, 1958.) An acoustic multiple-receiver correlation system is considered. See also 1959 of 1958.
- ANTENNAS AND TRANSMISSION LINES**
- 621.372.2 1067
Approximate Models for Transmission Lines and their Errors—R. R. Vierhout. (*Electronic Eng.*, vol. 31, pp. 94-95; February, 1959.)
- 621.372.2:538.569.2/.3 1068
The Absorption of Electromagnetic Waves in Absorbers and Lines with Sectionally Homogeneous Distributed Conductance—K. L.
- Lenz and O. Zinke. (*Z. angew. Phys.*, vol. 9, pp. 481-489; October, 1957.)
- 621.372.2:621.317.34 1069
Novel Method for Measuring Impedances on Surface-Wave Transmission Lines—G. Schulten. (*Proc. IRE*, vol. 47, pp. 76-77; January, 1959.)
- 621.372.22:621.372.51 1070
H.F. Exponential-Line Transformers—S. G. Young. (*Electronic Radio Eng.*, vol. 36, pp. 40-44; February, 1959.) The design and construction of four-wire tapered transmission lines to match balanced impedances are described.
- 621.372.8:621.3.018.75 1071
Pulse Waveform Degradation due to Dispersion in Waveguide—R. S. Elliott. (*IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-5, pp. 254-257; October, 1957.)
- 621.372.821 1072
Optimum Impedance and Dimensions for Strip Transmission Line—K. S. Packard. (*IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-5, pp. 244-247; October, 1957.)
- 621.372.824:621.372.831 1073
Low-Reflection Discontinuities in Diameter of Coaxial Lines—G. W. Epprecht. (*Tech. Mitt. PTT*, vol. 36, pp. 97-103; March 1, 1958.) Test results obtained on 50- Ω coaxial-waveguide junctions with changes in diameter ratio from 1.15:1 to 3:1 are given as a function of the distance between the transitions in the inner and outer conductors.
- 621.372.825 1074
Serrated Waveguide—(*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 270-283; July, 1957. Abstract, *Proc. IRE*, vol. 45, p. 1759; December, 1957.)
Part 1—Theory.—R. S. Elliott (pp. 270-275).
Part 2—Experiment.—K. C. Kelly and R. S. Elliott (pp. 276-283).
- 621.372.825 1075
Wave Propagation in a Diaphragm-Type and a Corrugated Waveguide—G. Piefke. (*Arch. elekt. Übertragung*, vol. 12, pp. 26-34; January, 1958.) The influence of diaphragms and periodic corrugations in a circular loss-free waveguide is investigated theoretically. See also 2625 of 1958.
- 621.372.826:538.614 1076
Surface Waves on the Boundary of a Gyrotropic Medium—M. A. Gintsburg. (*Zh. Eksp. Teor. Fiz.*, vol. 34, pp. 1635-1637; June, 1958.) Brief mathematical analysis for the case of a) a gyrotropic plate between two isotropic media and b) a channel between two gyrotropic media, with reference to the wave-retardation properties of the system. See 2564 and 2927 of 1954.
- 621.372.83 1077
Frequency Compensation for Simple Stepped Waveguide Transforming Sections—D. Wray. (*Electronic Eng.*, vol. 31, pp. 76-79; February, 1958.) Sections can be "broadbanded" by varying both broad and narrow waveguide dimensions at each step; this gives a better performance than sections designed on a single mid-band frequency.
- 621.372.83 1078
A Variable-Ratio Microwave Power Divider and Multiplexer—W. L. Teeter and K. R. Bushore. (*IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-5, pp. 227-229; October, 1957.)
- 621.372.832.8 1079
High-Power Ferrite Circulators—P. A. Rizzi. (*IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-5, pp. 230-237; October, 1957.)
- 621.372.837 1080
Waveguide Switches and Branching Networks—J. W. Sutherland (*Electronic Eng.*, vol. 31, pp. 64-68; February, 1959.) Parrel, shutter, gas-discharge and ferrite switches are compared. Branching tee properties and applications are discussed and a multiple branching network using hybrids is described.
- 621.372.852.21 1081
Broad-Band Quarter-Wave Plates—W. P. Ayres. (*IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-5, pp. 258-261; October, 1957.) Analytical expressions are given for the propagation constant for the two orthogonal dominant modes in a square waveguide loaded with a centered dielectric slab.
- 621.372.855 1082
Waveguide Termination with Commercial Film-Type Resistors—U. v. Kienlin and A. Kürzl. (*Nachrichtentech. Z.*, vol. 11, pp. 138-141; March, 1958.) The 20-w termination described consists of an arrangement of 20 1-k Ω , 1-w resistors parallel to the electric field. Methods of compensation and advantages of the system are discussed.
- 621.372.86 1083
End Sections for Matching Waveguide Directional Couplers of Periodic Structure—I. Lucas. (*Arch. elekt. Übertragung*, vol. 12, pp. 91-96; February, 1958.) Rules for determining the matching elements required at either end of the coupler are given. The bandwidth for a given matching condition is calculated for one and two end sections. Measurements on a directional coupler with power evenly shared between outlets indicate that one matching element at either end would be adequate.
- 621.396.67+621.396.11 1084
International Colloquium on Current Problems in Radio Wave Propagation, Paris, September 17-21, 1956. (See 1333.)
- 621.396.67 1085
The Current Distribution and Input Impedance of Cylindrical Antennas—E. V. Bohn. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 343-348; October, 1957.)
- 621.396.67-418 1086
Currents on Strip Aerials—T. B. A. Senior. (*Electronic Radio Eng.*, vol. 36, pp. 60-63; February, 1959.) An exact expression is obtained for the longitudinal distribution of current excited on a perfectly conducting strip by a normally incident plane wave. Computations are made for quarter- and half-wave antennas. Deduced total current distributions do not agree with those for a thin wire, but give a better approximation than the cosine variation.
- 621.396.67.012 1087
Evaluating Aerials Performance—L. A. Moxon. (*Wireless World*, vol. 65, pp. 59-65 and 139-144; February and March, 1959.) Various aspects of dipole, beam array and long-wire antenna design are discussed nonmathematically, using a simplified method for calculating gain and radiation resistance. Problems associated with transmission-line losses and with the maintenance of a high signal/noise ratio at the receiving antenna are considered.
- 621.396.67.029.62/.63:621.397.7 1087
Post-Installation Performance Tests of U.H.F. Television Broadcasting Antennas—D. W. Peterson. (*CA Rev.*, vol. 19, pp. 656-672; December, 1958.) The precautions neces-

- sary to reduce errors introduced by reflections from the earth's surface, obstacles, clutter and the earth's curvature, are discussed.
- 621.396.674.3** 1089
The Conductance of Dipoles of Finite Length and Thickness—K. Fränz, P. A. Mann and J. Vocoides. (*Arch. elekt. Übertragung*, vol. 12, pp. 49–53; February, 1958.) For a given frequency a family of dipoles is obtained which have the same conductance. The shape and conductance of a dipole are determined by evaluating a number of integrals and solving a first-order differential equation.
- 621.396.674.3.01** 1090
Formulation of the Antenna Circuit Theory based on the Variational Method—K. Furutsu. (*J. Radio Res. Labs., Japan*, vol. 5, pp. 315–333; October, 1958.) Radiation power and field strength are expressed as functions of current distribution in a dipole antenna system. These are stationary with respect to small variations in the distribution and are independent of the magnitude of the distribution.
- 621.396.677:629.19** 1091
Limitations of Satellite Antennas using Spherical Arrays—K. S. Kelleher and J. P. Shelton. (*Proc. IRE*, vol. 47, pp. 74–75; January, 1959.)
- 621.396.677.3:621-526** 1092
Servo Phase Control Shapes Antenna Pattern—E. W. Markow. (*Electronics*, vol. 32, pp. 50–52; January 2, 1959.) A design of phase-stabilized UHF amplifier and its application to the directivity control of a multi-element stationary array is described. The angular phase error between input and output is less than 2°.
- 621.396.677.71** 1093
Convergent Representations for the Radiation Fields from Slots in Large Circular Cylinders—L. L. Bailin and R. J. Spellmire. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 374–383; October, 1957.)
- 621.396.677.71:621.397.61** 1094
The Travelling-Wave V.H.F. Television Transmitting Antenna—M. S. Siukola. (*IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS*, vol. BTR-3, pp. 49–58; October, 1957. Abstract, *Proc. IRE*, vol. 46, p. 383; January, 1958.)
- 621.396.677.75** 1095
A Technique for Controlling the Radiation from Dielectric Rod Waveguides—J. W. Duncan and R. H. DuHamel. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 284–289; July, 1957. Abstract, *Proc. IRE*, vol. 45, p. 1759; December, 1957.)
- 621.396.677.8.012.12** 1096
On the Simulation of Fraunhofer Radiation Patterns in the Fresnel Region—D. K. Cheng. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 399–402; October, 1957.)
- 621.396.677.83** 1097
The Ideal Plane Reflector—M. Kummer. (*NachrTech.*, vol. 8, pp. 61–65; February, 1958.) The changes of the radiation pattern as a function of reflector distance are calculated for an infinitely large plane with a $\lambda/2$ dipole. The limitations in practical applications are considered.
- 621.396.677.832** 1098
A Circularly Polarized Corner-Reflector Antenna—O. M. Woodward, Jr. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 290–297; July, 1957. Abstract, *Proc. IRE*, vol. 45, p. 1759; December, 1957.)
- 621.396.677.832** 1099
Corner-Reflector Antennas with Arbitrary Dipole Orientation and Apex Angle—R. W. Klopfenstein. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 297–305; July, 1957. Abstract, *Proc. IRE*, vol. 45, pp. 1759–1760; December, 1957.)
- 621.396.677.85** 1100
Lens-Antenna Design—P. Foldes and L. Solymar. (*Electronic Radio Eng.*, vol. 36, pp. 73–75; February, 1959.) A simple geometrical method is described for the design of a lens which will give a prescribed amplitude and phase distribution in an aperture when the antenna is fed by a given primary source. The graphical method yields a polygon as the approximation to the lens contour.
- 621.396.677.85** 1101
Scanning Lens Design for Minimum Mean-Square Phase Error—E. K. Proctor and M. H. Rees. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-5, pp. 348–355; October, 1957.)
- AUTOMATIC COMPUTERS**
- 681.142** 1102
An Electronic Ratio Calculator—A. D. Boronkay. (*Electronic Eng.*, vol. 31, pp. 32–34; January, 1959.) Use is made of the principle that the amplitude ratio of two signals is approximately linearly related to the phase angle of the sum.
- 681.142:621.318.5:537.312.62** 1103
Digital-Analogue Conversion with Cryotrons—L. K. Wanlass and L. O. Hill. (*Proc. IRE*, vol. 47, pp. 100–101; January, 1959.)
- 681.142:621.318.57:538.221** 1104
A High-Speed Logic System using Magnetic Elements and Connecting Wire Only—H. D. Crane. (*Proc. IRE*, vol. 47, pp. 63–73; January, 1959.) Square-loop magnetic elements having multiaperture or multipath variations of the toroid shape are used to provide unilateral information-flow properties. Information-bit rates above $\frac{1}{2}$ mc have been obtained with commercial ferrite material. Advantages include simplicity of wiring and the inherent nondestructive read-out properties of the system.
- 681.142:621.383** 1105
Automatic Data Reduction of Spot Diagram Information—W. E. Goetz and N. J. Woodland. (*J. Opt. Soc. Amer.*, vol. 48, pp. 965–968; December, 1958.) The redistribution of energy falling on a photocathode during scanning by a mechanical image dissector is evaluated.
- CIRCUITS AND CIRCUIT ELEMENTS**
- 621.316.86:621.315.56** 1106
Measurements of Nonlinearity in Cracked-Carbon Resistors—G. H. Millard. (*Proc. IRE*, vol. 106, pp. 31–34; January, 1959.) Frequencies used were 1.06 kc, 3.3 kc, and 90 mc. Resistors measured ranged from 50–100 Ω , 50–100 Ω , $\frac{1}{8}$ w—55 w. The degree of nonlinearity would be negligible for most purposes.
- 621.318.57:621.318.134** 1107
Millimicrosecond Microwave Ferrite Modulator—A. H. Solomon and F. Sterzer. (*Proc. IRE*, vol. 47, pp. 98–100; January, 1959.) Description of a switching device similar to the resonance-type isolator described by Enander (17 of 1957), but with ferrite rings placed inside the helix and with a single straight wire along the axis carrying the modulating current.
- 621.318.57:621.383** 1108
Photoelectronic Circuit Applications—S. K. Ghandhi. (*Proc. IRE*, vol. 47, pp. 4–11; January, 1959.) The characteristics of switching circuits using combinations of electroluminescent cells and photoconductors are discussed. Applications are described in computer operations where the highest speed is not essential.
- 621.318.57:621.383.4:621.314.7** 1109
A Temperature-Stabilized Phototransistor Relay Circuit—J. C. Anderson and T. Winer. (*Electronic Eng.*, vol. 31, pp. 36–37; January, 1959.) A phototransistor and an ordinary junction transistor are used in a "long-tailed pair" arrangement to reduce the variation in dark current.
- 621.318.57:621.387** 1110
A Reversible Dekatron Counter—D. L. A. Barber. (*Electronic Eng.*, vol. 31, pp. 42–43; January, 1959.)
- 621.319.45** 1111
Recent Advances in the Solid-State Electrolytic Capacitor—A. V. Fraioli. (*IRE TRANS. ON COMPONENT PARTS*, vol. CP-5, pp. 72–75; June, 1958. Abstract, *Proc. IRE*, vol. 46, p. 1533; August, 1958.)
- 621.372.413:621.372.54.029.6** 1112
Design of Aperture-Coupled Filters—F. Shurer. (*IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-5, pp. 238–243; October, 1957.)
- 621.372.5** 1113
Network Synthesis—J. T. Allanson. (*Electronic Radio Eng.*, vol. 36, pp. 66–69; February, 1959.) A method is outlined for the synthesis of balanced asymmetrical RC networks, which may have a lower pad loss than the corresponding symmetrical types.
- 621.372.5.029.6:621.317.341** 1114
Definition of Lossy Quadrupoles by Voltage Mode Displacements in the Microwave Range—F. Gemmel. (*Arch. elekt. Übertragung*, vol. 12, pp. 76–80; February, 1958.) Equations are derived giving the voltage swr and the position of the voltage node on the input line of a lossy quadrupole in terms of the position of the short-circuit on its output line.
- 621.372.51 (083.57):621.314.7** 1115
Determination of Transistor Performance Characteristics at V.H.F.—G. E. Theriault and H. M. Wasson. (*IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS*, vol. BTR-3, pp. 40–48; June, 1957.) An "immittance chart," a development of the Smith Chart, is described, and its application in the design of adjustable matching networks for transistor measurements is illustrated.
- 621.372.54** 1116
General Physical Relations in Ladder-Type Filters—T. Laurent. (*Arch. elekt. Übertragung*, vol. 12, pp. 1–8; January, 1958.) The improvement of filter characteristics by mismatch conditions between the ladder sections is considered and a design method is outlined. See also 1931 of 1953.
- 621.372.54** 1117
Minimum-Insertion-Loss Filters—E. G. Fubini and E. A. Guillemin. (*Proc. IRE*, vol. 47, pp. 37–41; January, 1959.) The generalized criterion for Butterworth and Chebyshev filters of arbitrary bandwidth is chosen for minimum loss in the center of the band.
- 621.372.54** 1118
The Design of Two-Section Symmetrical Zobel Filters for Chebyshev Insertion Loss—W. N. Tuttle. (*Proc. IRE*, vol. 47, pp. 29–36; January, 1959.)
- 621.372.54** 1119
A Voltage-Controlled Continuously Variable Low-Pass Filter—A. J. Seyler and A. Korpel.

- (*Electronic Eng.*, vol. 31, pp. 16-22; January, 1959.) The filter has a 6 db/octave cut-off characteristic and the cut-off frequency, at -3 db, is continuously variable from 4.5 kc to 5 mc by a control-voltage variation of 85 v.
- 621.372.54** 1120
The Exact Design of Two Types of Single-Crystal, Wide-Band Crystal Filters—T. R. O'Meara. (*Trans. Inst. Radio Engrs.*, vol. CP-5, pp. 46-52; March, 1958. Abstract, *PROC. IRE*, vol. 46, Part 1, p. 931; May, 1958.)
- 621.372.54** 1121
The Symmetrical Transfer Characteristics of the Narrow-Bandwidth Four-Crystal Lattice Filter—T. R. O'Meara. (*IRE TRANS. ON COMPONENT PARTS*, vol. CP-5, pp. 84-89; June, 1958. Abstract, *PROC. IRE*, vol. 46, p. 1553; August, 1958.)
- 621.372.543.3** 1122
The Bifilar-T Circuit—T. Roddam. (*Wireless World*, vol. 65, pp. 66-71 and 145-148; February/March, 1959.) The circuit is critically examined, using lattice network theory. See also 2658 of 1958 (Hendry and McIntosh).
- 621.372.56.029.64:621.317.7.089.6** 1123
The Calibration of Microwave Attenuators by an Absolute Method—E. Laverick. (*IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES*, vol. MTT-5, pp. 250-254.) A bridge method is described and results of measurements at 3.2 cm λ are given.
- 621.373.4** 1124
Alternatives to the Wien Bridge—J. F. Young. (*Wireless World*, vol. 65, pp. 92-95; February, 1959.) Oscillator circuits with two-gang control, having higher output voltages and greater selectivity than the basic Wien bridge circuit are described.
- 621.373.421** 1125
A Wide-Band Voltage-Controlled Swept-Frequency RC Oscillator—R. S. Sidorowicz. (*A.T.E. J.*, vol. 14, pp. 88-112; April, 1958.) Oscillators based on the principle of parallel network tuning [see e.g. 2239 of 1955 (Stewart)] are examined. New equipment using this principle is described in detail. Frequency variations can be restricted to any limits in the range 20 cps-3 kc and the maximum duration of the sweep cycle can be varied from 4 to 20 min.
- 621.373.421.11** 1126
Incremental Frequency Control of RC Oscillators—D. L. A. Smith. (*Electronic Eng.*, vol. 31, pp. 38-39; January, 1959.) An auxiliary phase-shifting network is introduced to provide a continuously variable incremental frequency control. A practical method is described for systematically correcting the error which exists for large increments.
- 621.373.421.13:529.786** 1127
Construction and Properties of High-Quality Oscillator Crystals—G. Becker. (*Arch. elekt. Übertragung*, vol. 12, pp. 15-25; January, 1958.) Details are given of the shape, processing and mounting of crystals for use in quartz-clock standards. The influence of various design parameters on crystal characteristics and performance is discussed and a method of frequency control is described. See also 3468 of 1956 (Scheibe et al.).
- 621.373.43:621.314.7** 1128
A Monostable Circuit using a Transistor—H. Stinton. (*Electronic Eng.*, vol. 31, pp. 80-81; February, 1959.) A pulse generating circuit is described using a simple pulse transformer and a transistor. Pulse duration is adjustable over a wide range, greatly exceeding that usually associated with the pulse transformer used.
- 621.373.52:621.372.412** 1129
Crystal-Controlled Transistor Oscillators—H. Avender and A. Ludloff. (*Elektronische Rundschau*, vol. 12, pp. 75-80; March, 1958.) The design of parallel-resonant circuits is considered.
- 621.374.3:621.387** 1130
The Use of Dekatrons for Pulse Distribution—C. H. Stearman. (*Electronic Eng.*, vol. 31, pp. 69-71; February, 1959.) Specific methods are discussed for increasing the number of output lines available from cold-cathode selector tubes used for distributing pulses in a fixed time sequence.
- 621.374.32:621.314.7** 1131
A 1-mc Transistor Decade Counter—C. G. Bradshaw. (*Electronic Eng.*, vol. 31, pp. 96-97; February, 1959.)
- 621.374.4** 1132
Investigations on Harmonic Frequency Dividers to Determine the Prerequisite Conditions for Achieving High Division Ratios and Large Locking Ranges at High Frequencies—K. Schlichting. (*Z. angew. Phys.*, vol. 9, pp. 458-464; September, 1957.)
- 621.374.43:621.314.7** 1133
A Regenerative Modulator Frequency Divider using Transistors—F. Butler. (*Electronic Eng.*, vol. 31, pp. 72-75; February, 1959.)
- 621.374.5.538.652:621.396.96** 1134
Magnetostrictive Delay Line for Video Signals—G. I. Cohn, L. C. Peach, M. Epstein, H. O. Sorensen and D. P. Kanellakos. (*IRE TRANS. ON COMPONENT PARTS*, vol. CP-5, pp. 53-59; March, 1958. Abstract, *PROC. IRE*, vol. 46, Part 1, p. 931; May, 1958.)
- 621.375.018.75** 1135
Pulse Amplifier with Submillimicrosecond Rise Time—F. Sterzer. (*Rev. Sci. Instr.*, vol. 29, pp. 1133-1135; December, 1958.) The rise time is about 0.7 μ s. The dc pulses modulate a 2750-mc carrier, and the RF pulses are amplified by a traveling-wave valve. Demodulation then gives the amplified dc pulses. The maximum power gain is 20 db, and the maximum output voltage 2.1 v.
- 621.375.018.756:537.525.6** 1136
A Low-Noise Wide-Band Amplifier for the Investigation of Electron Avalanches—K. J. Schmidt-Tiedemann. (*Z. angew. Phys.*, vol. 9, pp. 454-458; September, 1957.)
- 621.375.024:621.317.321.027.21** 1137
Low-Level D.C. Amplifier with Whole-Loop Feedback—P. C. Hoell. (*Rev. Sci. Instr.*, vol. 29, pp. 1120-1124; December, 1958.) A 60-cps vibrating-contact modulator and whole-loop negative feedback are used, and a stabilized gain of 10^7 from dc to 10 cps, is obtained. The zero drift is less than $\pm 10^{-9}$ v per hour and less than 2×10^{-9} v per day.
- 621.375.2.012.6** 1138
Amplitude and Phase Response of an Amplifier Operating on an Interrupted Cycle—A. Sabbatini. (*Note Recensioni Notiz.*, vol. 7, pp. 277-287; May/June, 1958.) Analysis of class-C amplifier characteristics.
- 621.375.4** 1139
Transistor Amplifiers: Common Base versus Common Emitter—R. F. Purton. (*A.T.E. J.*, vol. 14, pp. 157-163; April, 1958.) The gain stability and frequency response of the two circuits are compared theoretically and experimentally. If a generator with a relatively low impedance is used, these characteristics need be no worse for the common-emitter than for the common-base amplifier.
- 621.375.4** 1140
Transistor Bias Design from Thermal Incremental Properties—L. M. Vallee. (*Electronic Eng.*, vol. 31, pp. 88-93; February, 1959.) Explicit formulae for design of single-stage bias networks are given. Two thermal parameters need to be known, one related to DI_{∞}/dI and the other to the rate of change of input current. Applications and experimental verification of the theory are included. See also 2078 of 1957.
- 621.375.4.01:512.831** 1141
An Introduction to Matrices and their Use in Transistor Circuit Analysis—J. S. Bell and K. Brewster. (*Electronic Eng.*, vol. 31, pp. 98-102; February, 1959.) Brief details are given for analyzing both simple and complex circuits using transistors, with emphasis on application to practical circuits.
- 621.375.432** 1142
Low-Noise Transistor Amplifier Stages with Negative Feedback—E. R. Hauri. (*Tech. Mitt. PTT*, vol. 36, pp. 103-108; March 1, 1958.) Theoretical investigation of the grounded-emitter connection with feedback resistance in the emitter circuit. The relation of source impedance to optimum noise figure is determined.
- 621.375.9:538.569.4.029.64** 1143
Maser Action in the Region of 60°K—C. R. Ditchfield and P. A. Forrester. (*Phys. Rev. Lett.*, vol. 1, pp. 448-449; December 15, 1958.) A three-level solid-state maser using a ruby with a 0.1 per cent Cr concentration has given a stable gain of 30 db at 9518 mc.
- 621.375.9:538.569.4.029.64** 1144
A Tunable Maser Amplifier with Large Bandwidth—R. J. Morris, R. L. Khyl and M. W. P. Strandberg. (*PROC. IRE*, vol. 47, pp. 80-81; January, 1959.) The bandwidth is 20 mc at 10 db gain, with a bath temperature of 4.2°K. The center frequency is tunable in the range 8400-9700 mc. Operation depends on a phonon-saturation mechanism in a ruby crystal.
- 621.375.9:538.569.4.029.64** 1145
A Chromium Corundum Paramagnetic Amplifier and Generator—G. M. Zverev, L. S. Kornienko, A. A. Manenkov and A. M. Prokhorov. (*Zh. Eksp. Teor. Fiz.*, vol. 34, pp. 1660-1661; June, 1958.) Note on the operation of a 3-kmc paramagnetic amplifier and generator using a single crystal of $Al_2O_3Cr_2O_3$. Levels characterized by the quantum numbers $M=3/2, \pm 1/2$ when the crystalline axis is oriented parallel to the constant external magnetic field, are used. 3-kmc absorption lines for different power levels of the 15-kmc auxiliary radiation are shown.
- 621.375.9:538.569.4.029.66** 1146
Molecular Amplifier and Generator for Submillimetre Waves—A. M. Prokhorov. (*Zh. Eksp. Teor. Fiz.*, vol. 34, pp. 1658-1659; June, 1958.) The rotational transitions of NH_3 molecules lie in the wavelength region below 1 mm. An amplifier can be constructed using a device in which radiation from one horn crosses a number of molecular beams and falls on a second horn. The maximum power obtainable is about 1 μ w.
- 621.375.9:621.385.029.6:537.533** 1147
New Method for Pumping a Fast Space-Charge Wave—Wade and Adler. (See 1412.)
- 621.375.9.029.62:621.3.011.23:621.314.63** 1148
Four-Terminal Parametric Amplifier—K. K. N. Chang. (*PROC. IRE*, vol. 47, pp. 81-82; January, 1959.) A lumped-constant model, using Ge junction diodes at a signal frequency

of 214 mc, has a gain of 8 db, a bandwidth of 0.25 mc and a noise factor of 2.5 db.

621.375.9.029.64:621.3.011.23:621.314.63 1149
Microwave Parametric Amplifier by means of Germanium Diode—B. Oguchi, S. Kita, N. Inage and T. Okajima. (PROC. IRE, vol. 47, pp. 77-78; January, 1959.) Construction and performance of an amplifier for 4 kmc, using gold- and silver-bonded Ge diodes.

621.375.9.029.64:621.3.011.23:621.314.63 1150
The Reactatron—a Low-Noise, Semiconductor Diode, Microwave Amplifier—F. A. Brand, G. Matthei and T. Saad. (PROC. IRE, vol. 47, pp. 42-44; January, 1959.) Two nonlinear-capacitor microwave p - n junction diodes are used in a balanced hybrid system. With a pump frequency of about twice the signal frequency of 2900 mc, power gains greater than 30 db with a noise figure of 2.7 db were obtained for a bandwidth of 0.5 mc.

621.376.2.029.64:621.318.134 1151
Amplitude Modulation of Centimetre Waves by means of Ferrocube—H. G. Beljers. (*Philips tech. Rev.*, vol. 18, pp. 82-86; September 14, 1956.) Losses in a ferrocube rod, caused by a modulating magnetic field, vary the amplitude of a centimeter wave reflected from it. The nature of the losses and the maximum frequency at which they occur are examined theoretically. An experimental circuit has been constructed, operating at 9300 mc with a modulation frequency of 1.3 mc; wide bandwidths are not possible.

621.376.23:621.317.373 1152
Diode Phase Detectors—S. Krishnan. (*Electronic Radio Eng.*, vol. 36, pp. 45-50; February, 1959.) The general case of unequal voltages applied to simple and balanced push-pull phase detectors is studied. Curves for design and evaluation of accuracy are given. The use of a signal greater than the reference voltage appears to be desirable in some applications.

621.376.23:621.385.029.6 1153
The Use of Beam Defocusing to Provide a Microwave Detector—Castro and Needle. (See 1413.)

GENERAL PHYSICS

537.226:539.23:538.566.2 1154
Stokes' Equations and their Application to the Refractivity of Thin Films—A. F. Wickersham, Jr. (*J. Opt. Soc. Amer.*, vol. 48, pp. 958-964; December, 1958.) The problem of determining the microwave optical properties of n parallel planes of artificial dielectric from the properties of a single planar array is solved by using an extension of Stokes' equations.

537.523:538.56 1155
Oscillographic Study of R.F. Oscillations in "Silent" Electrical Discharges—P. S. V. Setty. (*Proc. Nat. Inst. India*, Part A, vol. 24, pp. 235-239; July 26, 1958.) Oscillations produced under conditions of light or darkness in (a) iodine vapor and (b) hydrogen gas, in discharge tubes fitted with external "sleeve" electrodes are described.

537.525:538.569.4 1156
Certain Phenomena Observed on Coupling an Oscillator and a Tube of Ionized Gas Introduced in the Oscillator Inductance—T. V. Ionescu and O. C. Gheorghiu. (*Compt. rend. Acad. Sci., Paris*, vol. 246, pp. 3958-3601; June 30, 1958.) An experimental study is reported in which an oscillator coil carrying a current >20 ma surrounds an energized gas-discharge tube. The effects of the hf current on the discharge are considered, in particular the luminous

effects and the maintenance of the discharge after the energizing voltage is removed. See 1396 of 1958.

537.525.5:621.314.65 1157
A New Phenomenon of Electron Emission from Thin Mercury Films—H. von Bertele. (*Nature, (London)*, vol. 182, pp. 1148-1149; October 25, 1958.) Report of an experimental investigation of "film emission" in a specially designed discharge tube, indicating that neither extremely high current densities nor cathode evaporation are intrinsic requirements for emission from liquid metal surfaces.

537.54:621.396.822.029.6 1158
Spectral Distribution of Thermal Noise in a Gas Discharge—S. M. Bergmann. (IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-5, pp. 237-238; October, 1957.) "By means of thermodynamic considerations it is shown under which conditions microwave noise power generated by a gas discharge can be considered thermal. A critical analysis of Mumford's hypothesis is made."

537.56 1159
The Statistics of Plasma—P. S. Putter and F. Sauter. (*Ann. Phys., Lpz.*, vol. 1, Nos. 1-3, pp. 4-15; 1958.) Fundamental dynamic equations are derived comprising a system of partial differential equations.

537.56 1160
Calculation of Fields on Plasma Ions by Collective Coordinates—A. A. Broyles. (*Z. Physik*, vol. 151, pp. 187-201; April 23, 1958. In English.)

537.56:537.311.3 1161
The Electrical Conductivity of a Plasma—H. Schirmer and J. Friedrich. (*Z. Physik*, vol. 151, pp. 174-186 and 375-384; April 23 and May 8, 1958.) The exact theory of electrical conductivity of a plasma in the presence of a weak electric field is obtained by the integration of Boltzmann's equation for ionized gases, taking account of electron interaction.

537.56:538.63 1162
The Structure of Strong Collision-Free Hydrodynamic Waves—J. H. Adlam and J. E. Allen. (*Phil. Mag.*, vol. 3, pp. 448-455; May, 1958.) "A theoretical study has been made of the structure of strong "hydromagnetic" waves which are propagated across a magnetic field, in a low-density plasma where collisions can be neglected."

538.56 1163
Oblique Shock Waves in a Plasma with Finite Conductivity—M. I. Kiselev and V. I. Tseplyaev. (*Zh. Eksp. Teor. Fiz.*, vol. 34, pp. 1605-1609; June, 1958.)

538.56:537.56 1164
The "Fourth Reflection Condition" for Electromagnetic Waves in a Plasma—K. Rawer and K. Suchy. (*C. Rend. Acad. Sci., Paris*, vol. 246, pp. 3428-3430; June 23, 1958.) The introduction of a term generally neglected in applying Boltzmann's equation modifies the expression for refractive index normally used in magneto-ionic theory.

538.561:539.2 1165
Coherent Spontaneous Microwave Emission by Pulsed Resonance Excitation—L. E. Norton. (IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-5, pp. 262-265; October, 1957.) "This paper describes an investigation of the coherent microwave emission from pulse-excited ammonia molecules. Coherent and periodic pulses of near resonance frequency and 1- ω s duration excited the gas from its initial thermal equilibrium condition. Self-induced coherent emission (molecular ring-

ing) continued after the excitation field was removed. This radiation was observed during a period of 10 μ s. In an actual experiment performed, a new Doppler bandwidth reduction method was used in the gas cell. The observed spectral width of the ammonia 7.7 line was about 5 kc. The emission was used to stabilize the excitation signal source to a short-term frequency stability 2×10^{-10} ."

538.566 1166
Propagation of an Electromagnetic Field in a Medium with Spatial Dispersion—V. D. Shafronov. (*Zh. eksp. Teor. Fiz.*, vol. 34, pp. 1475-1489; June, 1958.) Investigation of the propagation of a transverse wave along a magnetic field in a plasma taking account of the thermal motion of electrons. Strong absorption occurs in the region at which Cherenkov radiation is possible.

538.566:535.4]+534.2 1167
On the Diffraction and Reflection of Waves and Pulses by Wedges and Corners—F. Oberhettinger. (*J. Res. Nat. Bur. Standards*, vol. 61, RP 2906, pp. 343-365; November, 1958.) "Various problems arising in the theory of the excitation of a perfectly reflecting wedge or corner by a plane, cylindrical, or spherical wave, are dealt with. The incident wave is represented by a line source (acoustic or electromagnetic) parallel to the edge. The spherical wave is emitted by an acoustic point source or by a Hertz dipole with its axis parallel to the edge. The case of an incident plane wave field is obtained as the limiting case (for large distances of the source from the edge) of the cylindrical or spherical wave excitation."

538.566:535.42 1168
The Diffraction of an Electromagnetic Plane Wave by a Perfectly Conducting Half-Plane—P. Poincelot. (*Compt. rend. Acad. Sci., Paris*, vol. 246, pp. 3418-3419; June 23, 1958.)

538.566.2:548 1169
Electromagnetic Waves in Isotropic and Crystalline Media Characterized by Permittivity with Spatial Dispersion—V. L. Ginzburg. (*Zh. Eksp. Teor. Fiz.*, vol. 34, pp. 1593-1604; June, 1958.) The necessity of taking account of spatial dispersion in non-gyrotropic media in relation to the permittivity tensor and also in the analysis of longitudinal (plasma) waves propagating in an isotropic medium or along the principal dielectric axes in crystals is stressed. Collective energy losses and the Cherenkov effect are also considered.

538.569.4.029.64 1170
Attempt to Interpret the Shape of Paramagnetic Resonance Signals by the Introduction of a High-Frequency Demagnetizing Field—A. Charru. (*Compt. rend. Acad. Sci., Paris*, vol. 246, pp. 3445-3447; June 23, 1958.) The theory accounts for the asymmetry of the resonance curve of organic radicals such as diphenylhydrazyl at 3 kmc.

538.569.4.029.64:061.3 1171
Microwave Physics—K. W. H. Stevens. (*Nature, (London)*, vol. 182, pp. 1121-1123; October 25, 1958.) Discussion of resonance techniques, particularly as applied for microwave amplification, with reference to three papers read at the British Association meeting (Section A) held in Glasgow, September 1958.

538.569.4.029.64:535.34 1172
Comparison of the Sensitivities of Paramagnetic-Electron-Resonance Spectrographs in the 3-cm Band—J. Roch. (*C. Rend. Acad. Sci., Paris*, vol. 247, pp. 59-62; July 7, 1958.) Modification of the coupling between a waveguide and a cavity is discussed for the following cases: (a) negligible background noise; (b) sig-

nal intensity of the same order as background noise.

539.2:548.0 1173
The Elastic Model of Lattice Defects—J. D. Eshelby. (*Ann. Phys., Lps.*, vol. 1, Nos. 1-3, pp. 116-121; 1958. In English.) It is shown that between rigid misfitting spheres in an elastic continuum there is a repulsive interaction if the spheres are bonded to the medium. The effect of other boundary conditions is discussed, and a simplified method is given for calculating the interaction between various types of defect.

539.2:548.0 1174
Energy States of a One-Dimensional Crystal with Vacant Lattice Sites—A. Mozumder. (*Proc. Nat. Inst. Sci. India*, Part A, vol. 23, pp. 288-294; September 26, 1958.)

539.2:548.4 1175
Relations between the Concentrations of Imperfections in Solids—F. A. Kröger and H. J. Vink. (*J. Phys. Chem. Solids*, vol. 5, pp. 208-223; May, 1958.) Imperfections having opposite effective charges tend to increase each other's concentrations. Foreign atoms are incorporated in a manner dependent both on energy levels caused by them and on the type of imperfection prevailing in the crystal in their absence.

GEOPHYSICAL AND TERRESTRIAL PHENOMENA

523.164.3 1176
On a Feature of Galactic Radio Emission—H. Tunmer. (*Phil. Mag.*, vol. 3, pp. 370-376; April, 1958.) "Observations of the intensity distribution of radio emission have shown a bright belt approximately normal to the galactic plane and passing through the anti-centre. An explanation is suggested in terms of the highly anisotropic radiation from relativistic electrons moving in the magnetic field of the local spiral arm. This suggestion avoids the supposition that the sun is in a special position in the galaxy."

523.164.3:535.417 1177
A Phase-Sensitive Interferometer Technique for the Measurement of the Fourier Transforms of Spatial Brightness Distributions of Small Angular Extent—R. C. Jennison. (*Mon. Not. R. Astr. Soc.*, vol. 118, pp. 276-284; September, 1958.) The RF interferometer described consists of three antenna systems tuned to 127 mc.

523.164.32:550.385.4 1178
On the Power Spectrum of Solar Radio Outburst and its Relation to S.W.F. (Dellinger Phenomenon) and Geomagnetic Storm—Y. Hakura. (*J. Radio Res. Labs, Japan*, vol. 5, pp. 283-293; October, 1958.) Solar radio outbursts are classified according to their power spectra. Those with power spectra increasing towards lower frequencies are associated with magnetic storms but not with S.I.D.'s, while those with spectra increasing towards higher frequencies are associated with S.I.D.'s without an accompany magnetic storm. Outbursts with flat power spectra appear to combine the attributes of the other two types.

523.75:523.165 1179
The Possibility of Forecasting Solar Phenomena and their Terrestrial Repercussions by the Study of the Intensity of Cosmic Radiation—J. P. Legrand. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 70-73; July 7, 1958.) An analysis of observations made from April 12, 1957 to March 31, 1958 shows that a decrease in cosmic-ray intensity of 1-3% occurs 24-48 h before chromospheric eruptions which are followed by intense geomagnetic and ionospheric

disturbances. This decrease also precedes or coincides with an increase in 200-mc solar RF noise. See also 2396 of 1958. (Brown).

523.75:550.385 1180
Discrimination between Chromospheric Eruptions associated with Geomagnetic Disturbances during the Present Solar Maximum—M. Lépineux. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 109-111; July 7, 1958.) An analysis of observations from April 15, 1957 to March 31, 1958 shows that some but not all solar flares are accompanied by a marked increase of RF noise at about 200 mc.

523.75:551.510.535 1181
A Simple Method of Detecting Solar Flare Effects in the Lower Ionosphere—E. Lauter and K. Sprenger. (*Z. Meteorol.*, vol. 12, pp. 205-210; July, 1958.) Field-strength records of transmitters operating at frequencies 150-300 kc over a distance of 1300 km provided a sensitive means of detecting and identifying solar flare effects.

523.755 1182
Intensities, Polarization and Electron Density of the Solar Corona from Photographs taken at the Total Solar Eclipse of 1952 February 25—H. von Klüber. (*Mon. Not. R. Astr. Soc.*, vol. 118, pp. 201-223; September, 1958.)

550.385.4 1183
Sudden Commencements of Magnetic at Tamanrasset—J. L. Bureau. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 112-114; July 7, 1958.) An analysis of observations made at Tamanrasset (22° 48'N, 5° 31'E) from 1950 to 1956 inclusive.

550.389.2:629.19 1184
Scale Height of the Upper Atmosphere—C. H. Bosanquet. (*Nature (London)*, vol. 182, pp. 1010-1011; October 11, 1958.) The effect of air drag on the orbit of an earth satellite is calculated allowing for the ellipticity of the earth.

550.389.2:629.19 1185
Sputnik as a Tool for Securing Geodetic Information—L. Gold. (*J. Frank. Inst.*, vol. 266, pp. 103-107; August, 1958.)

550.389.2:629.19 1186
A Possible Cause of Air Resistance Changes in Satellite Orbits—J. Bartels. (*Naturwiss.*, vol. 45, p. 181; April, 1958.) The variation in the extent of the atmosphere in accordance with geomagnetic activity, particularly in the auroral zones, is suggested as a possible cause of observed changes in satellite orbital period.

550.389.2:629.19 1187
The Rotation of the First Russian Earth Satellite—J. H. Thomson. (*Phil. Mag.*, vol. 3, pp. 912-916; August, 1958.) Fading records, obtained during transits over Cambridge at about 0330 U.T. October 16-22, 1957, have been examined. The 40 mc records show a discontinuous change in fading rate at the time of nearest approach. This is attributed to changes in polarization in certain aspects of the satellite relative to the observer, caused by the generation of a cone by the rotating aerial. It is concluded that the satellite rotated at 6.5 rev/min, generating a cone of semi-angle 45°, the direction of the rotating axis being at 23° to the direction of the orbit at perigee.

550.389.2:629.19 1188
Determination of the Orbit of an Artificial Satellite—N. Carrara, P. F. Ceccacci and L. Ronchi. (*Proc. IRE*, vol. 47, p. 75; January, 1959.) Theory of a method depending on Doppler measurements at four stations. No initial assumptions about the orbit are required.

550.389.2:629.19 1189
A New Method of Tracking Artificial Earth Satellites—A. P. Willmore. (*Nature (London)*, vol. 182, pp. 1008-1010; October 11, 1958.) A photoelectric method of optical tracking has been developed which appears to have a precision comparable with astronomical observations. A record of observations of the rocket of Sputnik III on July 23, 1958 at Tatsfield, Surrey, is shown.

550.389.2:629.19 1190
Tracking Orbits of Man-Made Moons—C. A. Schroeder, C. H. Looney, Jr, and H. E. Carpenter, Jr. (*Electronics*, vol. 32, pp. 33-37; January 2, 1958.) Description of the Mini-track interferometer-type system designed to measure satellite position to an accuracy within 20 seconds of arc with a time precision of one millisecond.

550.389.2:629.19 1191
Doppler Satellite Measurements—H. D. Tanzman, G. A. MacLeod, and W. T. Scott. (*Proc. IRE*, vol. 47, pp. 75-76; January, 1959.) A note on records obtained during successful and unsuccessful launchings.

550.389.2:629.19 1192
Further Radio Observations on Artificial Satellites—P. F. Ceccacci and C. Carreri. (*Ricerca sci.*, vol. 28, pp. 1817-1831; September, 1958.) Report on data obtained in Florence, Italy, from radio observations of satellites 1957 α , 1957 β , 1958 α , 1958 β 2 (Vanguard) and 1958 γ .

550.389.2:629.19 1193
Signal Strength Recordings of the Satellite 1958 I2 (Sputnik III) at College, Alaska—R. Parthasarathy and G. C. Reid. (*Proc. IRE*, vol. 47, pp. 78-79; January, 1959.)

550.389.2:629.19:002.6 1194
World Data Centre for Rockets and Satellites, Slough—(*Nature (London)*, vol. 182, p. 988; October 11, 1958.) It was agreed at the recent Moscow meeting of the Special Committee for the I.G.Y. to accept a British offer of a third center where data from British investigations would be collected and exchanged with the centers in Washington and Moscow.

551.508.8 1195
Radiosonde Measurement of Vertical Electric Field and Polar Conductivity—O. C. Jones, R. S. Maddever and J. H. Sanders. (*J. Sci. Instr.*, vol. 36, pp. 24-28; January, 1959.) A field mill and a Gerdian-type conductivity apparatus are described for use with a modified Kew Mk II radiosonde. These are used for measuring the electrical properties of the lower atmosphere and clouds.

551.510.535 1196
On the Semi-diurnal Lunar Variations in the Critical Frequency, Semi-thickness and Height of Maximum Ionization of the F₂ Layer in the Daytime—T. Yonezawa and Y. Arima. (*J. Radio Res. Labs., Japan*, vol. 5, pp. 303-314 October, 1958.) Results obtained for f_oF_2 show reasonable agreement with those of other workers but contradict a theory of F₂-layer formation previously postulated [128 of January (Yonezawa)]. Results for other parameters are doubtful, owing to statistical errors, but appear to be exactly opposite in phase to those of other workers.

551.510.535 1197
Back-Scatter Ionospheric Sounding Experiments—I. Ranzi. (*Note Recensioni Notiz.*, vol. 7, pp. 201-212; March/April, 1958. English summary, pp. 213-215.) Report on tests at 18.6 mc made near Rome from August to October, 1957. Large discrepancies are found with results

derived from vertical soundings. Subsequent tests are proceeding at 21.64 and 22.3 mc.

551.510.535:551.513 1198
Ionospheric Drift Measurements in the Long-Wave Range as a Contribution to the Problem of General Circulation of the Upper Atmosphere—K. Sprenger. (*Z. Meteorol.*, vol. 12, pp. 211–218; July, 1958.) Spaced-receiver measurements at 245 kc of ionospheric drift at altitude 90 km indicate distinct seasonal reversals in the circulation of the upper atmosphere. Small diurnal variations are superimposed on the normal drifts and an additional N-S component is observed during magnetic storms.

551.510.535"1958":621.396.11 1199
Ionosphere Review 1958—T. W. Bennington. (*Wireless World*, vol. 65, pp. 72–73; February, 1959.) Solar activity, F₂-layer ionization and ionospheric and magnetic disturbances were less than those recorded in 1957.

551.510.535 1200
The Ionosphere. [Book Review]—K. Rawer. Crosby Lockwood and Son, London, 202 pp., 1958; 42s. (*Nature (London)*, vol. 182, p. 1116; October 25, 1958.) A translation of the original German book (see 421 of 1954).

551.594 1201
Atmosphere Electricity [Book Review]—J. A. Chalmers. Pergamon Press, London, 327 pp., 63s. 1957. (*Met. Mag., London*, vol. 87, pp. 343–344; November, 1958.) Enlarged and revised edition of an earlier work (see 98 of 1950).

LOCATION AND AIDS TO NAVIGATION

621.396.93:061.3 1202
Radio Aids to Navigation—(*Nature (London)*, vol. 182, pp. 978–980; October 11, 1958.) Summaries of three papers read at the British Association meeting (Section G, Engineering) held in Glasgow, September, 1958.

621.396.933.2:621.396.677 1203
An Investigation of High-Frequency Direction-Finding Errors caused by Nearby Vertical Reradiators—C. W. McLeish and R. S. Roger. (*Proc. IRE, Part B*, vol. 106, pp. 58–60; January, 1959.) An expression for the maximum errors caused by reradiators is derived and the susceptibilities to error of three types of DF aerial are compared.

MATERIALS AND SUBSIDIARY TECHNIQUES

533.5 1204
A New Type of Titanium Getter Pump—L. Holland, L. Laursen and J. T. Holden. (*Nature (London)*, vol. 182, pp. 851–852; September 27, 1958.) Description of recent developments in the design of vapor sources in which refractory supports are avoided.

535.5:621.385.029.6 1205
The Titanium Pump: a Device for the Maintenance of Vacuum in Electron Tubes—H. Huber and M. Warnecke. (*Le Vide*, March/April, 1958, vol. 13, pp. 84–90; in French and English.) A new type of getter-ion pump is described.

535.215 1206
The Influence of Defect Levels on Photoemission—W. E. Spicer. (*RCA Rev.*, vol. 19, pp. 555–563; December, 1958.) Minimum thermionic emission, maximum conductivity and optimum effect of band-loading are produced by *p*-type defects near the top of the valence band. If photoemission from defect levels is desired, *n*-type levels near the bottom of the conduction band are needed. The maximum quantum efficiency from defect levels is about 1 per cent.

535.215:535.37 1207
The Intensity-Dependence of Photoconduction and Luminescence of Photoconductors in the Stationary State—II. A. Klasens. (J. Phys. Chem. Solids, vol. 7, pp. 175–200; November, 1958.) Experimental data are reviewed and two models for the electron-hole recombination processes are considered. The one-state model involving recombination either direct or via one impurity state, is shown to be generally inadequate, but the two-state model, with two discrete levels having different capture cross sections for electrons and holes, can explain quantitatively many aspects of photoconduction and luminescence. 60 references.

535.215 + 535.37]:546.482.21 1208
Fundamental Absorption Edge in Cadmium Sulphide—D. Dutton. (*Phys. Rev.*, vol. 112, pp. 785–792; November 1, 1958.) The absorption and reflection spectra of CdS have been determined in the temperature range 90°–340°K by photoelectric measurements on single crystals, using polarized light.

535.215:546.863.221 1209
Optical Properties of Antimony Trisulphide Films—C. Kunze. (*Ann. Phys. Lpz.*, vol. 1, pp. 165–172; 1958.) Measurements of absorption coefficient and refractive index on vapor-deposited Sb₂S₃ films were made at wavelengths of 400–1000 mμ.

535.215:546.863.221 1210
Photoelectric Properties of Antimony Trisulphide Films—C. Kunze. (*Ann. Phys., Lpz.*, vol. 1, pp. 173–182; 1958.) The spectral distribution of photoconductivity in Sb₂S₃ films of thickness 200–400 mμ was measured using a vidicon-type camera tube.

535.37 1211
Transient Response of Phosphors—G. I. Cohn and H. M. Musal. (*IRE TRANS. ON COMPONENT PARTS*, vol. CP-5, No. 2, pp. 90–101; June, 1958. Abstract, *PROC. IRE*, vol. 46, p. 1553. August, 1958.)

535.37:546.472.21 1212
ZnS:Sn,Li Phosphor—A. Wachtel. (*J. Electrochem. Soc.*, vol. 105, pp. 432–433; July, 1958.)

535.376 1213
The "Memory" Effect in the Enhancement of Luminescence by Electric Fields—G. Destriau. (*Z. Physik*, vol. 150, pp. 447–455; March 10, 1958.) See also 2752 of 1958.

535.376:546.472.21 1214
Electron Traps and the Electroluminescence Brightness and Brightness Waveform—F. F. Morehead, Jr. (*J. Electrochem. Soc.*, vol. 105, pp. 461–468; August, 1958.) "This work describes an investigation of the effect of temperature, voltage, and rise time on the size and shape of the brightness waves of electroluminescence; sawtooth and square pulse voltage waveforms were used to excite Cu-activated ZnS phosphors in a slightly conducting medium."

535.376:546.472.21 1215
Electroluminescence of Insulated Particles—K. Maeda. (*J. Phys. Soc. Japan*, vol. 13, pp. 1352–1361; November, 1958.) The local high electric field is attributed to disturbance of the applied field by conductive substances in the phosphor. Distributions of local field are calculated and the relation between brightness and applied voltage is discussed. The estimated luminous efficiency of a ZnS cell agrees with the observed value.

535.376:546.482.21 1216
The L_{2,3} Emission Spectrum and the Position of the L_{2,3} Absorption Edge of S in CdS—

G. Eichhoff. (*Ann. Phys., Lpz.*, vol. 1, Nos. 1–3, pp. 55–69; 1958.)

537.227 1217
Energy Loss Processes in Ferroelectric Ceramics—B. Lewis. (*Proc. Phys. Soc.*, vol. 73, pp. 17–24; January 1, 1959.) Measurements of permittivity, elastic compliance, and the electrical and mechanical loss coefficients of Ba(T_{0.9}Zr_{0.05})O₃ and (Ba_{0.94}N_{0.02})TiO₃, and their changes with time, are described. The observed losses can be accounted for entirely as macro- and micro-hysteresis loss associated with movement of domain boundaries.

537.227:547.476.3 1218
Upper Curie Temperature and Domain Structure in Various Regions of Rochelle Salt Crystals—H. E. Müser and H. Flunkert. (*Z. Physik*, vol. 150, pp. 21–32; December 21, 1957.)

537.228.1:549.514.51 1219
Mechanical Resonance Dispersion in Quartz at Audio-Frequencies—E. R. Fitzgerald. (*Phys. Rev.*, vol. 112, pp. 765–784; November 1, 1958.) Measurements of the complex shear compliance of single crystals of quartz and fused quartz at frequencies from 100 to 5000 cps have resulted in the discovery of sharp resonances in the compliance. Analysis of the data on the basis of a generalized stress-strain relation involving a linear combination of strain and its first and second time derivatives gives a close fit to the experimental curves.

537.311.3:538.22 1220
Electrical Resistivity of Compounds with Ordered Spin Arrangements—R. Parker. (*Phil. Mag.*, vol. 3, pp. 853–861; August, 1958.) "A summary is given of the published data of electrical resistivity of semiconductors which exhibit ordered spin arrangements below a Curie or Néel temperature. It is found generally that electrical anomalies accompany the advent of magnetic order only if the specific resistivity at the transformation temperature is less than a critical value. The implications of this rule are discussed in the light of recent theories of electrical conduction in spontaneously magnetized materials."

537.311.31 1221
The Quantum Theory of the High-Frequency Surface Impedance of a Metal—M. Ya. Azbel'. (*J. Phys. Chem. Solids*, vol. 7, pp. 105–117; November, 1958.)

537.311.31:[546.86 + 546.87] 1222
The Thermal and Electrical Resistivity of Bismuth and Antimony at Low Temperatures—G. K. White and S. B. Woods. (*Phil. Mag.*, vol. 3, pp. 342–359; April, 1958.)

537.311.33 1223
A Theory of Impurity Conduction: Part 1—T. Kasuya. (*J. Phys. Soc. Japan*, vol. 13, pp. 1096–1110; October, 1958.) Two ranges of impurity concentration in semiconductors such as Ge and Si are considered:—(a) a fairly large concentration, with impurity levels merged into the conduction bands; (b) a fairly small concentration, with impurity levels separated from conduction bands.

537.311.33 1224
A Theory of Impurity Conduction: Part 2—T. Kasuya and S. Koide. (*J. Phys. Soc. Japan*, vol. 15, pp. 1287–1297; November, 1958.) The case is considered of very low impurity concentration, where the fluctuation of local potential energy is much larger than the translational energy. In this case electron-phonon interaction is important since it enables electrons to move between localized states. Agreement between theory and experiment is fairly satisfactory. Part 1: 1223 above.

- 537.311.33 1225
Mobility of Electrons in Nondegenerate Semiconductors considering Electron-Electron Scattering—M. S. Sodha and P. C. Eastman. (*Z. Physik*, vol. 150, pp. 242–246; January 27, 1958. In English.)
- 537.311.33 1226
The Crystal Structure and Properties of the Group VB to VIIB Elements and of Compounds Formed between Them—E. Mooser and W. B. Pearson. (*J. Phys. Chem. Solids*, vol. 7, pp. 65–77; October, 1958.)
- 537.311.33 1227
Semiconductor Compounds of Predominantly Homopolar Character—H. Welker and H. Weiss. (*Z. Metallk.*, vol. 49, pp. 563–570; November, 1958.) Predominantly homopolar semiconductors are considered and the pronounced rectifier and transistor effects, as well as thermoelectric and photoelectric effects, exhibited by compounds such as InSb are discussed.
- 537.311.33 1228
Time-Dependent Changes in Excess Carrier Concentrations in the Presence of Surface Recombination—J. D. Nixon and P. C. Banbury. (*Proc. Phys. Soc.*, vol. 73, pp. 54–58; January 1, 1959.) The decay of excess carriers in a thin slab of semiconductor is considered, assuming that surface recombination takes place through Shockley-Read type centers (420 of 1953). Numerical calculations for Ge show that the time constants differ from the steady-state lifetime except in the case of low trap densities.
- 537.311.33:535.215 1229
Role of Traps in the Photoelectromagnetic and Photoconductive Effects—R. N. Zitter. (*Phys. Rev.*, vol. 112, pp. 852–855; November, 1958.) Expressions for the steady-state photoelectromagnetic and photoconductive currents are obtained which show that in certain cases the photoelectromagnetic current is determined by a lifetime different from the one determining the photoconductive current. This theory has been applied to *p*-type InSb.
- 537.311.33:535.215 1230
Optical Method for Determining Carrier Lifetimes in Semiconductors—L. Hultdt. (*Phys. Rev. Lett.*, vol. 2, pp. 3–5; January 1, 1959.) Photoelectrically liberated cause an increase in the infrared absorption of semiconductors. This increase is used to measure carrier lifetime in Ge.
- 537.311.33:537.32 1231
The Power Limits of Thermoelectric Effects in Semiconductors—O. Böttger. (*Z. Physik*, vol. 151, pp. 296–306; May 8, 1958.) The efficiency of a loaded thermocouple and the maximum temperature reduction of a Peltier element are calculated. A criterion for assessing the suitability of materials for thermoelectric devices is defined. Suitable semiconductors are restricted to the conductivity range 10^2 – $3 \times 10^4 \Omega^{-1} \text{cm}^{-1}$.
- 537.311.33:538.615 1232
Theory of Impurity Photo-ionization Spectrum of Semiconductors in Magnetic Fields—R. F. Wallis and H. J. Bowlden. (*J. Phys. Chem. Solids*, vol. 7, pp. 78–89; October, 1958.) The theory shows two types of photo-ionization spectrum depending on the impurity ionization for zero field being larger or smaller than $LeH/4\pi^*mc$. Detailed calculations are presented of the high-field spectrum for donor impurities in InSb.
- 537.311.33:538.63 1233
Surface Transport Theory—J. N. Zemel. (*Phys. Rev.*, vol. 112, pp. 762–765; November 1, 1958.) A theory is presented for the dependence of the galvanomagnetic parameters on the surface potential of a semiconductor. The expressions for the conductivity reduce to that given by Schrieffer (2322 of 1955) when the magnetic field is zero. Formal equations for the magnetoconductivity and Hall coefficient are derived.
- 537.311.33:538.632 1234
Influence of the Geometry on Hall Effect and Magneto-resistance Effect in Rectangular Semiconductor Specimens—H. J. Lippmann and F. Kuhrt. (*Naturwiss.*, vol. 45, pp. 156–157; April, 1958.)
- 537.311.33:538.632 1235
Hall Effect in High Electric Fields—J. F. Gibbons. (*Proc. IRE*, vol. 47, p. 102; January, 1959.) Discussion of the limitation of Hall voltage in different semiconductor materials due to the terminal velocity of the charge carriers.
- 537.311.33:539.2 1236
Energy-Band Interpolation Scheme based on a Pseudopotential—J. C. Phillips. (*Phys. Rev.*, vol. 112, pp. 685–695; November 1, 1958.) The effective potential for electrons near the Fermi level is split into two parts, the part due to the core, and the part due to the other valence electrons. It is assumed that the relative effect of the core is small enabling two-parameter pseudopotentials to be constructed for diamond and silicon that give good agreement with orthogonalized plane-wave calculations and experiment at special points of the Brillouin zone, and also yield reasonable results for the bands at other points of the zone.
- 537.311.33:546.24:538.632 1237
Effect of Hydrostatic Pressure on the Anomalous Hall-Coefficient Reversal in Single-Crystal Tellurium—A. Nussbaum, J. Myers and D. Long. (*Phys. Rev. Lett.*, vol. 2, pp. 6–9; January, 1959.)
- 537.311.33:[546.26 + 546.28 + 546.289] 1238
The Electronic Structure of Diamond, Silicon and Germanium—G. G. Illall. (*Phil. Mag.*, vol. 3, pp. 429–439; May, 1958.) "The equivalent orbital method of describing the electronic structure of valence crystals, introduced in an earlier discussion of diamond [see *Proc. roy. Soc. A.*, vol. 202, pp. 336–344; August 7, 1950], is extended in various ways."
- 537.311.33:546.26–1:538.63 1239
Magnetic Field Dependence of the Hall Effect and Magnetoresistance in Graphite Single Crystals—D. E. Soule. (*Phys. Rev.*, vol. 112, pp. 698–707; November 1, 1958.)
- 537.311.33:546.26–1:538.63 1240
Analysis of Galvanomagnetic de Haas-van Alphen Type Oscillations in Graphite—D. E. Soule. (*Phys. Rev.*, vol. 112, pp. 708–714; November 1, 1958.)
- 537.311.33:546.26–1:538.63 1241
Analysis of Multicarrier Galvanomagnetic Data for Graphite—J. W. McClure. (*Phys. Rev.*, vol. 112, pp. 715–721; November 1, 1958.)
- 537.311.33:[546.28 + 546.289] 1242
Theoretical Calculation of Distribution Coefficients of Impurities in Germanium and Silicon, Heats of Solid Solution—K. Weiser. (*J. Phys. Chem. Solids*, vol. 7, pp. 118–126; November, 1958.) Distribution coefficients are calculated for group III, IV and V impurities in Ge and Si by estimating the change in energy involved in transferring the impurity atoms from a crystalline reservoir to the solid and molten host material respectively. Good agreement with experimental data is obtained.
- 537.311.33:[546.28 + 546.289] 1243
The Effect of Pressure on the Optical Absorption Edge of Germanium and Silicon—T. E. Slykhouse and H. F. Drickamer. (*J. Phys. Chem. Solids*, vol. 7, pp. 210–213; November, 1958.)
- 537.311.33:546.28 1244
Absorption Spectrum of Arsenic-Doped Silicon—H. J. Hrostowski and R. H. Kaiser. (*J. Phys. Chem. Solids*, vol. 7, pp. 236–239; November, 1958.) Experimental results agree well with results obtained from the effective mass theory. Bands not obeying the calculated term scheme are shown to arise from other donor impurities. Optical and thermal ionization energies appear to be very similar for both group III and group V impurities.
- 537.311.33:546.28 1245
Electropolishing Silicon in Hydrofluoric Acid Solutions—D. R. Turner. (*J. Electrochem. Soc.*, vol. 105, pp. 402–408; July, 1958.)
- 537.311.33:546.28:535.215 1246
Measurement of Surface Recombination Velocity in Silicon by Steady-State Photoconductance—H. M. Bath and M. Cutler. (*J. Phys. Chem. Solids*, vol. 5, pp. 171–179; May, 1958.) In the method described, excitation can be accomplished by penetrating as well as non-penetrating light, so that the components of lifetime which depend respectively on the bulk and surface properties of the sample can be emphasized. Application of the method to Si indicates that surface recombination processes are different from those of Ge.
- 537.311.33:546.289 1247
New Observations on Dislocations in Germanium—H. Dorendorf. (*Z. angew. Phys.*, vol. 9, pp. 513–519; October, 1957.) The origin and extent of dislocations and their elimination by annealing is discussed with reference to series of photographs.
- 537.311.33:546.289 1248
High-Vacuum Studies of Surface Recombination Velocity for Germanium—H. H. Madden and H. E. Farnsworth. (*Phys. Rev.*, vol. 112, pp. 793–800; November 1, 1958.) Details are given of the observed changes in the value of the surface recombination velocity of minority carriers for (100) faces of Ge crystals under various vacuum conditions.
- 537.311.33:546.289 1249
Free-Carrier Absorption in *n*-Type Ge—R. Rosenberg and M. Lax. (*Phys. Rev.*, vol. 112, pp. 843–852; November 1, 1958.) The structure of the conduction band of Ge is completely taken into account in calculating the cross-section in second-order Born approximation. An excellent fit to data at 450°K is obtained.
- 537.311.33:546.289 1250
Comparison of Radio-Copper and Hole Concentrations in Germanium—K. Wolfstirn and C. S. Fuller. (*J. Phys. Chem. Solids*, vol. 7, pp. 141–145; November, 1958.) The changes in carrier concentration for four ⁶⁴Cu concentrations were determined in Ga-doped and As-doped Ge, by resistivity measurements. The results are in reasonable agreement with calculations based on published values of the ionization energy levels for Cu in Ge [2192 of 1957 (Woodbury and Tyler)].
- 537.311.33:546.289 1251
Melted-Layer Crystal Growth and its Application to Germanium—F. H. Horn. (*J. Electrochem. Soc.*, vol. 105, pp. 393–395; July, 1958.) Single-crystal material of constant impurity distribution may be grown conveniently from a doped melted layer maintained above solid retained in a crucible.

- 537.311.33:546.289 1252
Etching and Polishing of Germanium Surfaces—D. Grist and E. Preuss. (*Z. angew. Phys.*, vol. 9, pp. 526–531; October, 1957.) Comparison of the effectiveness of various methods.
- 537.311.33:546.289:535.215:538.639 1253
A Quadratic Photoelectromagnetic Effect in Germanium—M. Cardona and W. Paul. (*J. Phys. Chem. Solids*, vol. 7, pp. 127–140; November, 1958.) The quadratic effect occurs when the magnetic field is rotated about an axis lying in the illuminated face and perpendicular to the original direction of the field. Its relation to the usual linear photoelectromagnetic effect is investigated theoretically and experimentally, and satisfactory agreement obtained. The ratio of the quadratic to the linear photoelectromagnetic field is, for small light intensities, a function only of the Hall angles, the magnetoresistance coefficients of electrons and holes, the impurity concentration, and the magnetic field. See also 2460 of 1958 (Kikoin and Bykovskii).
- 537.311.33:546.289:539.164 1254
Alpha-Particle Irradiation of Ge at 4.2°K—G. W. Gobeli. (*Phys. Rev.*, vol. 112, pp. 732–739; November 1, 1958.) The radiation removed carriers from the samples, no measurable thermal recovery being observed below 22°K. Two distinct regions of thermal recovery were observed, the maximum rates of change occurring near 33°K and 67°K.
- 537.311.33:546.289:621.396.822 1255
Study on the Correlation between the Noise by Hole Generation and Surface Recombination Velocity at Ge Fused Junction—K. Komatsubara. (*J. Phys. Soc. Japan*, vol. 13, pp. 1409–1410; November, 1958.)
- 537.311.33:546.289:621.396.822 1256
1/f Noise in Germanium Devices—T. B. Watkins. (*Proc. Phys. Soc.*, vol. 73, pp. 59–68; January 1, 1959.) Simultaneous field-effect and noise measurements were made on specially prepared Ge junction diodes to test current theories of 1/f noise. As a result a mechanism is proposed which is based on Petritz's theory involving local barrier breakdown.
- 537.311.33:546.289:621.396.822 1257
Excess Noise in Narrow Germanium p-n Junctions—T. Vajima and L. Esaki. (*J. Phys. Soc. Japan*, vol. 13, pp. 1281–1287; November, 1958.) Alloyed junctions less than 200 Å wide, showing inverted rectification and negative resistance in the forward direction, gave considerable 1/f noise in a certain forward-bias range. The noise was stable and surface-insensitive and mainly associated with excess forward current. Measurements were made at temperatures down to 77°K. The interpretation of results is discussed.
- 537.311.33:546.492.241 1258
Preparation and Electrical Properties of Mercury Telluride—T. C. Harman, M. J. Logan and H. L. Goering. (*J. Phys. Chem. Solids*, vol. 7, pp. 228–235; November, 1958.) Details are given of the preparation, by direct reaction of Hg vapor with liquid Te, zone melting and annealing of HgTe. The composition, near stoichiometry, could be altered by the anneal. Hall coefficient and resistivity were measured as a function of temperature.
- 537.311.33:546.681.19 1259
Diffusion of Zinc in Gallium Arsenide—J. W. Allen and F. A. Cunnell. (*Nature (London)*, vol. 182, pp. 1158; October 25, 1958.) A theoretically derived curve showing impurity concentration as a function of depth for a diffusion temperature of 1000°C, is confirmed experimentally.
- 537.311.33:546.682.86 1260
Electrical Conduction in p-Type InSb between 100° and 2°K—E. H. Putley. (*Proc. Phys. Soc.*, vol. 73, pp. 128–131; January 1, 1959.)
- 537.311.33:546.682.86 1261
Nuclear Magnetic Resonance in Indium Antimonide: Part 1—The Effect of Impurities—E. H. Rhoderick. (*Phil. Mag.*, vol. 3, pp. 545–563; June, 1958.) In n-type (Te-doped) and p-type (Zn-doped) samples of InSb the amplitude and width of the ¹¹⁵In nuclear magnetic resonance are very sensitive to the presence of donors or acceptors. This is ascribed to broadening arising from the interaction of the nuclear quadrupole moment with the Coulomb field of the ionized impurity atoms.
- 537.311.33:546.682.86 1262
Nuclear Magnetic Resonance in Impure Indium Antimonide—M. H. Cohen. (*Phil. Mag.*, vol. 3, pp. 564–566; June, 1958.) An analysis of experimental results (1261 above) on nuclear quadrupole broadening in InSb based on a detailed theory of line shape.
- 537.311.33:546.682.86:538.63 1263
Magnetically Induced Impurity Banding in n-InSb—R. J. Sladek. (*J. Phys. Chem. Solids*, vol. 5, pp. 157–170; May, 1958.) "A detailed study of the effect of a strong magnetic field on donor levels in n-InSb has been made by means of Hall-effect and resistivity measurements down to 1.5°K, using field strengths up to 28,000 G. The results of these measurements indicate that donor levels which are split off from the conduction band under the influence of the magnetic field form a narrow impurity band with finite mobility. The position of the impurity band (i.e. donor ionization energy) and the mobility of electrons in the impurity band have been determined as a function of magnetic field strength."
- 537.311.33:546.682.86:538.63 1264
An Oscillatory Transverse Magnetoresistance Effect in n-Type InSb—E. H. Putley. (*Proc. Phys. Soc.*, vol. 73, pp. 131–133; January 1, 1959.)
- 537.311.33:546.682.86:538.63 1265
Interpretation of the Transverse Magnetoresistance in p-type Indium Antimonide at Liquid-Nitrogen Temperature—C. H. Champness. (*Phys. Rev. Lett.*, vol. 1, pp. 439–440; December 15, 1958.)
- 537.311.33:546.714–31 1266
Study of the Semiconducting Properties of Pyrolusite—J. N. Das. (*Z. Phys.*, vol. 151, pp. 345–350; May 8, 1958. In English.)
- 537.311.33:546.817.221 1267
Magnetoresistance in PbS, PbSe, and PbTe at 295°, 77.4°, and 4.2°K—R. S. Allgaier. (*Phys. Rev.*, vol. 112, pp. 828–836; November 1, 1958.) Magnetoresistance measurements at 77.4°K conformed to the general phenomenological weak-field theory. Saturation effects and deviations from the sinusoidal behavior were observed at 4.2°K. The data generally favored a model with considerable mass anisotropy.
- 537.311.33:546.863.241:537.32 1268
Thermoelectric Properties of Antimony Telluride and the Solid Solutions Sb₂Te₃–Bi₂Te₃—H. Benel. (*C. Rend. Acad. Sci., Paris*, vol. 247, pp. 584–587; August 4, 1958.) Thermoelectric power and electrical resistivity were measured and the effect of doping was investigated.
- 537.311.33:548.0 1269
Ternary Semiconducting Compounds with Sodium-Chloride-Like Structure: AgSbSe₂,
- AgSbTe₂, AgBiS₂, AgBiSe₂—S. Geller and J. H. Wernick. (*Acta Cryst.*, vol. 12, Part 1, pp. 46–54; January 10, 1959.) Crystallographic data on high- and low-temperature modifications are discussed.
- 537.311.33:669.018.13 1270
Constitution of the AgSbSe₂–AgSbTe₂–AgBiSe₂–AgBiTe₂ System—J. H. Wernick, S. Geller and K. E. Benson. (*J. Phys. Chem. Solids*, vol. 7, pp. 240–248; November, 1958.)
- 537.311.4 1271
Induced Conductivity at the Surface of Contact between Metals—S. D. Chatterjee and S. K. Sen. (*Phil. Mag.*, vol. 3, pp. 839–852; August, 1958.) A report is given of experimental investigations into the variations in contact resistance between potassium and aluminum surfaces when the junction is under the influence of an electrical stress.
- 537.32:546.23–17 1272
Thermoelectric Observations on Grey Selenium—J. I. Carasso and R. W. Pittman. (*Nature (London)*, vol. 182, p. 1011; October 11, 1958.) Experiments on spectrographically pure Se showed that the thermoelectric power of single crystals is independent of temperature while that of oriented polycrystals increases with temperature.
- 538.22 1273
The Influence of Hydrostatic Pressure on the Magnetizability of Au₂Mn.—K. H. v. Klitzing and J. Gielessen. (*Z. Physik*, vol. 150, pp. 409–414; March 10, 1958.)
- 538.22:537.311.3 1274
The Magnetic Susceptibility and Electrical Resistivity of some Transition-Metal Silicides—D. A. Robins. (*Phil. Mag.*, vol. 3, pp. 313–327; April, 1958.)
- 538.221 1275
Thermal Activation of Ferromagnetic Domains—F. D. Stacey. (*Proc. Phys. Soc.*, vol. 73, pp. 136–138; January 1, 1959.) By similar analysis to that used in the theory of electrical noise it is shown that the factor *C* in the expression *C* exp (–*E*/*kT*) for the probability of energy *E* being available to overcome potential barriers in domain wall movements, is proportional to the absolute temperature *T*.
- 538.221 1276
Orientational Superstructure arising from Mechanical Deformation of an Fe-Ni Alloy—R. Vergne. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 197–200; July 16, 1958.) Samples of an Fe-Ni alloy under tensile stress were heated above the Curie point for 75 min. and then quenched. The energy of magnetization subsequently observed was found to be a function of the tension.
- 538.221:538.632 1277
Theory of the Hall Effect in Ferromagnetic Substances—J. M. Luttinger. (*Phys. Rev.*, vol. 112, pp. 739–751; November 1, 1958.) The Hall effect is computed on the basis of a simple model making use of the transport theory of Kohn & Luttinger (1105 of 1958). The calculation is rigorous, but assumes a slowly varying scattering potential, a simple band, and very few conduction electrons. None of these assumptions is very realistic, but they enable a discussion to be given of the types of contribution which can occur.
- 538.221:538.65 1278
The Damping of High-Frequency Elastic Waves in Cubic Ferromagnetic Single Crystals—G. Simon. (*Ann. Phys., Lpz.*, vol. 1, Nos. 1–3, pp. 23–35; 1958.) The inhomogeneity of magnetization due to domain formation is shown to give rise to a damping maximum,

when the depth of skin-effect penetration becomes comparable with the mean dimensions of the domains.

538.221:539.23:538.632 1279
Ferromagnetic Thin Films. Hall Effect in Thin Films of Nickel—G. Goureaux, P. Huet and A. Colombani. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 189–193; July 16, 1958.)

538.221:539.23:538.632 1280
Measurements of the Hall Effect in Ferromagnetic Nickel Films—L. Reimer. (*Z. Physik*, vol. 150, pp. 277–286; February 14, 1958.) Results obtained with vapor-deposited films of thickness 20–2000 Å are discussed.

538.211:539.234:538.61 1281
The Observation of Weiss Domains in Polycrystalline Material using the Magnified Magneto-optical Kerr Rotation—J. Kranz and W. Drechsel. (*Z. Physik*, vol. 150, pp. 632–639; April 8, 1958.) Application of the technique described in 841 of 1957 (Kranz) to observation of the specimen at an angle of about 55°.

538.221:621.3.042.15:538.662 1282
Investigations on the Temperature Coefficient of the Initial Permeability of Dust Core Materials—H. Henniger. (*Nachr. Tech.*, vol. 8, pp. 66–75; February, 1958.) Test equipment is described and the results of measurements on cores prepared from various materials are analyzed in order to improve manufacturing techniques.

538.221:[621.318.124+621.318.134] 1283
Domain Configurations on Ferrites—D. J. Craik and P. M. Griffiths. (*Proc. Phys. Soc.*, vol. 73, pp. 1–13; January 1, 1959, plates.) The domain structures of Mn, MnZn, Ni, Cu and Ba ferrite crystals were examined by conventional and modified Bitter figure techniques. The Mn, Ni and Ba ferrites gave patterns as expected from their known magnetic constants. The Cu ferrite patterns could not be interpreted exactly. The MnZn ferrite gave anomalous patterns considered to be due to a strain-induced uniaxial anisotropy.

538.221:621.318.124 1284
Magnetic Materials with Perminvar Effect—A. v. Kienlin. (*Z. angew. Phys.*, vol. 9, October and December, 1957.) [Part I—Contribution to the Understanding of the Perminvar Effect and of its Dependence on Temperature (pp. 520–526).]
 Part 2—The Significance of Composition for the Perminvar Effect in Ferrites with Low Cobalt Content (pp. 631–640).

538.221:621.318.134 1285
Phenomenological Theory of Kinetic Processes in Ferromagnetic Dielectrics—M. I. Kaganov and V. M. Tsukernik. (*Zh. Eksp. Teor. Fiz.*, vol. 34, pp. 1610–1618; June, 1958.) Calculation of the relaxation time resulting from the interaction of spin waves.

538.221:621.318.134 1286
Electromagnetic Fields in Ferrite Ellipsoids—R. A. Waldron. (*Brit. J. appl. Phys.*, vol. 10, pp. 20–22; January, 1959.) The field in a ferrite immersed in a uniform magnetic field is calculated with the aid of the permeability tensor for a number of ferrite shapes.

538.221:621.318.134 1287
Effect of Oxygen Pressure on Microstructure and Coercive Force of Magnesium Ferrite—R. E. Carter. (*J. Amer. Ceram. Soc.*, vol. 41, pp. 545–550; December, 1958.)

538.221:621.318.134 1288
Dispersion in the Complex Transverse Susceptibility of Lithium Ferrite in the Frequency

Range 10–10,000 mc—F. Voigt. (*Ann. Phys., Lpz.*, vol. 1, Nos. 1–3, pp. 86–101; 1958.) The results of measurements on LiFe₂O₄ ring cores are discussed.

538.211:621.318.134 1289
The Growth of Single Crystals of Magnetic Garnets—J. W. Nielson and E. F. Dearborn. (*J. Phys. Chem. Solids*, vol. 5, pp. 202–207; May, 1958.) Crystals of magnetic garnets, Y₃Fe₅O₁₂, Sm₃Fe₅O₁₂, Er₃Fe₅O₁₂ and Gd₃Fe₅O₁₂ suitable for many research purposes have been grown from iron-rich melts of lead oxide.

538.221:621.318.134 1290
Interpretation of the Magnetic Properties of Erbium Garnet in which the Al³⁺ and Cr³⁺ Ions have been Substituted for Fe³⁺ Ions—G. Villers, J. Loriers and R. Pauthenet. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 587–590; August 4, 1958.)

538.221:621.318.134 1291
Initial and Remanent Permeability Spectra of Yttrium Iron Garnet—R. D. Harrington and A. L. Rasmussen. (*Proc. IRE*, vol. 47, p. 98; January, 1959.)

538.221:621.318.134:534.232 1292
Effect of Divalent Ion Substitutions on the Magnetomechanical Properties of Nickel Ferrite—S. F. Ferebee and C. M. Davis, Jr. (*J. Acoust. Soc. Amer.*, vol. 30, pp. 747–750; August, 1958.) Magnetostrictive properties of Co-substituted Ni ferrous ferrite are investigated to determine its suitability for use in electromechanical transducers.

538.221:621.318.134:621.317.411 1293
Measurements of the Permeability Tensor for "Ferrocube" at 24,000 mc—G. G. Robbrecht and J. L. Verhaeghe. (*Nature, London*, vol. 182, p. 1080; October 18, 1958.) Measurements have been made on spherical and disk samples of ferrocube IVB in a TE₁₁₂ resonant cavity excited by circularly polarized waves.

538.23:538.221 1294
The Influence of Temperature in the Creep of Asymmetric Hysteresis Cycles—Nyugen-Van-Dang. (*C. Rend. Acad. Sci., Paris*, vol. 247, pp. 56–59; July 7, 1958.) Experiments at 290°, 77° and 20°K show that temperature has little influence on creep.

539.234:546.72 1295
Stability of Evaporated Films—M. W. Roberts. (*Nature, London*, vol. 182, pp. 1151–1152; October 25, 1958.) Thin Fe films show a rapid reduction of surface area to 23 per cent of the original value as a result of the chemisorption of O₂.

621.315.61–41 1296
Tape and Film Insulation for Electronic Equipment—G. Sideris. (*Electronics*, vol. 32, pp. 42–43; January 2, 1959.) A brief review of the properties and applications of some of the newer forms of insulating material.

MATHEMATICS

512:621.316.5 1297
The Solution of Equations in Switching Algebra—H. Zemanek. (*Arch. elekt. Übertragung*, vol. 12, pp. 35–44; January, 1958.)

517.512.2 1298
Fourier Analysis, a New Numerical Method—L. Hyvärinen. (*Acta Poly. scand.*, No. 248, Ma 2, 19 pp.; 1958.) Evaluation is facilitated by the use of templates.

MEASUREMENTS AND TEST GEAR

621.3.018.41(083.74) 1299
A Very Precise Short-Period Comparator for 100-kc Frequency Standards—A. M.

Thompson and R. W. Archer. (*Proc. IRE*, Pt. B, vol. 106, pp. 61–64; January, 1959.) The instrument described uses addition, amplification and phase detection; an accuracy of frequency comparison within 1 part in 10¹³ has been obtained.

621.3.018.41(083.74) 1300
A Microwave Frequency Standard—B. H. L. James and M. T. Stockford. (*Electronic Eng.*, vol. 31, pp. 2–7 and 82–87; January and February, 1959.) Reference frequencies within the range 7–20 mc are derived as harmonics from a 100-kc temperature-controlled crystal oscillator with a short-term stability better than 1 in 10⁷. Several lower-frequency references are provided by frequency division down to 50 cps. Full descriptions of the circuits are given.

621.3.018.41(083.74):529.786:525.35 1301
The Random Variation in the Speed of Rotation of the Earth—A. Stoyko and N. Stoyko. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 182–185; July 16, 1958.) A comparison of U.T.2 with the Cs resonator of the National Physical Laboratory is made for 1955–1958 based on the work of Essen *et al.* (3195 of 1958).

621.3.089.6:061.6 1302
Services and Facilities of the Electronic Calibration Center—(*Tech. Bull. Nat. Bur. Stand.*, vol. 42, pp. 223–229; November, 1958.) Details of the department recently opened at Boulder, Colorado, for the calibration of laboratory standards in terms of the national standards maintained by the N.B.S. See 858 of 1958.

621.317(083.7):061.3 1303
Conference on Electronic Standards and Measurements—(*Tech. Bull. Nat. Bur. Stand.*, vol. 42, pp. 209–217; November, 1958.) Report on the conference held at Boulder, Colorado, August 13–15, 1958. The technical program consisted of the following sessions:

- The Relationship of Standards to Physical Constants;
- Frequency and Time Interval Standards;
- Direct-Current and Low-Frequency Standards;
- Radio-Frequency Standards;
- Microwave Standards;
- Organization and Operation of Standards Laboratories.

621.317.3:621.314.7 1304
Measurement of Transistor Characteristics in the 3–250-Megacycle Frequency Range—J. H. O'Connell and T. M. Scott. (*RCA Rev.*, vol. 19, pp. 598–616; December, 1958.) Description of apparatus and techniques developed for the measurement of the input, output and unilateralization impedances, transconductance and current gain.

621.317.335 1305
Simple Apparatus for Measuring Dielectric Constants and Losses from 10 cps to 50 kc—J. C. S. Richards. (*J. Sci. Instr.*, vol. 36, pp. 22–23; January, 1959.) A bridge circuit with a battery-operated detector is described in which no coupling transformer or Wagner earth is used. Measurements are made to an accuracy within about ±1 per cent.

621.317.336:537.311.62 1306
Surface Impedance—J. C. Anderson. (*Electronic Radio Eng.*, vol. 36, pp. 56–60; February, 1959.) The surface impedance of a conducting material is measured at VHF by the use of a coaxial transmission line terminated by a disk of the material. Theory, experimental method, and results for pure Ni are given. A calibrated RF meter is not required.

- 621.372.2:621.317.34 1307
Novel Method for Measuring Impedances on Surface-Wave Transmission Lines—G. Schulten. (Proc. IRE, vol. 47, pp. 76-77; January, 1959.)
- 621.317.34.029.64:621.317.755 1308
A Microwave Reflectometer Display System—J. C. Dix and M. Sherry. (*Electronic Eng.*, vol. 31, pp. 24-29; January 1959.) The instrument described gives a continuous display of the matching of an X-band waveguide component over the frequency band 7500-11,000 mc.
- 621.317.35:621.396.822 1309
Spectrum Analysis of Random Noise Generators—L. G. Polimerou. (*Commun. and Electronics*, pp. 672-679; November, 1958.) A method is described using a selective filter and a variable chopper-carrier. Possible limitations and errors are discussed.
- 621.317.351:621.317.619:621.396/.397 1310
A Method of Measurement and Oscillographic Recording of Phase Characteristics for Wide-Band Transmission Systems in the Range 0.1-10 mc—W. Rietz. (*Z. angew. Phys.*, vol. 9, pp. 489-495; October, 1957.) The equipment described [see also 219 of 1958 (Kroebel and Wegner)] has a phase error of $\pm 0.05^\circ$ at 13 kc for input amplitude changes of 20 db.
- 621.317.361:529.786 1311
Efficiency of Frequency Measurements with an Atomic Clock—M. Peter and M. W. P. Strandberg. (Proc. IRE, vol. 47, pp. 92-93; January, 1959.)
- 621.317.382.029.64:538.632 1312
Power Measurement at 4 Gc/s by the Application of the Hall Effect in a Semiconductor—L. M. Stephenson and H. E. M. Barlow. (Proc. IRE, Part B, vol. 106, pp. 27-30; January, 1959.) Description of an instrument developed from an earlier design [see 1488 of 1956 (Barlow and Stephenson)]. The range at 4 mc is 30 mw-20 w with a ± 3 per cent error at any swr between unity and 0.1. Only about 3.4 per cent of the power measured is adsorbed. Measurements made using a crystal within the power-carrying waveguide are also described.
- 621.317.44 1313
A New Method of Measuring Susceptibility—C. Hilsun and A. C. Rose-Innes. (*Nature (London)*, vol. 182, p. 1082; October 18, 1958.) A magnetic analogue of a Wheatstone bridge with a permanent magnet as the source of magnetomotive force is described. It may be used to measure a very weak susceptibility such as that of deionized water.
- 621.317.44:538.248:539.23 1314
The Measurement of Residual Magnetism of Thin Films—S. Yamaguchi. (*Naturwiss.*, vol. 45, p. 205; May, 1958.) A method based on the Lorentz effect is described.
- 621.317.7.089.6:621.372.56.029.64 1315
The Calibration of Microwave Attenuators by an Absolute Method—E. Laverick. (See 1123.)
- 621.317.733.011.4 1316
The Influence of the Leads on the Capacitance of Capacitors—G. Zickner and W. Wiessner. (*Ann. Phys., Lpz.*, vol. 1, Nos. 1-3, pp. 70-85; 1958.) Error sources in capacitance-bridge measurements by substitution methods are discussed. The elimination of the error due to the separation of the lead from the capacitor is described, and the magnitude of the effect is investigated by means of measurements on two- and three-plate capacitors.
- 621.317.755 1317
Simple, High-Sweep-Speed, Single-Stroke Oscilloscope—W. P. Baker. (*J. Sci. Instr.*, vol. 36, pp. 30-31; January, 1959.) The advantages of applying the accelerating voltage to the cathode ray tube in the form of a long-tailed pulse instead of using a steady dc source and grid bias are noted, and circuit arrangements are shown.
- 621.317.755:621.372.5 1318
Image Distortion by RC Quadripoles of a Cathode-Ray Oscilloscope and its Correction at Low Frequencies—H. Wittke. (*Elektron. Rundschau*, vol. 12, pp. 89-93; March, 1958.) Correction formulae are derived for reconstituting the original signal waveform from the output of a RC network. Difficulties of correcting circuit distortion in cathode ray oscilloscopes are outlined.
- 621.317.77.029.4 1319
A Low-Frequency Phasemeter—N. Hambley. (*Electronic Eng.*, vol. 31, pp. 13-15; January, 1959.) The phasemeter described gives a quick and accurate measurement in the frequency range 1-100 cps.
- 621.317.784 1320
A Precision Thermoelectric Wattmeter for Power and Audio Frequencies—J. J. Hill. (Proc. IRE, Part B, vol. 106, p. 64; January, 1959.) Discussion on 1226 of 1958.
- 621.317.79.029.6:535.322.4 1321
Compact Microwave Refractometer for Use in Small Aircraft—M. C. Thompson, Jr. and M. J. Vetter. (*Rev. Sci. Instr.*, vol. 29, pp. 1093-1096; December, 1958.) The refractometer operates on the principle developed by Birnbaum (see *Ibid.*, vol. 21, No. 2, pp. 169-176; February, 1950) using a frequency-modulated klystron oscillator. A two-channel recording system is described which provides a simultaneous indication of profile and turbulence data. The rms noise levels are approximately 0.05 and 0.01 N units respectively.
- 621.317.794.029.62/.63 1322
An Aperiodic Barretter Probe for Power Measurements at $f=30$ to 1500 mc and its Application—H. Rieck. (*Nachr. Tech.*, vol. 8, pp. 50-55; February, 1958.) A balanced twin-barretter mount is described for use as a bolometer at powers up to 10 mw.
- 621.317.799:621.397.6 1323
A Television Waveform Generator using Transistors—Rozner. (See 1373.)
- 621.317.799:621.397.62:535.623 1324
A One-Tube Crystal-Filter Reference Generator for Colour TV Receivers—R. H. Rausch and T. T. True. (IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS, vol. BTR-3, pp. 2-7; June, 1957.)

OTHER APPLICATIONS OF RADIO AND ELECTRONICS

- 611.84.001.57:621.372.029.64 1325
Characteristics of a Model Retinal Receptor Studied at Microwave Frequencies—J. M. Enoch and G. A. Fry. (*J. Opt. Soc. Amer.*, vol. 48, pp. 899-911; December, 1958.) A simplified microwave model of the human eye is described. Tests were carried out at 2.42 and 3.20 cm λ .
- 612.3:621.3.083.7 1326
New Method of Exploring the Digestive System by means of an Ingestible Radio Capsule—M. Marchal and M. T. Marchal. (*C. Rend. Acad. Sci. Paris*, vol. 246, pp. 3519-3520; June, 1958.)
- 616.621.3.083.7:621.314.7 1327
Experiments with an Ingestible Intestinal Transmitter—M. v. Ardenne and H. B. Sprung,

(*Naturwiss.*, vol. 45, pp. 154-155; April, 1958.) The transmitter, which is 26 mm long and 10 mm in diameter, operates at 1.5 kc, this frequency can increase up to 5 kc in accordance with pressure. See also 3611 of 1957 (Mackay and Jacobson).

- 621.384.62 1328
Radio-Frequency Aspects of Electro-Nuclear Accelerators—A. F. Harvey. (*Proc. IEE*, Part B, vol. 106, pp. 43-57; January, 1959.) Survey of various machines with emphasis on RF aspects of design and performance. 176 references.

- 621.384.622.2 1329
Anomalous Attenuation in Linear Electron Accelerators—V. J. Vanhuysse. (*Nature (London)*, vol. 182, pp. 1081-1082; October 18, 1959.) Anomalous attenuation is explained on the basis of the "multipactor" effect.

- 621.385.833 1330
Emission Microscopy with Secondary Electrons (16-keV Primary Electrons)—W. Bayh. (*Z. Physik*, vol. 150, pp. 10-15; December 21, 1957.)

- 621.385.833 1331
Method of Obtaining Images in Electron Diffraction and Microscopy without Photographic Emulsion—J. J. Trillat and L. Tertian. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 582-584; August 4, 1958.) The photographic plate in an electron microscope or diffractograph is replaced by a sheet of insulating material which is treated with a fine positively charged powder to reveal the image.

- 621.387.4 1332
The Spreading of the Discharge in Self-Quenching Counter Tubes: Part 2—E. Huster and E. Ziegler. (*Z. Physik*, vol. 149, pp. 583-593; November 25, 1957.) Experimental results are interpreted on the basis of the theory suggested in Part 1 (1561 of 1957).

PROPAGATION OF WAVES

- 621.396.11+621.396.67 1333
International Colloquium on Current Problems in Radio Wave Propagation, Paris, 17th-21st Sept. 1956—(*Ann. Télécommun.*, vol. 12, pp. 140-216; May, 1957.)
 Ionospheric Propagation Problems:
 a) **Theory of the Interaction between Two Waves in an Ionized Gas**—M. Bayet, J. L. Delcroix, and J. F. Denisse (pp. 140-141).
 b) **Graphical Method for the Routine Determination of the Vertical Distribution of Electron Density in Ionospheric Layers from Frequency-Sweep Recordings**—W. Becker (pp. 141-145). See 115 of 1957.
 c) **Study of Ionospheric Propagation by Hyperfine Recording of the Field of Standard Transmitters**—J. Bouchard (pp. 146-150).
 d) **The Self-Demodulation of Waves in the Ionosphere**—M. Cutolo (pp. 150-155).
 e) **Pulse Propagation Experiments at Oblique Incidence**—W. Dieninger (pp. 155-159).
 f) **Ionospheric Wind Observations**—F. Harnischmacher (pp. 159-161).
 g) **The Transmission of Solar Radio Bursts**—C. de Jager (pp. 161-164).
 h) **A Proposal for Improving the Accuracy of Ionospheric Forecasts**—C. M. Minnis (pp. 164-168).
 i) **Some Current Problems in Ionospheric Forecasting**—K. Rawer (pp. 169-172).
 j) **The Influence of the Auroral Zone on Radio Communications**—J. Rybner and E. Ungstrup (pp. 172-173).
 k) **On the Apparent Velocity of Short Waves**—A. Stoyko (pp. 173-174).
 l) **Influence of Propagation on Radio-navigation Systems based on Phase Measurement**—E. Vassy (pp. 175-176).

- Radio Meteorology:
 m) **Influence of Meteorological Factors on the Propagation of Ultra High Frequencies**—M. Anastassiades and L. Carapiperis (pp. 177-180).
 n) **The Refractive Index of Humid Air for Microwaves**—A. Battaglia, G. Boudouris and A. Gozzini (pp. 181-184).
 o) **The Reflection of V.H.F. Waves by Sharp Variations of Humidity in the Troposphere**—P. Beckmann (pp. 184-186).
 p) **Study of the Interaction of Two Adjacent Spheres Placed in an Electromagnetic Field**—J. Mevel (pp. 186-188).
 q) **Influence of Frontal Discontinuities on the Propagation of Decimetric and Centimetric Waves**—P. Misme (pp. 189-194).
 Influence of Irregular Surfaces; Aerial Problems:
 r) **New Types of Aerial for Long-Range Radio Links**—G. Broussaud (pp. 195-197).
 s) **Optimum Apertures of Aerials for Randomly Distributed Fields**—W. C. Hoffman (pp. 198-200).
 t) **Measurements of Gain across Mountainous Obstacles in California (U.S.A.)**—R. E. Lacy (pp. 200-204).
 u) **The Centipede Aerial**—D. K. Reynolds, J. Lignon and P. A. Szente (pp. 205-210).
 v) **The Reflection of a Plane Electromagnetic Wave by a Perfectly Conducting Rough Surface**—J. P. Schouten and A. T. De Hoop (pp. 211-214).
 w) **Calculation of Reflection Coefficients and Differential Scattering Cross-Sections for Irregular Surfaces**—V. Twersky (pp. 214-216).
- 621.396.11 1334
The Sommerfeld Ground Wave—B. Kockel. (*Ann. Phys., Lpz.*, vol. 1, pp. 145-156; 1958.) A discussion of Sommerfeld's treatment of surface-wave propagation.
- 621.396.11 1335
Some Observations on Scattering by Turbulent Inhomogeneities—M. Balsler. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-5, pp. 383-390; October, 1957.)
- 621.396.11:551.510.52 1336
Propagation at Great Heights in the Atmosphere—G. Millington. (*Marconi Rev.*, vol. 21, No. 131, pp. 143-160; 4th Quarter 1958.) A theoretical study is made of tropospheric propagation through an atmosphere which is standard at small heights, but in which the refractive index approaches unity asymptotically at great heights. This leads, in the geometrical-optics region, to a reduced horizon distance and in the diffraction region to a reduction in height-gain.
- 621.396.11:551.510.52 1337
Diurnal Influences in Tropospheric Propagation—M. W. Gough. (*Marconi Rev.*, vol. 21, No. 131, pp. 198-212; 4th Quarter 1958.) Signal-strength recordings at 80 mc combined with meteorological soundings, have been maintained for six months over a 137-km non-optical path close to the Persian Gulf. Results show that there are three main types of propagation. Examples are given of the association between high signal levels and surface ducts, particularly of the intense nocturnal type, most of which are capable of trapping the first transmission mode.
- 621.396.11:551.510.535 1338
The Propagation of Electromagnetic TE(H) Waves Generated by a Horizontal Current-Carrying Conductor Loop in the Cavity between Two Concentric Spheres—C. Rösner. (*Z. angew. Phys.*, vol. 9, pp. 448-453; September, 1957.) The system earth-air-ionosphere is considered.
- 621.396.11:551.510.535 1339
Ionospheric Propagation: the C.C.I.R. at Warsaw 1956 and the Method of Professor Gea—R. Gea Sacasa. (*Rev. Telecomun. (Madrid)*, vol. 12, pp. 13-20; September, 1958.) A further comparison of I.F.R.B. predictions with those made by Gea's method. See also 3952 of 1958.
- 621.396.11:551.510.535 1340
The Determination of Skip Distances by Back-Scatter Sounding—I. Ranzi and A. Porreca. (*Note Recensioni Notiz.*, vol. 7, pp. 294-299; May/June, 1958. English translation, pp. 300-305.) Back-scatter minimum equivalent paths for 22.3 mc are compared with vertical-sounding data obtained about half-way along the path at the time when reception of the 21.64-mc transmission from Daventry fails in Rome. Skip distances determined by back-scatter sounding agree with observations; those deduced from vertical soundings appear to be excessive. See also 1197 above.
- 621.396.11:551.510.535 1341
Determination of H.F. Sky-Wave Absorption—G. L. Pucillo. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-5, pp. 314-315; July, 1957.) A graphical method is described.
- 621.396.11:551.510.535 1342
Characteristics of F₂-Layer Multiple Reflections: Part 2—Y. Nomura, S. Katano, Y. Echizenya, R. Nishizaki, K. Ishizawa, T. Mori and T. Kokaku. (*J. Radio Res. Labs., Japan*, vol. 5, pp. 295-302; October, 1958.) Focusing effects due to the curvature of the ionosphere are considered to be the cause of multiple reflections. For Part 1 see 894 of 1956 (Echizenya *et al.*).
- 621.396.11:551.510.535 1343
Round-the-World Echoes—G. A. Isted. (*Marconi Rev.*, vol. 21, pp. 173-183; 4th Quarter 1958.) Information on round-the-world echoes, including those attributed to transmissions from satellite 1957a, is reviewed. Investigations show that circulating signals can be propagated through areas of low ionization density, and it is suggested that the mechanism is one in which transmission is made possible by launching a wave at an extremely small angle of incidence against a tilted layer, to produce a succession of reflections which are confined solely to the ionosphere.
- 621.396.11:551.510.535:523.5 1344
Meteor Activity as a Factor in Ionospheric Scatter Propagation—G. A. Isted. (*Marconi Rev.*, vol. 21, pp. 161-172; 4th Quarter 1958.) An analysis has been made of the results of experimental work on transient echoes and of the characteristics of the received signal. Weather-cloud discharges or some other mechanism may prove to have a greater influence on forward scatter propagation than meteor activity.
- 621.396.11:551.510.535:629.19 1345
Measurement of the Scattering Matrix with an Intervening Ionosphere—H. Brysk. (*Commun. and Electronics*, pp. 611-612; November, 1958.) An expression for the received field is derived for radar scattering from earth satellites when Faraday rotation occurs.
- 621.396.11.029.63 1346
Propagation Measurements at 858 mc over Paths up to 585 km—G. C. Rider. (*Marconi Rev.*, vol. 21, pp. 184-197; 4th Quarter 1958.) Results are given of measurements of signal strength and fading parameters which have been accompanied by polarization, height-gain and space-diversity tests, and an investigation of multipath time delays. These results give an attenuation ratio of 0.103 db/km, show no dependence of fading-rate upon distance and diversity distances larger than expected.
- 621.396.11.029.64:621.396.677 1347
Some Observations of Antenna-Beam Distortion in Trans-horizon Propagation—A. T. Waterman, Jr., N. H. Bryant and R. E. Miller. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-5, pp. 260-266; July, 1957. Abstract, *Proc. IRE*, vol. 45, p. 1759; December, 1957.)
- ### RECEPTION
- 621.396.62:621.376.3 1348
The R.F. Protection Ratios required for Modern V.H.F. F.M. Receivers—B. Gramatke, R. Netzband and E. Paulsen. (*Rundfunktech. Mitt.*, vol. 2, pp. 41-53; April, 1958.) Subjective and objective methods of measuring interference on commercial receivers are described. Measurements were made for transmitters modulated with different programs and for transmitters radiating the same program. The reduction of protection ratios for shared-channel operation is discussed.
- 621.396.62:621.314.7 1349
Special Circuits for Transistor Receivers—W. E. Sheehan and W. H. Ryer. (*Electronics*, vol. 32, pp. 56-57; January 9, 1959.) Design data on four portable transistors receiver circuits are summarized in tabular form.
- 621.396.62:621.376.3 1350
Dynamic Trap Captures Weak F.M. Signals—E. J. Baghdady and G. J. Rubissow. (*Electronics*, vol. 32, pp. 64-66; January 9, 1959.) The use is described of a high-Q trap in a reactance-valve circuit for the suppression of a stronger undesired signal.
- 621.396.81.029.62 1351
220 Mc/s Radio-Wave Reception at 700-1000 Miles—L. A. Ames and T. F. Rogers. (*Proc. IRE*, vol. 47, p. 86; January, 1959.) Fields measured on the ground and in the air were higher than would be expected from theories of tropospheric scatter, and showed rapid fading. They are attributed to ionospheric scatter, and reasons are given for the increase in field strength above that estimated by extrapolation from lower-frequency measurements.
- 621.396.812.3 1352
Theoretical Considerations of A11-Teleggraphy Reception Disturbed by Fading—B. Betzenhammer. (*Arch. elekt. Übertragung*, vol. 12, pp. 81-90; February, 1958.)
- 621.396.82.029.64:621.396.96 1353
Radar Interference to Microwave Communication Services—R. D. Campbell. (*Commun. and Electronics*, pp. 717-722; November, 1958.) The radiated frequency spectrum of high-power radars has been studied and the effect of these emissions on microwave systems has been determined. Suggestions are made for avoiding interference, including the suppression of spurious radar frequencies.
- 621.396.821 1354
Average Power of Impulsive Atmospheric Radio Noise—S. V. C. Aiya. (*Proc. IRE*, vol. 47, p. 92; January, 1959.)
- 621.39 1355
Comprehensive Comparisons in the Planning of Telecommunication Systems—R. Krzyckewski. (*A.T.E. J.*, vol. 14, pp. 120-128; April, 1958.) A method is described for evaluating, by weighting, the relative merits of radio, coaxial-cable and balanced-pair carrier cable systems when applied to a hypothetical trunk telecommunication problem.

STATIONS AND COMMUNICATION SYSTEMS

- 621.391 1356
Sampling of Signals without D.C. Components—A. R. Billings. (*Electronic Radio Eng.*, vol. 36, pp. 70-72; February, 1959.) A signal contained in a finite frequency band can only be sampled unambiguously at frequencies in certain permitted bands. The Shannon-Hartley law restricting permissible sampling frequencies is generalized.
- 621.391 1357
Maximum Number of Signals of Fixed Total Energy among which a Discrimination to within ϵ can be made in the Presence of White Gaussian Noise—P. Béthoux. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 573-575; August 4, 1958.) A note on the determination of $n(E, \epsilon)$, the maximum number of different signals of energy $\leq E$ in a limited spectrum which it is possible to distinguish with an error probability $\leq \epsilon$.
- 621.391:621.396.3 1358
Notes on Error-Correcting Techniques: Part I—Efficiency of Single-Error-Correcting Codes with a Constant Bit Rate of Transmission—J. Dutka. (*RCA Rev.*, vol. 19, pp. 628-641; December, 1958.)
- 621.391:621.396.8 1359
On the "Minimum Loss Operation Time" for Short-Wave Communication: Part 2—I. Shibata. (*J. Radio Res. Labs., Japan*, vol. 5, pp. 335-340; October, 1958.) This concept is not only a useful one for prediction purposes, but it also allows the integration of all possible modes of propagation. The importance of the E_s layer is stressed. Part 1: 258 of January.
- 621.396.41:551.510.52(083.57) 1360
Tropospheric-Scatter System Design—L. P. Yeh. (*Commun. and Electronics*, pp. 707-716; November, 1958.) Generalized curves have been obtained from experimental and theoretical data for channel signal/noise ratio, scatter loss, terrain effects and fading. Calculations show that for a 200-mile path a s.s.b. system will be more reliable than a FM system. See also 1552 of 1958.
- 621.396.41:621.396.65 1361
Design of Tropospheric-Scatter Multichannel Telephone Links—B. Peroni and G. Bianconi. (*Note Recensioni Notiz.*, vol. 7, pp. 147-180; March/April, 1958.) A method of calculating the system parameters for FM radio links is given with design curves.
- 621.396.41:621.396.65 1362
New Techniques Allow Wider Use of Radio Carrier in Telephone Systems—E. V. Hird. (*Can. Electronics Eng.*, vol. 2, pp. 28-33; August, 1958.) Methods of high-frequency bridging, branching and interconnecting of multichannel R/T systems are described.
- 621.396.5:534.76 1363
A Compatible System of Stereo Transmission by F.M. Multiplex—M. G. Crosby. (*J. Audio Eng. Soc.*, vol. 6, pp. 70-73; April, 1958. Discussion.) A sum-and-difference technique is described in which an additive combination of the microphone outputs is transmitted on the main channel and a subtractive combination on the subcarrier channel. Advantages include a balanced program for the monaural listener and equal signal/noise ratios for the stereophonic system, with an effective gain for the subcarrier channel of 6 db.
- 621.396.65 1364
New Aspects in the Planning of Radio Links—G. Megla. (*Nachr. Tech.*, vol. 8, pp. 50-55; February, 1958.) Choice of route and range,

and the use of special aeriels are discussed on the basis of recent experimental findings.

SUBSIDIARY APPARATUS

- 621.3.078:621.318.3 1365
Electronic Magnetic-Field Stabilizer with Feedback—C. Fric. (*Compt. rend. Acad. Sci., Paris*, vol. 246, pp. 3602-3605; June 30, 1958.) Fluctuations in the magnetizing current of an electromagnet are integrated using a vibrator-amplifier in a circuit with a time constant of 10 ms, and the derived signal energizes a compensating coil.
- 621.311.69:537.311.33:535.215 1366
The Effect of Radiation on Silicon Solar-Energy Converters—J. J. Loferski and P. Rappaport. (*RCA Rev.*, vol. 19, pp. 536-554; December, 1958.) The solar cells are unaffected by irradiation with ultraviolet light and X rays. High-energy electrons, protons and α -particles produce a time decay in such characteristics as maximum power output, open-circuit voltage and short-circuit current. The lifetime of the cells, to 75 per cent of the initial power output, is estimated to be 10^6 years when exposed to the radiations assumed to exist at earth-satellite orbit heights.
- 621.314.63:621.3.014.36 1367
Forward Current Surge Failure in Semiconductor Rectifiers—F. E. Gentry. (*Commun. and Electronics*, pp. 746-750; November, 1958. Discussion, pp. 754-755.)
- 621.316.721.078:621.314.7 1368
A Semiconductor Current Limiter—R. M. Warner, Jr., W. H. Jackson, E. I. Doucette and H. A. Stone, Jr. (*PROC. IRE*, vol. 46, pp. 44-56; January, 1959.) A constant-current device is described which uses the field effect at a junction. The reverse bias applied controls the conductance of a thin sheet of semiconductor. Various models and a particular design are discussed and some applications to switching and waveform generation are given.
- 621.316.712.078:621.385.3 1369
Compensation for Fluctuations in the Anode Current of a Valve due to Heating—C. Curie and Y. Descamps. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 278-280; July 21, 1958.) In the method described, an antiphase rectified voltage proportional to heater current variations is applied to the control grid of the valve to stabilize the anode current.
- TELEVISION AND PHOTOTELEGRAPHY
- 621.397.331.2:771.35 1370
Zoom Lenses for Closed-Circuit Television—F. G. Back. (*J. Soc. Mot. Pict. Telev. Eng.*, vol. 67, pp. 598-600; September, 1958. Discussion.) The use of zoom lenses is described and the optical and mechanical properties of various lenses are compared.
- 621.397.5:535.623:778.5 1371
Colour TV Recording on Black and White Lenticular Film—J. M. Brumbaugh, E. D. Goodale and R. D. Kell. (*IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS*, vol. BTR-3, pp. 71-75. Abstract, *PROC. IRE*, vol. 46, p. 383; January, 1958.) See also 2562 of 1958 (Kingslake).
- 621.397.6:535.623 1372
Colour Television Experiments—N. Mayer. (*Rundfunktech. Mitt.*, vol. 2, pp. 75-85; April, 1958.) Test equipment for the N.T.S.C. system is described.
- 621.397.6:621.317.799 1373
A Television Waveform Generator using Transistors—F. Rozner. (*Electronic Eng.*, vol. 31, pp. 8-12. January, 1959.) A small light-weight instrument designed for use with transportable television broadcasting equipment is described. Four outputs are provided:—composite synchronizing waveform, line drive, field drive and composite blanking waveform.
- 621.397.61:535.88 1374
The Television Transmission of Nontransparent Still Pictures by means of the Flying-Spot Scanning System (Television Episcopy)—R. Theile and F. Pilz. (*Rundfunktech. Mitt.*, vol. 2, pp. 54-63; April, 1958.) An experimental episcopy is described which has facilities for variable magnification and automatic focusing. Suggestions are made for the design of a television episcopy.
- 621.397.611.2:771.35 1375
Vidicon Camera Lenses—G. H. Cook. (*J. Soc. Mot. Pict. Telev. Eng.*, vol. 67, pp. 596-598; September, 1958.) The suitability of standard cinematographic lenses is discussed and a new range of lenses is described.
- 621.397.611.2:771.35 1376
A New Series of Lenses for Vidicon-Type Cameras—J. D. Hayes. (*J. Soc. Mot. Pict. Telev. Eng.*, vol. 67, pp. 593-595; September, 1958.)
- 621.397.62 1377
The Importance of the Video Receiver in Black-Level Transmission in Television—H. Grosskopf. (*Rundfunktech. Mitt.*, vol. 2, pp. 64-74; April 1958.) The importance of the picture monitor in controlling the black level in transmission is outlined and the need for standardization of characteristics and adjustments is stressed. Improvements in black-level reproduction are suggested taking account of the subjective assessment of picture quality.
- 621.397.62 1378
Synchronous and Exalted-Carrier Detection in Television Receivers—J. Avins, T. Brady and F. Smith. (*IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS*, vol. BTR-4, pp. 15-23, February, 1958. Abstract, *PROC. IRE*, vol. 46, p. 801; April, 1958.)
- 621.397.62:535.623 1379
Magnetic Demodulators for Colour TV—M. Cooperman. (*Electronics*, vol. 32, pp. 56-58; January 2, 1959.) The basic design of a magnetic demodulator is given and its performance is discussed.
- 621.397.62:535.623:621.317.799 1380
A One-Tube Crystal-Filter Reference Generator for Colour TV Receivers—R. H. Rausch and T. T. True. (*IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS*, vol. BTR-3, No. 1, pp. 2-7; June, 1957.)
- 621.397.62:621.373.52 1381
Transistorized Television Vertical-Deflection System—W. F. Palmer and G. Schiess. (*IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS*, vol. BTR-3, No. 2, pp. 98-105; October, 1957. Abstract, *PROC. IRE*, vol. 46, No. 1, pp. 383-384; January, 1958.)
- 621.397.62:621.374.33 1382
A New Noise-Gated A.G.C. and Sync System for TV Receivers—(*IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS*, vol. BTR-3, June, 1957.)
Part 1—J. G. Spracklen & W. Stroh (pp. 28-31).
Part 2—G. C. Wood (pp. 32-34). See 2928 of 1957.
- 621.397.62:621.375.1 1383
Design Considerations of a Developmental U.H.F. Tuner using an R.F. Amplifier—J. B. Quirk. (*IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS*, vol. BTR-4, pp.

5-11; February, 1958). Investigation of the performance of a ceramic triode incorporated in an existing UHF tuner.

621.397.62:621.375.1 1384

A Constant-Input-Impedance R.F. Amplifier for V.H.F. Television Receivers—H. B. Yin and H. M. Wasson (IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS, vol. BTR-3, pp. 65-70; October, 1957. Abstract, PROC. IRE, vol. 46, p. 383, January, 1958.)

621.397.62:621.375.4 1385

Transistor Design for Picture I.F. Stages—R. J. Turner and P. Hermann. (IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS, vol. BTR-3, pp. 76-80; October, 1957. Abstract, PROC. IRE, vol. 46, p. 383; January, 1958.)

621.397.621:535.623:621.385.832 1386

Techniques of Colour Purity Adjustments in Receivers employing the "Apple" Cathode-Ray Tube—R. C. Moore, A. Hopengarten and P. G. Wolfe. (IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS, vol. BTR-3, pp. 23-27; June, 1957.)

621.397.621:621.318.4:621.373.43 1387

A Line Transformer with Tuned High-Voltage Winding—H. Reker. (*Nachrichtentech. Z.*, vol. 11, pp. 147-153; March, 1958.) A method of preventing parasitic oscillation in line-deflection transformers is described. See 3673 of 1957 (Beauchamp).

621.397.621:621.397.82 1388

The Synchronization Characteristics of Television Receivers in the Presence of Interference—E. Lüdike. (*Arch. elekt. Übertragung*, vol. 12, pp. 8-14; January, 1958.) Methods of reducing the susceptibility to noise of the sweep control systems in television receivers are described.

621.397.6212:621.385.032.263 1389

Drive Factor and Gamma of Conventional Kinescope Guns—Gold and Schwartz. (See 1417.)

621.397.621.2001.4 1390

A Method of Measuring the Optical Sine-Wave Spatial Spectrum of Television Image Display Devices—O. H. Schade. (*J. Soc. Mot. Pict. Telev. Eng.*, vol. 67, pp. 561-566; September, 1958.) The resolution characteristic of kinescopes has been measured with electrically generated sine-wave patterns giving consistent results with varying beam current. Measurements are independent of phosphor decay time.

621.397.8:535.623 1391

On the Quality of Colour-Television Images and the Perception of Colour Detail—O. H. Schade, Sr. (*RCA Rev.*, vol. 19, pp. 495-535; December, 1958.) The contrast range and color saturation of commercial color kinescopes in the N.T.S.C. system give a color space larger than that of commercial color motion pictures. More than 60% of the fine detail color information is transmitted and reproduced.

621.397.8:535.623 1392

Colour-Signal Distortions in Envelope Type of Second Detectors—B. D. Loughlin. (IRE TRANS. ON BROADCAST AND TELEVISION RECEIVERS, vol. BTR-3, pp. 81-93; October, 1957. Abstract, PROC. IRE, vol. 46, p. 383; January, 1958.)

621.397.813 1393

Test Pattern for Assessing the Quality of Television Images and Projection Systems—W. Kroebel. (*Naturwissenschaften*, vol. 45, pp.

205-206; May, 1958.) The pattern consists of a series of equidistant vertical grey wedges decreasing in contrast along the abscissa. It is designed for subjective assessment of picture quality.

621.397.9 1394

Industrial Television Installations—W. Mayer. (*Frequenz*, vol. 12, pp. 45-49; February, 1958.) Outline of the basic principles of design and operation.

621.397.9 1395

New Applications of Industrial Television—E. F. Spiegel. (*Frequenz*, vol. 12, pp. 33-38; February, 1958.) Accessories and special equipment of German manufacture are briefly described.

TRANSMISSION

621.396.61.001.4:621.396.3 1396

I.R.E. Technical Committee Report: Methods for Testing Radiotelegraph Transmitters (below 50 Mc/s)—(Proc. IRE, vol. 47, pp. 57-63; January, 1959.) Report 58 IRE 15. TR1.

621.396.712.2:621.396.66 1397

A New Method of Automatic Monitoring of Radio Broadcast Channels—F. Enkel. (*Nachrichtentech. Z.*, vol. 11, pp. 142-147; March, 1958.) An automatic system is described for the supervision of a transmission system by monitoring the radiated signal. Measurements of level, noise and distortion are carried out during natural breaks in modulation. Tests of nonlinear distortion are made using the interval signal, and facilities for recording interference on magnetic tape are provided.

TUBES AND THERMIONICS

621.314.63 1398

The Influence of Adjacent Connections on the D.C. and A.C. Characteristics of p-n Junctions—H. Beneking. (*Z. angew. Phys.*, vol. 9, pp. 626-631; December, 1957.) The effect of the external connections closely approaching the barrier layer is calculated; the result is in agreement with experimental findings.

621.314.7 1399

Transistors Reach for Higher Frequencies—A. L. Barry and P. J. Coppen. (*Can. Electronics Eng.*, vol. 2, pp. 19-27; August, 1958.) A review is given of progress in the development of amplifying devices for high-frequency operation.

621.314.7:621.317.3 1400

Measurement of Transistor Characteristics in the 3-250-Megacycle Frequency Range—O'Connell and Scott. (See 1304.)

621.314.7:621.318.57 1401

The Theory of the Switching Transistor—W. v. Münch. (*Z. angew. Phys.*, vol. 9, pp. 621-625; December, 1957.) Discussion of the characteristics of the special transistor described in 3244 of 1956 (Salow and v. Münch).

621.314.7.002.2 1402

Selective Electrolytic Etching of Germanium and Silicon Junction Transistor Structures—I. A. Lesk and R. E. Gonzalez. (*J. Electrochem. Soc.*, vol. 105, pp. 469-472; August, 1958.) Selective etching may be used to expose part of the base region and thereby facilitate attachment of the transistor base lead.

621.314.7.01 1403

On the Usefulness of Transconductance as a Transistor Parameter—H. L. Armstrong. (Proc. IRE, vol. 47, pp. 83-84; January, 1959.)

621.383.2 1404

Image Converters for Quantity Production and the Feasibility of Defining their Emission Gain Characteristic—P. Görlich, A. Krohs, H. J. Pohl and G. Zerst. (*Z. angew. Phys.*, vol. 9, pp. 561-566; November, 1957.) Three types of converter, two with Cs-Sb and one with caesium oxide cathode, and their applications are described.

621.383.4 1405

A Simple Photoelectric Amplifier—H. Oswald and H. Straubel. (*Z. angew. Phys.*, vol. 9, pp. 438-442; September, 1957.) A differential arrangement of two photoconducting cells in a bridge circuit is described. The two cells are formed by dividing a single CdS crystal after mounting.

621.383.4 1406

The Conduction Mechanism of CdS Sandwich Photocells—G. Ecker and J. Fassbender. (*Z. Phys.*, vol. 149, pp. 571-582; November 25, 1957.) A theory is developed based on a model of photoconductivity in semiconductors and compared with results of measurements on photocells.

621.383.4 1407

Behaviour of Lead Sulphide Photocells in the Ultraviolet—A. Smith and D. Dutton. (*J. Opt. Soc. Amer.*, vol. 48, pp. 1007-1009; December, 1958.) Spectral response was measured in the wavelength range 0.2-2.0 μ . The increase in quantum yield observed below 0.6 μ is attributed to electron multiplication by secondary internal emission.

621.383.4:535.371.07 1408

A Feedback Light-Amplifier Panel for Picture Storage—B. Kazan. (Proc. IRE, vol. 47, pp. 12-19, January, 1959.) A panel incorporating optical feedback is described in which separate phosphor layers having 40 elements per inch are used for viewing and feedback. The effect of varying the feedback factor is shown.

621.385.029.6 1409

Generation of Second Harmonic in a Velocity-Modulated Electron Beam of Finite Diameter—F. Paschke. (*RCA Rev.*, vol. 19, pp. 617-627; December, 1958.) The analysis, taking into account the beam fringe fields, shows that the amplitudes of the second-harmonic current and velocity do not vary periodically with distance.

621.385.029.6 1410

Development of High-Power Pulsed Klystrons for Practical Applications—M. Chodorow, E. L. Ginzton, J. Jasberg, J. V. Lebacqz and H. J. Shaw. (Proc. IRE, vol. 47, pp. 20-29; January, 1959.) The design procedure is described for sealed-off, tunable klystrons giving peak output powers of 1-2 mw and operating in the X, S and L frequency bands. The electron beam and RF structure are considered and cavity tuning methods appropriate to the three bands are discussed. The measured performance is shown by curves of power output, gain and efficiency.

621.385.029.6 1411

The Effect on Gain and Stability of a Travelling-Wave Valve of the Choice of Surface Resistance of its Attenuating Layer—W. Eichlin and G. Landauer. (*Nachrichtentech. Z.*, vol. 11, pp. 131-137; March, 1958.) Maximum attenuation in coaxial cylindrical attenuators with normal helix dimensions is obtained with surface resistance about 2.5 k Ω . Noncoaxial arrangements have attenuation maxima corresponding to lower surface resistances.

621.385.029.6:537.533:621.376.9 1412

A New Method for Pumping a Fast Space-Charge Wave—G. Wadd and R. Adler. (PROC. IRE, vol. 47, pp. 79–80; January, 1959.) A technique for low-noise amplification by interaction between a fast wave and an electron beam is illustrated by a mechanical analogy of a pendulum with periodically applied vertical forces.

621.385.029.6:621.376.23 1413

The Use of Beam Defocusing to Provide a Microwave Detector—P. S. Castro and J. S. Needle. (PROC. IRE, vol. 47, pp. 82–83; January, 1959.) A note on the detection action which occurs when a density-modulated or velocity-modulated electron beam is allowed to drift through a field-free space inside a cylindrical collector of resistive material. Wide-band characteristics are obtainable, with sensitivity comparable to that of a crystal diode. See also 2259 of 1956 (Mendel).

621.385.029.63 1414

Electrostatically Focused Travelling-Wave Tube—D. J. Blattner and F. E. Vaccaro. (*Electronics*, vol. 23, pp. 46–48; January 2, 1959.) Details are given of the construction and performance of a light-weight valve for operation at frequencies between about 2 and 3 kmc.

621.385.029.65 1415

Helix-Type Travelling-Wave Amplifier for

48 kmc/s—T. Miwa, M. Mishima and I. Yanaka. (PROC. IRE, vol. 47, pp. 89–90; January, 1959.)

621.385.032.26 1416

The Focal Length of a Diaphragm of Finite Aperture for Electron Beams with Finite Space Charge—K. Pöschl and W. Veith. (*Arch. elekt. Übertragung*, vol. 12, pp. 45–48; January, 1958.)

621.385.032.263:621.397.62.12 1417

Drive Factor and Gamma of Conventional Kinescope Guns—R. D. Gold and J. W. Schwartz. (*RCA Rev.*, vol. 19, pp. 564–583; December, 1958.) A new expression for the beam-current/drive-voltage relation is derived which gives values for gamma and drive factor that agree more closely with experimental results than the values derived from previous expressions.

621.385.2.032.213.1 1418

Surface Structure of Saturated-Diode Filaments—F. A. Benson and M. S. Seaman. (*Electronic Eng.*, vol. 31, pp. 40–41; January, 1959.)

621.385.3.029.63 1419

Design and Performance of a New Low-Noise Triode for Use up to 1000 mc/s—A. D. Williams and D. C. Gore. (PROC. IRE, vol. 106,

pp. 35–42; January, 1959.) A glass-base valve Type A2521 is described which is stable up to 1 mc. Measured noise factor in optimum conditions increases linearly from 6 db at 300 mc to 11.5 db at 900 mc. Mutual conductance is 15 ma/v and amplification factor is 60. Used as an oscillator the valve has a power output of 150 mw at 1250 mc.

MISCELLANEOUS

621.3-71:629.13 1420

Cooling Airborne Electronic Equipment—L. A. Williamson. (*Wireless World*, vol. 65, pp. 87–91; February, 1959.) Conditions under which equipment must operate in aircraft and the limitations of air-cooling are examined. Experimental chassis to facilitate cooling by liquids have been constructed using either an electroforming honeycomb process or roll-bonded ducted aluminum sheet. The choice of coolants and heat-exchangers is discussed and the advantages of the system are summarized.

621.396 1421

[I.E.E.] Radio and Telecommunication Section: Chairman's Address—G. Millington. (PROC. IRE, Part B, vol. 106, pp. 11–14; January, 1959.) Topics discussed include technical literature, "rigour" in mathematics, systems of notation, the limitations of statistics, and the use of "speculative reasoning."