# PROCEEDINGS OF THE IRE®

## Published Monthly by

## The Institute of Radio Engineers, Inc.

## CONTENTS

THE COVER—Blazing gases, erupting to heights of 200,000 miles or more above the surface of the sun, give dramatic evidence of the tremendous power and violence of the thermonuclear reaction going on within the sun. Here on earth, man has succeeded in producing a thermonuclear reaction on a much tinier scale—the hydrogen bomb. He is now hard at work trying to find a way of controlling this reaction so that it may be put to work as a useful, nonhazardous, inexhaustible source of energy. What this will mean to the future of mankind and the formidable technical problems that must first be surmounted is lucidly described in the article which starts on page 134 of this issue.

*Photo—High Altitude Observatory of the University of Colorado*

*Copyright: © 1957, by the Institute of Radio Engineers. Inc.*

# Yasujiro Niwa

Vice-President, 1957

Yasujiro Niwa was born in Matsuzaka City, Japan in April, 1893. He was graduated from the University of Tokyo in 1916. He received the bachelor of engineering degree in 1916 and the doctor of engineering degree in 1926 from the same university. In 1916 he joined the Electrotechnical Laboratories, Department of Communication, and did research on magnetic materials. The permeameter bearing his name and the calculation formula of inductances of square coils were contributions made during this period. In 1924, he moved to Nippon Electric Co., where he worked until 1947. At Nippon Electric he has been successively director of research, chief engineer, and finally managing director in charge of engineering. Since 1949 he has been President of Tokyo Electrical Engineering College. He is also a member of the Radio Regulatory Council, Radio Technical Council and the Standards Council of Japan.

Among Dr. Niwa's many achievements should be counted his contributions to the development of facsimile transmission, and radio and television broadcasting in Japan. His system of picture transmission has been widely adopted in Japan since the inauguration of facsimile transmission in 1928.

He was President of the Institute of Electrical Communication Engineers of Japan in 1942–1943, President of the Institute of Electrical Engineers of Japan in 1950–1951, President of the Japan Radio Society in 1941, and President of the Television Society in 1954–1956. He is presently the President of the Facsimile Society and the Television Research Committee.

Dr. Niwa has received many awards such as the Imperial Prize from the Japan Academy of Science in 1937, the Imperial Prize in 1930 and the Grand Prize in 1920 from the Japan Invention Association, the Medal of Merit from the Institute of Electrical Communication Engineers of Japan in 1942, the Asano Prize from the Institute of Electrical Engineers of Japan in 1929, the Order of Rising Sun from the Japanese Government in 1940, and the Medal of Cultural Merit from the Japan Broadcasting Association in 1956. He was elected a Fellow of the IRE in 1951, which cited him for "his leadership in radio engineering in Japan and contributions to vocational education."

# Poles and Zeros
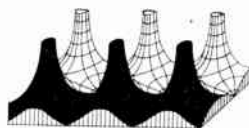


**Standards.** During 1956, we published in the PROCEEDINGS no fewer than ten new IRE Standards, in such diverse fields as feedback systems, microwave tubes, audio measurements, electron devices, facsimile, semiconductors, tr tubes, broadcast receivers, computers, and transistors. Ahead of us is much more of the same, including a new master index of all current IRE Standard Definitions of Terms to be published in the near future.

No other material published by IRE takes so much coordinated effort by so many workers and is, perhaps, so little appreciated by the membership at large. The IRE Standards perforce deal with established technology, old enough at least for competing methods and points of view to have appeared and to require mediation—so the appeal of novelty is seldom present. Moreover, much standardization work deals with fine points, precision of concept and word, which are largely lost on the casual reader. Everyone gives lip service to the IRE Standards, but how many read and use them? Everyone who ever worked on a standards subcommittee has had this nagging doubt: Is the work worth it?

It must be agreed that the IRE Standards Committee and the Technical Committees that feed it perform a function which is uniquely the duty and obligation of a professional society. One can imagine conventions being staged by others and there are many alternate outputs for technical papers. But the publication by a professional society of agreed-upon standards confers status and engenders respect that would be hard to come by if they were published by any other agency. The publication does not always settle the arguments that preceded standardization, but by presenting the majority point of view, it permits us all to get on with the work, using a common set of tools. This, we think, is the most important aspect of our Standards program; it settles issues, at least for a reasonable period of time, and provides a common basis of understanding and procedure.

Standards need not, and indeed should not, embalm obsolete concepts. A case in point is the Standards on Letter Symbols for Semiconductor Devices (July, 1956). Your editor was at first dismayed in reading this Standard to note its departure from standard practice in using a single symbol ($h$) to mean current ratio, voltage ratio, resistance, impedance, admittance and conductance, depending on a rather complicated set of subscripts for distinction. But on reflection, we realized that "standard practice" meant our own preconceived, and quite probably, obsolete notions. Transistors are complicated; they are better understood today, after eight years, than

vacuum tubes were after their first twenty-eight years. And the symbols reflect this better, more comprehensive understanding. So we withdrew our objection.

The late Eric Shipton said of mountains that they are important to mountaineers because they are there. So, too, a Standard is important because it is there, because its subject matter is precisely, if sometimes arbitrarily, standardized. Standards don't have to be based on older notions; they don't even have to be logical (although this helps). But, to be worth a tinker's dam, they have to be *accepted* and *used*. So, we urge every worker in the IRE vineyard to know the Standards in his field, and to use them in his thinking, in his lab work and in his publications.

**Upside Down.** Errors in printing, particularly in printing highly technical material, are so difficult to guard against that, despite careful reading by authors, reviewers and staff, occasional errors occur. In the first Color Television Issue (1951), a passage which originally read "red, green and blue" was printed "red, yellow and blue," because an editorial assistant was certain that the primary colors used in television were the same as those used in printing and so corrected what she thought was a slip of the typewriter. The difference between additive and subtractive primaries was thereafter explained, with gestures, to the staff, and the second Color Television Issue (1954) appeared without incident.

Once in a blue moon, our printer commits an error. The only such to come to mind recently concerns the cut, which appeared on page 1355 of the special issue on ferrites, October, 1956. After 2000 copies of this issue had been printed, the error was discovered, bells rang, the presses were stopped, the offending cut turned right side up, and the remaining 58,000 copies printed. After
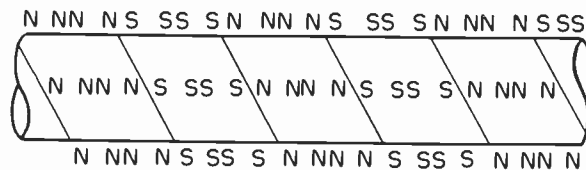


mature deliberation, it was decided to release the 2000 "error" copies, which are now well distributed throughout the IRE membership. To assist each member in determining whether he has a collector's item in his attic, we reprint the cut in question, right side up. No further comment, except to note that we have a very conscientious printer.—D.G.F.

# Radio and Electronics in Brazil

## ALLEN H. SCHOOLEY, FELLOW, IRE

In view of the recent formation of an IRE Section in Rio de Janeiro, Brazil, the following report on the status of radio engineering and electronics in Brazil is of special timeliness and interest. Allen H. Schooley, Vice-Chairman of the Washington, D. C. Section, is on leave from his position as Superintendent of the Electronics Division of the U. S. Naval Research Laboratory to assist the Brazilian Navy in establishing a Naval Research Institute.—*The Editor*

Brazil is larger than the United States and extends over nearly half the area of South America. Sixty million Portuguese-speaking Brazilians outnumber the total inhabitants of all other South American countries. Even so, the average population density is low; being 17 persons per square mile. This compares with 50 persons per square mile in the United States and 200 in France.

Brazil's climate is varied. Large areas of the Amazon equatorial plains have the steamy heat of the tropical jungle. Relatively few people live there. The great industrial cities of São Paulo, Rio de Janeiro, Belo Horizonte, and Porto Alegre have relatively mild even-temperature weather. These four busy, growing cities are the home of about six million Brazilians. There are also cool southern uplands where snow comes annually and wheat is an important crop.

The Brazilian Navy, with 20,000 miles of navigable rivers to supervise and a coastline to protect that is greater than the combined Atlantic and Pacific coasts of the United States, was first to see the need for radio communication in the vastness that is Brazil. In 1903, two years after Marconi transmitted the letter "S" across the Atlantic Ocean, the Navy initiated plans that resulted in the installation of a shore station in Rio de Janeiro and wireless stations aboard five ships in 1904 using German equipment. From this beginning, radio and television has become an important factor in Brazilian cultural and economic life. There are now about 3300 radio stations operating in all parts of Brazil. This includes 6 television stations, 750 broadcast stations, and 650 commercial stations. The remainder are distributed among railroad, post office, police, Army, Navy, Air Force and other federal and state governmental services.

The electronics manufacturing industry was started in 1933 with the formation of a company called "Cacique S/A" which started the commercial assembly of radio receivers in Brazil. This company name has not survived but a successor is today very active in the manufacture of radio, industrial, and medical electronic equipment. There are now about 65 companies engaged in the production of electronic components and systems and the need for importing these things is decreasing each year. Some of the principle materials and components being imported at present are transformer steel, litz wire, alnico, rf core material, plastic molding materials, selenium rectifiers, vibrators, dry batteries, three-speed phonograph motors, tubes, toggle switches, and electrolytic, ceramic trimmer, and high voltage capacitors. Most of the more elaborate electronic systems such as radars, sonars, television transmitting equipment, and computers are imported. The extent of the electronics industry in Brazil is indicated by the value of production figures for 1954. It is estimated that the equivalent value of radio receiver production approached 40 million dollars and the value of television receiver production reached about 15 million dollars. To attain these figures about 5 million dollars worth of imports were required.

Much of the radio and electronics equipment produced in Brazil is adapted from North American and European designs. The changes are usually directed toward using available components, simplifying help in servicing, and improving performance under conditions of high humidity and wide-line voltage fluctuations. Creative engineering work is severely limited by the number of electronic engineers available. A large percentage of the electronics engineers have had training abroad. Although there are about a dozen schools that have courses in physics and engineering, there are only a few that give appreciable electronics training. An excellent five-year course is conducted by the Electronics Division of the Instituto Tecnológico de Aeronáutica at São José dos Campos near São Paulo. This is a part of the Centro Técnico de Aeronáutica that has been established by the Brazilian Air Force since World War II. The students are selected by competitive examinations and pay no fees, even for board or room, and in addition are given a small cash stipend. There is no obligation for graduates to work with the Air Force unless they so desire. The graduates are in great demand and, as in the United States, the supply of electronic engineers is not adequate.

Electronics research is just starting in Brazil. At present, a small amount is being done at a few universities and other agencies. The Brazilian Navy is in the process of increasing its research and development activity by establishing a Naval Research Institute overlooking Guanabara Bay in Rio de Janeiro. The reinforced concrete structure is now 30 per cent complete and the building will be ready for occupancy by the end of the year. Although the Institute will conduct work in all the physical sciences, a major part of the program will be in electronics.

The United States and Brazilian Governments are now cooperating on the installation of nearly two million dollars worth of electronics equipment designed to improve the rapidity and safety of air traffic in the Rio de Janeiro, São Paulo, and Porto Alegre areas. Further cooperation is in the field of atomic energy that depends heavily on electronics for instrumentation and controls. Brazil is to receive $350,000 from the United States toward the cost of a 5000-kilowatt pool-type research reactor to be constructed by a United States firm at the University of São Paulo.

The formation of the first IRE Section in Brazil at Rio de Janeiro comes at a time when interest in radio and electronics developments and applications is increasing. Before petitioning for the new Section, there were 58 IRE members in all Brazil, 18 of whom resided in Rio. The possibility of forming a section was met with considerable enthusiasm and 16 new Associate plus 9 new Member Grade Members joined the original 18 to make a total of 43 names on the petition. Aside from the engineering and scientific benefits accruing from Section activities, there is also the very practical benefit of making it considerably simpler to arrange a favorable exchange rate for transmitting subscription fees to the IRE in New York.

It is believed that the Rio de Janeiro Section of the IRE will contribute much to the future advancement of radio and electronics in Brazil, and also IRE will be strengthened by this new activity under the Southern Cross. More IRE sections in Brazil may be expected as the electronics industry expands. Already São Paulo has sufficient members to submit a petition and there are indications that Belo Horizonte and Porto Alegre will also become important electronics centers.

# Controlled Fusion Research—An Application of the Physics of High Temperature Plasmas*

RICHARD F. POST†

For the past several years, physicists have been quietly conducting one of the most important and challenging investigations in the history of modern science—a search for a means of producing a controlled thermonuclear, or fusion, reaction. The search is indeed challenging, since a suitable fuel must be "heated" to kinetic temperatures of 100,000,000°C, and confined at that temperature long enough to permit the fuel nuclei to undergo fusion. Imposing as this challenge is, the advantages that the fusion process offers over nuclear fission and other present-day methods of producing energy are equally imposing: an inexhaustible supply of fuel, freedom from radiation hazards, a process whereby energy may be extracted directly in electrical form without first going through cumbersome heat-exchange equipment.

Most remarkable is the nature of the currently attractive methods to achieve controlled fusion. In contrast to fission power reactors, which use the techniques of the chemical and mechanical engineer, controlled fusion may use largely electronic techniques. The work, therefore, is of predominant interest to the IRE membership.

Last July the *Reviews of Modern Physics* published an excellent article which reviewed the need for new energy sources, and described the physical processes by which a controlled fusion reaction may occur and the technical problems that must be surmounted. The Editorial Board felt the discussion was so illuminating that it should be reprinted in the following pages for the benefit of all PROCEEDINGS readers. To assist readers who may not be familiar with certain terms used in nuclear physics, the editors have added a glossary (opposite, on page 135), prepared with the help of E. O. Johnson of RCA Laboratories.

As a final comment, it would be well for readers to note the close relationship of the thermonuclear power developments to gas discharge phenomenon and microwaves, subjects which have been of interest for a long time to radio engineers. The microwave relation is mentioned on page 158 of the article; although it is only a small part of the paper, it is a very important tool and may even play a role in power transmission from one region to another.—*The Editor*

*Summary*—Some of the long-range implications and advantages of achieving the production of power from controlled fusion reactions between isotopes of hydrogen, helium, and lithium are set forth. The physical conditions which seemingly must be established to accomplish this are presented. These are shown to lead to a situation in which the fusion fuel will exist in the form of a very hot, tenuous, fully ionized gas—*i.e.*, a plasma. It is shown that the required conditions can be maintained only by the use of a force field, for example, a magnetic field, to confine the plasma. A simple example of magnetic confinement—the pinch effect—is described. The electrodynamic properties of a fully ionized gas interacting with a magnetic field are briefly outlined. General scaling laws which would probably apply to any operable controlled fusion reactor are given. Examples are cited of fruitless approaches to the achievement of controlled fusion. Some aspects of the problem of "plasma diagnostics" are described to illustrate means for an experimental approach to high temperature plasma physics.

## INTRODUCTION

ACHIEVEMENT of the controlled release of fusion energy on a large scale will represent a truly permanent solution to mankind's expanding need for energy sources. Among the several light elements which might be used as fuel in a fusion reactor, the deuterium in the world's oceans alone would sustain an energy production rate 1000 times the world's present capacity for more than a billion years. However, the technical problems to be solved seem great indeed.

When made aware of these, some physicists would not hesitate to pronounce the problem impossible of solution.

Scattered references to the problem of fusion power have appeared in the literature for many years. It is known that several countries are working on the problem, and at the 1955 Geneva Conference H. J. Bhabha of India ventured the prediction that the problem will be solved in twenty years. In the fall of 1955, Chairman Strauss of the U. S. Atomic Energy Commission announced officially that the Commission was supporting, under the code name of Project Sherwood, a long-range research program aimed toward achieving controlled fusion for peacetime uses. It was disclosed that the major portion of the experimental work was being carried out at three sites (Princeton University, and two laboratories operated for the U. S. Atomic Energy Commission by the University of California: the Los Alamos Scientific Laboratory, and the University of California Radiation Laboratory at Livermore), and that supporting research was being carried out at Oak Ridge and New York University. Although 1951 was given as the beginning date for Project Sherwood, the problem of fusion power had actually been under study at Commission Laboratories since before the end of World War II. Much of the basic physical theory needed was formulated at Los Alamos by Edward Teller, Enrico Fermi, James Tuck, and others. Some of the

## GLOSSARY

**barn** — A unit of area, equal to $10^{-24}$ cm², used for measuring nuclear cross sections.

**bremsstrahlung** — The electromagnetic radiation emitted by a charged particle during deceleration.

**coulomb barrier** — The electrostatic barrier, or "hill," associated with the inverse square electric field surrounding a charged particle.

**coulomb field** — The inverse square law electric field surrounding an electric charge.

**cross section** — The effective target area of one particle with respect to a collision, or interaction, with another oncoming particle.

**Debye length** — In an ionized medium or plasma a characteristic distance from any randomly chosen particle A, beyond which other charged particles are effectively shielded from the electric field of A. The shielding or screening action is caused by charges of opposite polarity, surrounding A, whose aggregate electric field effectively neutralizes that of particle A.

**Debye sphere** — A sphere whose radius is a Debye length.

**deuteron** — A hydrogen nucleus containing a neutron besides the usual proton. It is the nucleus of deuterium.

**deuterium** — An atom whose nucleus is a deuteron. This atom is an isotope of hydrogen having an atomic weight of two.

**electron volt (ev)** — The kinetic energy acquired by a particle, having one electron charge unit, that has been accelerated by a potential difference of one volt. It is $1.6 \times 10^{-19}$ watt-seconds. This is a convenient definition for relating mechanical to electrical energy.

**fission reaction** — A nuclear reaction where very heavy nuclei are broken into smaller nuclei. As in the fusion reaction, energy is released from the annihilation of a small amount of matter according to the famous equation $mc^2$ = energy. This shows that one atomic mass unit is equivalent to 932 million electron volts of energy.

**fusion or thermo-nuclear reaction** — A nuclear reaction where light nuclei are combined into heavier nuclei.

**kev** — Thousand electron volts.

**kinetic temperature** — See temperature.

**Maxwellian distribution** — A particular, but very commonly observed, distribution of energies among the particles of a classical system. It is characteristic of systems whose particles engage in perfectly random or "heat motion."

**mean free path** — In a system of particles the mean distance that one particle travels before it collides, or interacts, with another particle.

**mev** — Million electron volts.

**plasma** — A region of relatively high but essentially equal densities of freely moving positive and negative charge-carrying particles.

**triton** — A hydrogen nucleus containing two neutrons besides the usual proton. Tritium is an atom whose nucleus is a triton. Tritium is an isotope of hydrogen having an atomic weight of three.

**temperature, kinetic temperature** — The concept of temperature as normally used describes the kinetic condition, or energy, of a system of particles in kinetic equilibrium with themselves and their surroundings. The temperature appears as a form factor in the particle energy distribution function. It turns out, from an integration of this function for classical systems (Maxwellian energy distribution), that the mean kinetic energy of the particles is three-halves the temperature times a constant, the Boltzman constant. The concept of kinetic temperature arises from an extrapolation of the temperature concept to describe the energy of particles which are not in equilibrium with their surroundings. The kinetic temperature is two-thirds the mean particle energy (in a Maxwellian energy distribution) divided by the Boltzman constant. In terms of electron volts, one electron volt = $1.16 \times 10^4$ degrees Kelvin of kinetic temperature.

**D** — Symbol for a deuteron.

**DD** — Symbol for a deuteron-deuteron reaction.

**DT** — Symbol for a deuteron-triton reaction.

**He³** — A helium nucleus of atomic weight 3. It contains two protons and a neutron.

**He⁴** — A helium nucleus of atomic weight 4. It contains two protons and two neutrons.

**Li⁶** — A lithium nucleus of atomic weight 6. It contains three protons and four neutrons.

**Li⁷** — A lithium nucleus of atomic weight 7. It contains three protons and four neutrons.

**n** — Symbol for neutron.

**p** — Symbol for proton.

**STP** — Abbreviation for standard temperature and pressure which is taken as 0°C. and 760 mm of mercury.

**T** — Symbol for triton.

results of this theory were later to be applied in the Los Alamos Sherwood experimental program, initiated and directed by J. L. Tuck. Also, early in 1951, Princeton University astrophysicist Lyman Spitzer, Jr., unaware of the classified Los Alamos work, conceived a different approach to a fusion reactor from others then under consideration in this country and submitted his idea to the U. S. Atomic Energy Commission. The Commission in turn initiated support under Spitzer's direction of what is now Project Matterhorn at Princeton. Shortly thereafter Herbert York, of the University of California Radiation Laboratory in Berkeley, having learned of the work of the Los Alamos and Princeton groups, suggested several new approaches. Thus, along with his job of organizing the new Livermore laboratory, York formed a small experimental group, with the author of this article as head, for the purpose of exploring and expanding on these new approaches to the problem.

In June of 1952, T. H. Johnson, Director of the Research Division of the U. S. Atomic Energy Commission, convened a classified conference on Controlled Thermonuclear Reactions in Denver, Colo.

Under the chairmanship of Edward Teller, the conference produced a summary of the problem and the then most promising attacks on it.

The great hopes and equally great technical difficulties of the controlled fusion problem were brought into focus, and under the Research Division, coordination and expansion of the Commission's controlled fusion research was begun.

Since the Denver meeting, many additional conferences have been held covering the experimental and theoretical aspects of the problem. At present the Sherwood scientific program is guided by a national steering committee, which has as its members: Edward Teller of the University of California, James L. Tuck, of the Los Alamos Scientific Laboratory, Lyman Spitzer, Jr., of Princeton University, and William Brobeck of the University of California Radiation Laboratory. The coordination and administration of the over-all program was placed under the direction of Amasa S. Bishop, head of the Sherwood branch of the Division of Research.

In many of its parts this article will represent merely a distillation of the work of the many physicists who have been contributing diligently to the Sherwood effort. The thorough understanding of the basic physical principles of the controlled fusion reaction which exists today is in large part a measure of their individual and collective genius.

Undergirding the entire field is, of course, the earlier work of the astrophysicists, who may now look forward to the possibility of practical peacetime applications of ultra-high temperature processes, which have in the past only been known to exist in the stars, or in the core of the atomic bomb.

## MOTIVATIONS

It has been said that the achievement of a controlled fusion reaction represents one of the great scientific challenges of the century. This fact alone would provide sufficient stimulus to the physicists of today to tackle the problem with vigor. But a broader look will show that the motivations are much deeper. Sobering studies of population trends and the great humanitarian needs for improving the health and standard of living of the majority of the world's people show the need for an astounding sustained growth in power production capabilities in the world.[1]

If these extrapolations prove to be valid, fossil fuels such as coal and oil would fall far short of satisfying the demands in less than a century. Even the recoverable solar energy incident on all the accessible land area of the earth would scarcely equal some predictions of the power requirements a century from now. If the predicted demand is to be satisfied and maintained, there remains only nuclear fission or nuclear fusion as sources of large quantities of energy. It is estimated that the energy reserve of fissionable fuels which are economically processable is roughly 25 times that of fossil fuels. Large though this reserve is, it would represent only a relatively few decades of energy needs in the possible world of a century from now. By comparison, the energy requirements of that future era could be supplied for a billion years by fusion of only the deuterium in the oceans. Fusion fuel is essentially inexhaustible.

Another significant comparison is in the cost and cost-trends of nuclear and fossil fuels. Even by the extraction methods in use today the per-unit-energy cost of deuterium as a fuel is only a few per cent of the cost of coal. While the unit cost of all other fuels including fissionable materials will become increasingly higher as lower grade reserves are tapped, the cost of deuterium recovered from sea water can be expected to drop or at least remain roughly constant indefinitely.

Again in the framework of the possible world of a century from now, the prospect of supplying some hundreds of billions of kilowatts of power from fission reactors alone is not a pleasant one from the standpoint of radiation accidents and disposal of radioactive wastes. At such levels of power production some $10^{13}$ curies of long-lived radioactive solid and gaseous fission products would be produced per year and would have to be safely disposed of. The full impact of this problem will of course not be felt in this generation, so that these extrapolations clearly do not influence the present desirability of constructing power generating plants based on the fission reaction.

The same questions of explosion hazard and radioactive waste disposal can be asked about a fusion reactor. There the most probable answer is that no dis-

[1] P. C. Putnam, "Energy in the Future," D. Van Nostrand Co. Inc., New York, N. Y.; 1953.

cernible explosion hazard will exist, and by the nature of the reactions involved no appreciable amount of radioactive waste is likely to be produced.

To anticipate later discussions, the problem of achieving a controlled thermonuclear reaction is to heat a suitable nuclear fuel to kinetic temperatures of 100,000,000°C or more, and then to controllably confine it somehow at these temperatures for a sufficiently long time to permit the fuel nuclei to undergo fusion—with the consequent release of energy in excess of the losses from the reaction region. The excess energy flux would then be available to be harnessed as useful power. The generation of fusion power differs greatly from that of fission power. No initial heating or heat insulation of fissionable fuel is needed to cause a nuclear fission reaction to become self-sustaining, whereas heating and confinement of the fuel are the prime problems of the fusion reaction.

Although reasonably good estimates can be made of the operating expenses of a fission power plant, no estimate of similar accuracy can be made for a fusion reactor. It can probably be said, however, that there is no reason to believe that the cost of fusion power cannot be made competitive with other sources, given sufficient development, especially in the light of rising costs for all other types of fuel.

One important possibility implicit in the nature of the controlled fusion reaction is that methods can be envisaged by which much of the fusion energy could be extracted directly in electrical form. This possibility could introduce a revolutionary reduction in the cost and complexity of electrical power generating installations, through elimination of the conventional thermal cycle.

It appears that the continued growth of civilization as we know it will demand the achievement of practical fusion power within a period substantially less than one hundred years. It further appears that a primary fuel for fusion power, deuterium, is essentially inexhaustible and that even its utilization on a grand scale should present no problems of radioactive waste disposal. Beclouding this rosy prospect are the truly formidable scientific and technical problems yet to be solved in achieving a practical controlled fusion reactor.

## What Is a Controlled Fusion Reactor?

The answer to the question, "What is a controlled fusion reactor?" is best given today by merely saying that it is a device within which appropriate isotopes of light elements could be caused to undergo nuclear fusion, the end result being the controlled production and extraction of useful quantities of energy, in excess of that required to operate the device. Ingenious suggestions have been advanced for defining some possible forms such a device might take. Study of these devices and the associated physical phenomena and apparatus

important to their performance is the reason for the existence of Project Sherwood. Some physical problems common to any research efforts in the field will be sketched in this article. In this section, some general aspects of the problem will be discussed which are predominantly related to the nuclear reactions themselves.

### Nuclear Fusion Reactions

Among the nuclear reactions which appear promising for a controlled fusion reactor are those involving the various isotopes of hydrogen, helium, and lithium. Some of these reactions are listed in Table I.

TABLE I

FUSION REACTIONS

| | | |
|---|---|---|
| 1) | $D + D \rightarrow He^3 + n +$ | 3.25 mev |
| 2) | $D + D \rightarrow T + p +$ | 4 mev |
| 3) | $T + D \rightarrow He^4 + n +$ | 17.6 mev |
| 4) | $He^3 + D \rightarrow He^4 + p +$ | 18.3 mev |
| 5) | $Li^6 + D \rightarrow 2\ He^4 +$ | 22.4 mev |
| 6) | $Li^7 + p \rightarrow 2\ He^4 +$ | 17.3 mev |

Reactions 1) and 2) occur with roughly equal probability. Reactions 3) and 4) are interesting because of their high energy yield and also because they involve the reaction products of 1) and 2).

The cross sections for reactions 1), 2), 3), and 4) have been measured down to low energies. Fig 1 (next page) summarizes some recent data on these reactions.[2]

The strong influence of the Coulomb barrier on the cross section at low energies is to be noted. This effect would, of course appear even more pronounced for reactions 5) and 6), making the practical utilization of these reactions more difficult.

### Energy Balance Considerations

Since reaction cross sections not far below the maximum values are achieved [particularly for reaction 3)] in the range of energies between 10 and 100 kev, whereas the energy yield per reaction is several mev, it is apparent that a substantial margin exists for the production of a net energy gain. That is, even if only a few per cent of a group of energetic deuterons and tritons were to undergo fusion, more reaction energy would be produced than the total kinetic energy of the original group.

As is well known, the energy balance still falls far short of permitting a net energy gain by simply bombarding a target of, say, deuterium with a beam of deuterons. In such a case the bulk of the kinetic energy of the impinging beam is dissipated uselessly by ionization, radiation, and energy transfer to the atomic elec-

[2] W. R. Arnold, J. A. Phillips, G. A. Sawyer, E. J. Stovall, Jr., and J. L. Tuck, "Cross sections for the reactions $D(d, p)T$, $D(d, n)He^3$, $T(d, n)He^4$, and $He^3(d, p)He^4$ below 120 kev," *Phys. Rev.*, vol. 93, pp. 483–497; February, 1954.
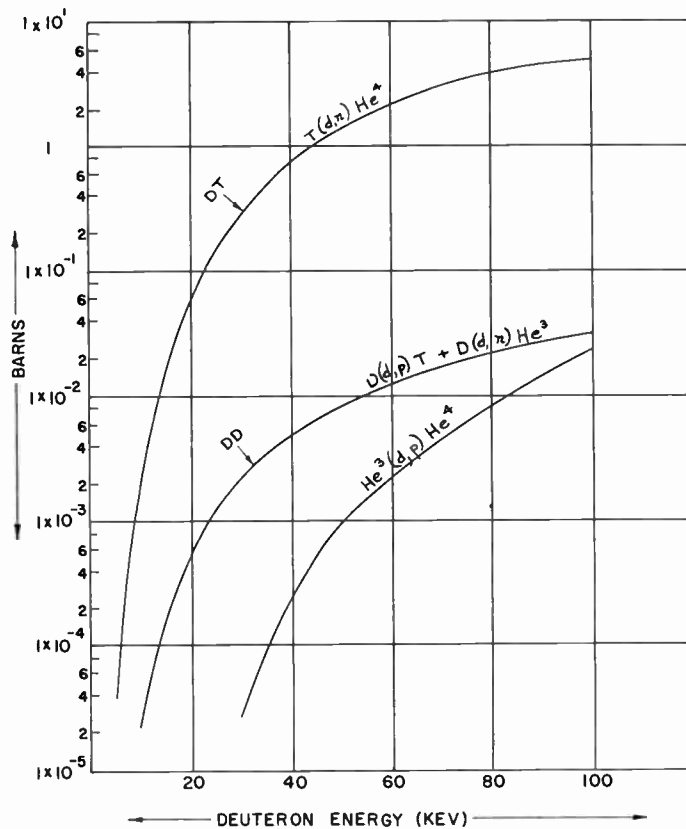
Fig. 1—Nuclear fusion reaction cross sections as a function of relative particle energy.

trons of the target, resulting in a yield of only $10^{-5}$, or about $10^{-4}$ of that necessary to produce a favorable energy balance.

A possible answer to this problem, taking a leaf from the book of the astrophysicist, is that the entire fuel charge must be heated to a kinetic temperature sufficient to produce a substantial reaction rate by virtue of the mutual collisions of these fuel nuclei. These kinetic temperatures correspond to mean particle energies of tens or perhaps even hundreds of kilo electron volts.[3] At such temperatures all matter is completely ionized, and ionization losses, dominant in the example of the deuteron beam incident on a target, are negligible. Achievement of a net power balance under such conditions then becomes primarily a problem of competition between the nuclear energy produced within the volume of the reacting fuel charge and the energy lost through its outer surface. Here again the astrophysicists will answer that, since the volume to surface ratio increases with the radius, sufficiently large reacting regions can always maintain a net energy balance, *viz.*, the sun and the stars. For earthlings this method of solving the problem is not a very attractive one and other means must

---

[3] For a Maxwellian particle distribution the mean particle energy $\overline{W}$ is equal to $(\frac{3}{2})kT$. In discussions of fusion reactions it is convenient to speak of kinetic temperatures in terms of kilo electron volts. In these units, 1 kev kinetic temperature $=1.16\times10^{7\circ}$ Kelvin, and is $\frac{2}{3}$ of the mean particle energy in a Maxwellian distribution appropriate to that temperature.

be found. These means must consist in picking out the most favorable fusion reactions and in creating especially effective circumstances for the reduction of surface energy losses, not only from the standpoint of achieving a favorable energy balance, but also of preventing the material walls surrounding the reaction region from being vaporized.

As has been said, the heating of matter to ultrahigh kinetic temperatures results in a state of complete ionization, *i.e.*, a gas composed of free electrons and ions in equal numbers, charge-wise. In this report, a fusion fuel in such a state is called a *plasma* after the original definition by Langmuir. The study of the dynamics of totally ionized gases or plasmas is a new and important field of physics which is receiving much attention in connection with astrophysical problems. This is perhaps not surprising when it is realized that all but a tiny fraction of the matter in the universe exists in the plasma state. Controlled fusion research represents an attempt to apply astrophysical plasma dynamics on an earthly scale.

### Theory of Binary Reactions

The fusion reaction is a binary one, *i.e.*, two-body collisions are involved. In the case of a fusion reaction in a heated plasma, continuous mutual collision processes are involved among all the particles. It will be useful to review the theory of such processes.

Suppose that a group of fuel ions with a kinetic temperature $T$ are mutually colliding. The probability that two fuel ions will react in passing close to each other is, of course, describable in terms of a mutual reaction cross section $\sigma$, which is a function of the relative velocity between the ions, $v_{12}$. The probability per unit time that a given ion of type 1 will react with another of type 2 will be given by the product of the reaction cross section $\sigma$, the relative velocity $v_{12}$, and the particle density of atoms of type 2, $n_2$ ions/cm³. Since there will exist a distribution of relative velocities, rather than a fixed relative velocity, the product $\sigma v_{12}$ must be averaged over the distribution appropriate to the kinetic temperature of the plasma. Thus the rate per particle of type 1 is

$$R_1 = n_2\langle\sigma v_{12}\rangle_{Av}. \tag{1}$$

The total reaction rate per unit volume is then found by multiplying $R_1$ by the particle density of ions of type 1

$$R_{12} = n_1 R_1 = n_1 n_2\langle\sigma v_{12}\rangle_{Av} \text{ reactions/cm}^3/\text{sec.} \tag{2}$$

If ion types (1) and (2) are identical (as in the DD reaction), then the expression becomes

$$R_{11} = \tfrac{1}{2}n^2\langle\sigma v\rangle_{Av} \text{ reactions/cm}^3/\text{sec.} \tag{3}$$

The reaction power density is simply the reaction rate times the energy release per reaction, $W_{12}$ or $W$.

$$p = n_1 n_2\langle\sigma v_{12}\rangle_{Av}W_{12}, \tag{4}$$

or

$$p = \tfrac{1}{2}n^2\langle\sigma v\rangle_{Av}W. \qquad (5)$$

Where the fuel plasma velocity distribution is known, the computation of $\langle\sigma v\rangle_{Av}$ is straightforward, though lengthy. Where a Maxwellian velocity distribution can be assumed, useful analytic expressions can be derived which predict $\langle\sigma v\rangle_{Av}$ at low and medium energies.

In (2) it is shown that reaction cross sections at low energies can be fitted accurately by a Gamow barrier penetration expression. For the total DD reaction the expression is (energies in kev, cross sections in barns, 1 barn $= 10^{-24}$ cm$^2$)

$$\sigma_{DD} = \frac{288}{W} \exp\,|-45.8W^{-1/2}|. \qquad (6)$$

The value of $\langle\sigma v\rangle_{Av}$ can be obtained by integrating the product of the above expression and the particle velocity distribution over all possible relative velocities. For a Maxwellian distribution with a temperature $T$ in kilovolts, the result obtained is[4]

$$\langle\sigma v_{DD}\rangle_{Av} = 260 \times 10^{-16}T^{-2/3} \exp\,|-18.76T^{-1/3}|$$
$$T < 50 \text{ kev.} \quad (7)$$

A similar but somewhat more complicated expression can be obtained for the DT reaction. The values of $\langle\sigma v_{DD}\rangle_{Av}$ and $\langle\sigma v_{DT}\rangle_{Av}$ for Maxwellian particle distributions are shown in Fig. 2 for the energy range 1 to 100 kev.

It is interesting to note that at very low temperatures (say below 5 kev) the important reacting particles are those few with energies several times the mean particle energy. This is a consequence of the exceedingly steep falloff of the reaction cross section at low energies.

A rough estimate of this effect for the DD reaction may be obtained merely by locating the maximum of the function

$$\exp\,-|45.8W^{-1/2} + WT^{-1}|, \qquad (8)$$

which is the product of the exponential terms in the reaction cross section and Maxwellian velocity distribution expressions.

By differentiation, the value of $W_m$, the energy of ions giving the maximum contribution to the reaction rate, can be found. Expressed as a ratio of $W_m$ to the

$$\frac{W_m}{T} = \frac{8.1}{T^{1/3}}. \qquad (9)$$

Thus at $T = 1$ kev this expression indicates that most of the reactions are contributed by particles with energies about eight times the kinetic temperature energy, *i.e.*, at low temperatures the tail of the Maxwellian wags the dog.

[4] See G. Gamow and E. Teller, "Rate of selective thermonuclear reactions," *Phys. Rev.* vol. 53, p. 608; April, 1938, for a derivation of the functional form. The numerical values above were obtained by C. Leith (private communication).
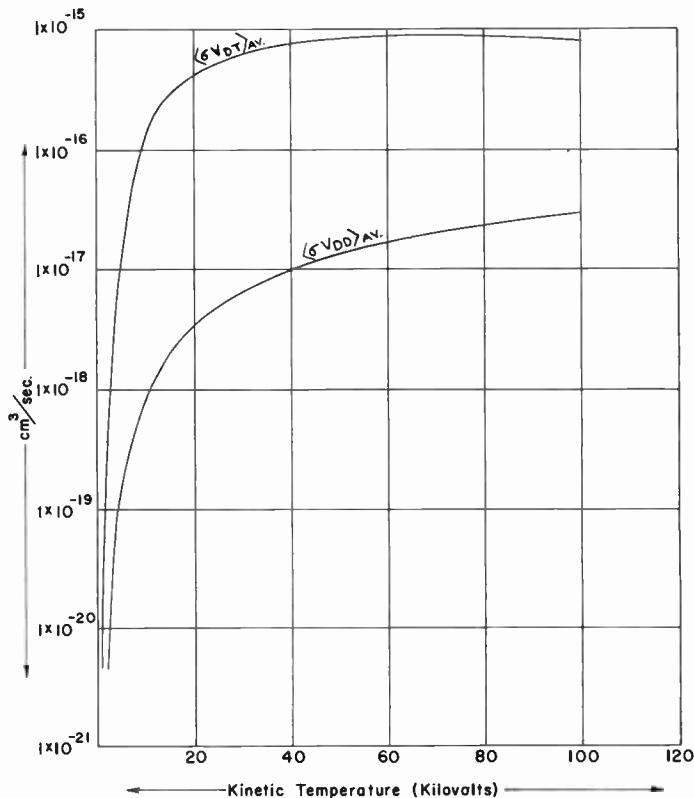


Fig. 2—$(\sigma v_{DD})_{Av}$, $(\sigma v_{DT})_{Av}$—reaction rate parameters for a Maxwellian particle distribution.

### Reaction Power Density

It is of interest to calculate representative reaction powers, mean reaction times, and reaction mean free paths for some possible situations. Consider the DD and DT reactions, supported at a temperature of 100 kev. (This temperature is chosen merely because the cross sections for both DD and DT are well developed and slowly varying at 100 kev, so that the figures quoted are not particularly temperature sensitive. It will later be shown that operation at temperatures an order of magnitude lower is at best marginal.) From the curves the value of $\langle\sigma v_{DD}\rangle_{Av}$ is seen to be about $3\times10^{-17}$ cm$^3$/sec. The mean reaction energy for the total reaction is, from Table I $(3.25+4)\div2 = 3.6$ mev or about $6\times10^{-13}$ joule/reaction. Thus the mean total reaction power density is

$$p_t = \tfrac{1}{2}n_D{}^2 \times (3 \times 10^{-17}) \times (6 \times 10^{-13})$$
$$= 9 \times 10^{-30}n_D{}^2 \text{ watts/cm}^3\ (T = 100 \text{ kev}). \quad (10)$$

The important feature of the expression is that it depends on the *square* of the deuteron density.

In Fig. 3, $p_t$ is plotted against the particle density of the fuel. The horizontal line at 100 watts/cm$^3$ (2800 kilowatts per cubic ft) represents a typical power density which might be achieved in a fission power reactor. The dashed vertical line at $n_D = 2.7\times10^{19}$ marks the particle density of a gas at STP.

Also plotted is the reaction power density for a 50 per
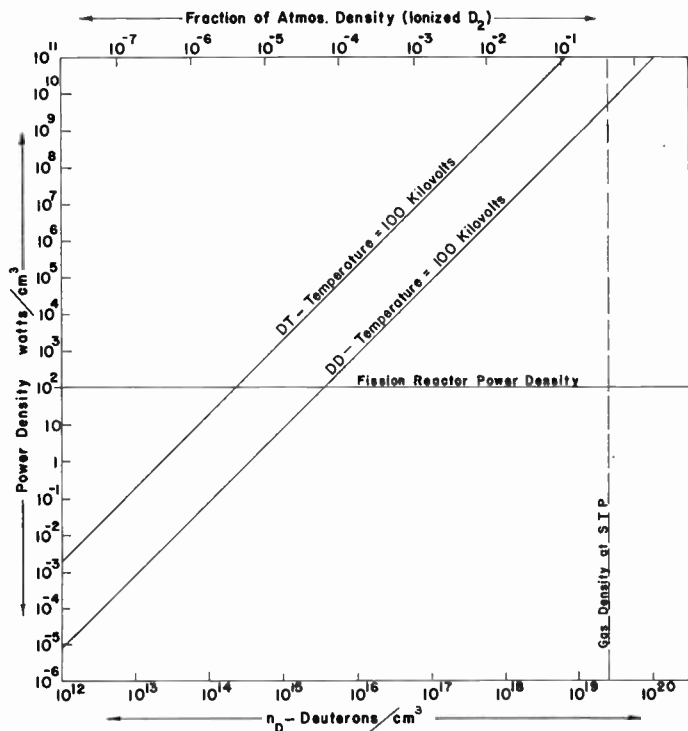
Fig. 3—DD and DT total reaction power density as a function
of deuteron particle density (temperature 100 kilovolts).

cent mixture of D and T at a temperature of 100 kilo-
volts, calculated from (2) and Fig. 2.

Examination of the curves for DD and DT shows
that power densities of the same order of those in fission
reactors are achieved at fuel particle densities of about
$10^{-4}$ and $10^{-5}$ times atmospheric densities, respectively.
These particle densities are perhaps surprisingly low,
since they are almost as low as are encountered in many
laboratory vacuum systems ($10^{-3}$ mm Hg $\approx 10^{-6}$ atmos-
phere). Now it is clear that an exact parallel cannot be
drawn between the problems of pressure and heat trans-
fer in a fission reactor which limit the working power
density, and the related problems in a fusion reactor.
Nevertheless, similar limitations on power transfer will
appear in any continuously operating fusion reactor. It
follows that continuous operation at a temperature of
100 kev and at densities even approaching atmospheric
densities seems impossible. A measure of this is the fact
that a power of 500,000 kilowatts represents about the
power output of a large steam-powered electrical gener-
ating plant. With deuteron particle densities equal to
atmospheric density, and at 100 kev temperature, Fig. 3
indicates that this power is equalled by the DD reaction
in a reacting volume of only 0.03 cubic centimeter! At
the same time the gas kinetic pressure exerted by the
fuel would be about $10^7$ atmospheres or $1.5 \times 10^8$ psi.

At the other extreme are the consequences of operat-
ing at too low a fuel density. Many proposed schemes
for producing thermonuclear power fail, not qualita-
tively but quantitatively, merely because insufficient
density can be achieved. When the density has fallen

to $10^{12}$ particles/cm³ for example, the curves show that
DD power has fallen to only about $10^{-5}$ watt/cm³,
which is far too low to be economically interesting. A
density of $10^{12}$ particles/cm³ is typical of plasma den-
sities generated in ordinary gas discharges.

### Reaction Mean Free Path—Energy Division

Operation of a fusion reactor at lower or higher tem-
peratures than 100 kev is surely conceivable. However,
it will be shown in a later section that operation below
a certain minimum temperature is not possible. Having
specified $\langle \sigma v \rangle_{Av}$ one may find the mean lifetime of a fuel
ion. From (1) the collision rate per particle is $R_1 = n\langle \sigma v \rangle_{Av}$.
Thus the mean lifetime before reaction is $\tau = 1/R_1$
$= 1/n\langle \sigma v \rangle_{Av}$. This is plotted in Fig. 4 as a function of
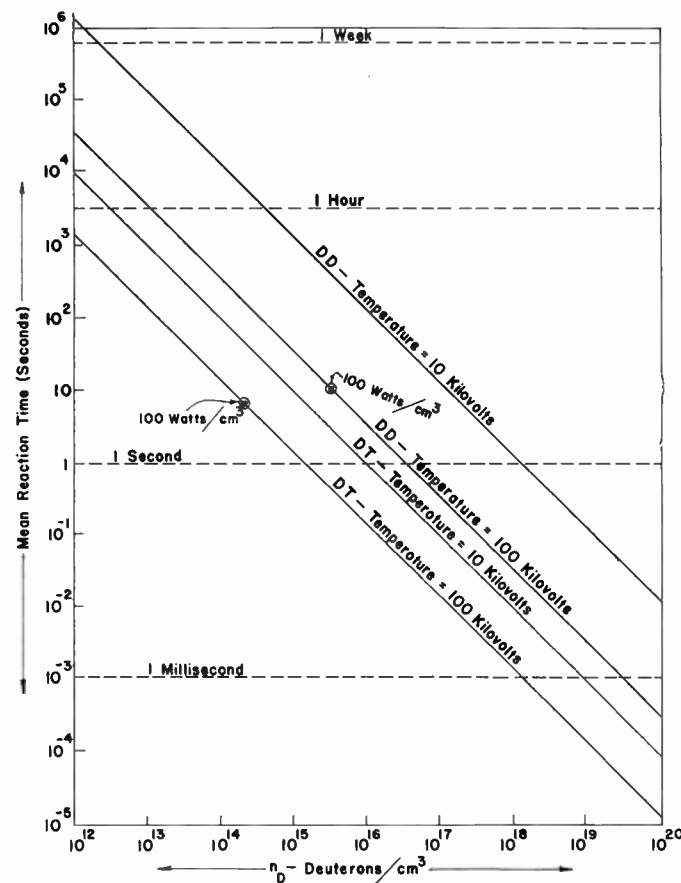


Fig. 4—Mean reaction time at various temperatures as a
function of deuteron particle density.

the deuteron particle density, $n_D$, for DD and DT at
10 kev and at 100 kev. The indicated points are located
at densities appropriate to a power density of 100
watts/cm³ (fission reactor values). From these it can be
seen that in what might constitute typical cases the
mean lifetime of a particle before reaction could be
many seconds. It follows that the mean life of a fuel ion
before it is lost from the reaction region must not be too
short compared to this time in order to maintain the
power balance.

The significance of a necessarily long reaction time can be more fully appreciated from the curves for the reaction mean free path, $\lambda = 1/n\bar{\sigma}$. These are plotted in Fig. 5, also for DD and DT at 10 and 100 kev. It is
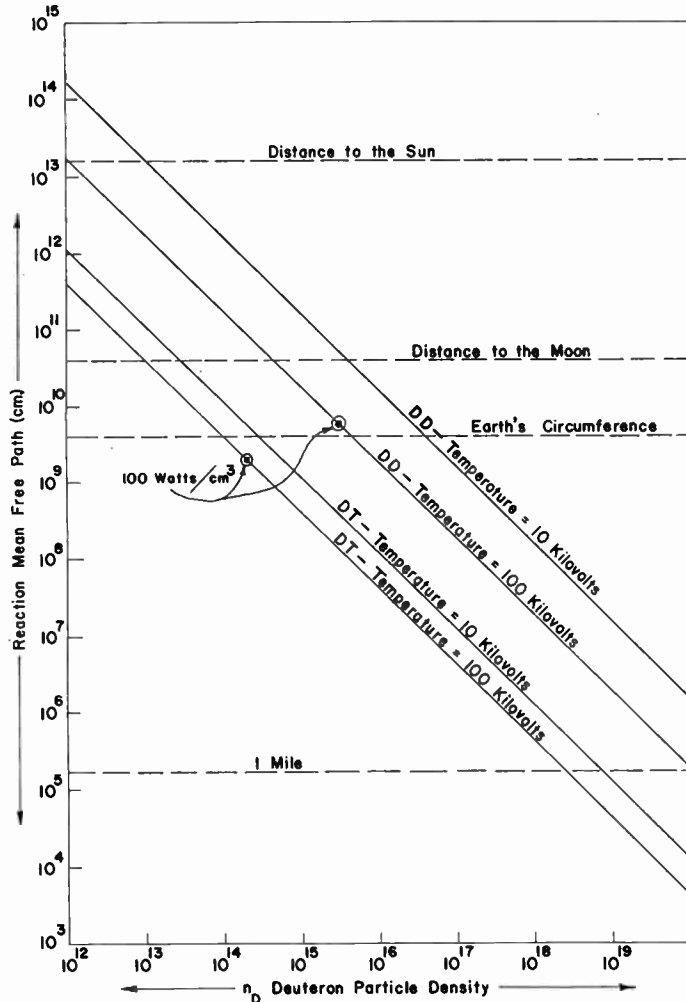


Fig. 5—Reaction mean free paths as a function of deuteron particle density.

to be noted that if the power density is assumed equal to 100 watts/cm³ the mean free path for DD reactions at 100 kev, for example, is about $5 \times 10^9$ cm, or a distance about equal to the circumference of the earth!

Thus far nothing has been said about the immediate fate of the energy produced when a fusion reaction takes place. The kinetic energy of the reaction is of course divided between the product particles, with the lion's share going to the lighter particle. If the lighter particle happens to be a neutron it will almost surely leave the reacting region (which has low density) and deposit its kinetic energy elsewhere. Charged reaction products may or may not leave the reacting region depending on the circumstances. In any event, the total particle energy of the charged reaction product is all that is available internally to supply energy losses and to sustain the reaction. The partition of energy between charged

particles and neutrons is listed below for DD and DT reactions (neglecting initial kinetic energies).

$DD \rightarrow (T + 1.0 \text{ mev}) + (p + 3.0 \text{ mev})$    } Roughly equal probability of occurrence.

$DD \rightarrow (He^3 + 0.8 \text{ mev}) + (n + 2.45 \text{ mev})$

$DT \rightarrow (He^4 + 3.6 \text{ mev}) + (n + 14.1 \text{ mev})$
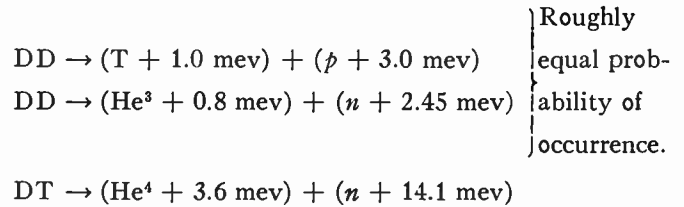
In the case of the total DD reaction, on the average 66 per cent of the reaction energy is transferred to charged reaction products and 34 per cent to neutrons. Thus the *internal* power density generated by the DD reaction is 0.66 times the value given by (10) and the DD curve of Fig. 3. In the case of the DT reaction only 20 per cent is transferred to charged products and 80 per cent is imparted to neutrons.

To simplify the discussion, except where noted, in this and all other sections of this paper where the DD reaction is considered, the possible contribution of secondary fusion reactions involving the charged reaction products themselves (*i.e.*, T and He³) is neglected.

## COMPETING PROCESSES

The processes in which energy is lost are in competition with the nuclear yields. The role of these processes is crucial. Some of these loss processes are common to any system. Among these is radiation from the electron cloud which, together with the fuel ions, forms the fuel plasma. If the more usual situation of radiation equilibrium were to be blindly assumed, the radiation flux from a reacting region of the fuel would be given by the usual blackbody relationship $I = \sigma T^4$ ergs/cm²/sec. At 10 kev (about $10^8$ degrees Kelvin), this radiation flux would amount to some $10^{21}$ watts/cm²! We can safely conclude that no controlled fusion device can operate in a state of radiation equilibrium between the particles and the radiation field. However, radiation equilibrium is only achieved when the mean free path for internal absorption of the radiation in question is smaller than the dimensions of the system. Fortunately this condition is far from satisfied in the plasma of a fusion reactor, whence the use of the phrase "kinetic temperature" rather than merely "temperature" to describe the energy state of a plasma.

The situation is similar to that in the tenuous outer parts of the sun's corona. Although it has been found that the kinetic temperature of the electron gas in the corona is of the order of $10^6$ degrees Kelvin, the actual radiation from the sun is effectively dominated by deeper lying layers and corresponds to that from a blackbody at a temperature value of only about 5000°K. In the outer parts of the corona, the mean free path for the important components of the self-radiation is much larger than the thickness of the layer so that radiative equilibrium is not attained.

*Effective Radiation Rate—Minimum Temperature for a Self-Sustaining Fusion Reaction*

For systems where dimensions are small compared to the mean free path for absorption of photons at an energy corresponding to the kinetic temperature, a rough rule is that the actual radiation from the medium will be of the order of the blackbody value diminished by the ratio of dimensions of the system to the mean free path for absorption of the photons. At $10^8$ degrees K the absorption mean free path for plasma densities of, say, $10^{15}$ particles per cm³, is about $10^{20}$ cm so that the reduction is of the same order. More precisely stated, the dominant radiation process from the plasma will be ordinary bremsstrahlung or X-ray emission occurring when the fast-moving electrons of the plasma are deflected by the Coulomb field of the fuel ions. The theory of this radiation is discussed by Heitler[5] and others. Upon averaging over a Maxwellian electron distribution an expression can be found for the radiation power density emitted as X rays from a completely ionized gas ($T_e$ = electron temperature in kilovolts).

$$p_r = 0.48 \times 10^{-30} Z^2 n_e^2 T_e^{1/2} \text{ watts/cm}^3. \qquad (11)$$

For a hydrogenic plasma $Z = 1$ and $n_e = n_i$ ($n_e$ = electron particle density, $n_i$ = ion particle density). So that for the DD reaction, for example,

$$p_r = 0.48 \times 10^{-30} n_D^2 T_e^{1/2} \text{ watts/cm}^3. \qquad (12)$$

Comparison with (5) for the rate of power generation shows that both vary as $n_D^2$. If electron and ion temperatures are assumed equal, it is seen that even in the absence of all other losses (except neutrons) a controlled fusion reaction sustaining itself from its internal energy generation cannot be accomplished below a certain minimum temperature, which is independent of the density. Since $\langle \sigma v \rangle_{Av}$ for the reaction increases exponentially with temperature while the radiation rate increases only as $T^{1/2}$, power production may exceed radiation above a critical temperature $T_c$ at which $p_r = p_i$ (we count only the *charged* reaction products since neutrons will escape). We find from (5), (7), and (12) the ratio of fusion to radiated power (an energy contribution from secondary TD reactions has been assumed)

$$p_i/p_r = 1.83 \times 10^4 (T)^{-7/6} \exp \left[ -18.76 T^{-1/3} \right]. \quad (13)$$

Setting $p_i/p_r = 1$, one finds $T_c$ to be about 35 kilovolts. The value of $T_c$ for DT is considerably lower, about 4 kev.

Because of the $Z^2$ dependence of $p_r$, it is apparent that the presence of even a small amount of ionized high-$Z$ elements could greatly increase the radiated power density. It follows that high purity is a prime requirement for the plasma of a fusion reactor. For this and other reasons utmost attention must be paid to the role of the material walls of the reactor in introducing impurities through bombardment and heating effects which might otherwise be considered unimportant.

*Need for Confinement—Possible Modes of Operation*

Radiation from the plasma electrons thus establishes a minimum temperature of operation of a fusion reactor, but does not introduce any specific limitations on size of the reaction region or on number density of the plasma (so long as radiative equilibrium is not approached), since both radiation power density and fusion power density vary in the same way with particle density. We have seen that, in a continuous operating reactor, power density limitations would probably appear, which in turn limit the particle density. It will be shown in a later section that the mean free path for ordinary collisions between particles of the plasma is very long when the plasma is at thermonuclear temperatures. But unless other means are used, collisions between the particles represent the only mechanism inhibiting their escape from the reaction region. Thus at densities such as $10^{15}$ particles per cm³ there is no *a priori* reason why a hot fuel ion may not escape from the reactor immediately.

It is clear that a dilemma exists in probing further into the concept of a controlled fusion reaction. Two general ways out of the difficulty can be mentioned. One possibility is that the idea of a continuously reacting plasma must be abandoned and a transient mode of operation be substituted. Here high densities would be used so that, for a short time at least, collisional diffusion velocities would be sufficiently slow to prevent the plasma from escaping rapidly enough from the region in which fusion is taking place to quench the reaction. The logical extrapolated limit of this process is a hydrogen bomb, which is not exactly "controlled" in the usual sense of the word. The second general attack is to interpose between the material walls of the reactor and the plasma some force field capable of exerting a pressure, so that the reaction heat is not immediately quenched by contact with a low temperature region. Even at reduced densities it is clear that the high temperatures of the fuel implies that substantial pressures must be sustained. For example, at 100 kev kinetic temperature and a total particle density of $6 \times 10^{15}$ particles per cm³, the kinetic pressure exerted by the fuel would be about 1000 atmospheres or 15,000 lb/in². A gravity field is clearly too weak, except when applied on a stellar scale. The outstanding remaining possibility is the electromagnetic field, which is capable of transferring momentum and thus can exert a force. Since the fusion fuel will be essentially totally ionized, the charged particles constituting the fuel are capable of direct interaction with such a field. The gamut of possibilities here is large and their discussion is clearly outside the scope of this article, but brief mention of an example will be made in the next section.

A controlled fusion reactor must operate far from radiative equilibrium, under which condition radiation

[5] W. Heitler, "*Quantum Theory of Radiation*," Oxford University Press, New York, N. Y., third edition; 1954.

losses are limited to ordinary bremsstrahlung. At a sufficiently high operating temperature reaction power will exceed radiation, permitting the achieving of a self-sustaining fusion reaction if no other significant losses are present. At the same time, to overcome excessive losses of particles (and thus energy) by direct transport out of the reaction region, either a high density (and therefore highly transient) mode of operation must be contemplated, or containment for useful lengths of time by some type of electromagnetic field must be sought.

## An Example of a Possible Means for Magnetic Confinement—The Pinch Effect

A classic example of an electromagnetic field which can act to confine a group of charged particles is the so-called "pinch effect"—the self-constriction of a group of charged particles moving in such a way as to produce a unidirectional current. The effect is another example of the familiar fact that parallel circuits carrying current in the same direction attract each other. A theory of the pinch effect was first set forth by Bennett[6] and amplified by Tonks[7] and Tonks and Allis.[8] Some experimental work on the subject has been reported in the literature[9,10] and recently a more direct connection has been cited between the pinch phenomenon and controlled fusion research.[11,12]

The theoretical work of Bennett and others showed that in the ordinary pinch effect at high currents the conducting column would become closely concentrated at a central axis, producing a plasma column and surrounding magnetic field as shown in Fig. 6. The fact that the current becomes pinched is direct evidence that the carriers are being confined in the radial direction by the effect of the magnetic field.

To illustrate the principles behind magnetic confinement of this type let us adopt an over-simplified picture of the pinch and derive from it the conditions for kinetic pressure balance between plasma and the effects of the pinch magnetic field. We assume, as shown in Fig. 6 that the pinched current is confined to a thin cylindrical shell or boundary of outer radius $a$ and thickness $\epsilon$. The plasma filled region inside the shell will be characterized by a uniform density $n_i$ of singly charged ions per cm³ and $n_e$ electrons per cm³, and a kinetic temperature $T$. $n_i$ will be taken equal to $n_e$ (necessarily so, as

[6] W. H. Bennett, "Magnetically self-focussing streams," *Phys. Rev.*, vol. 45, pp. 890–897; June, 1934.

[7] L. Tonks, "Theory and phenomena of high current densities in low pressure arcs," *Electrochem. Soc.*, vol. 72, p. 167; 1937.

[8] L. Tonks and W. Allis, "Theory of magnetic effects in plasma of arc," *Phys. Rev.*, vol. 56, pp. 360–373; August, 1939.

[9] A. A. Ware, "A study of a high-current toroidal ring discharge," *Trans. Roy. Soc.*, vol. A243, pp. 197–220; 1951.

[10] S. W. Cousins and A. A. Ware, "Pinch effect oscillations in a high current toroidal ring discharge," *Proc. Phys. Soc. (London)*, vol. A64, pp. 159–166; February, 1951.

[11] "Perhapsatron to Columbus," *Nucleonics*, vol. 13, p. 23; December, 1955, discusses Los Alamos Controlled Fusion Research Program under J. L. Tuck and discloses investigation of properties of the pinch effect.

[12] "Thermonuclear power . . . a search for ideas," *Nucleonics*, vol. 14, p. 42; February, 1956, includes a review of unclassified work on pinch effects at Tufts and University of Southern California.

Fig. 6—Schematic representation of pinch effect in a plasma.

will be seen later). Since it is a gas, the plasma within exerts a kinetic pressure $P = (n_i + n_e)kT$. Outside the shell the density and thus the kinetic pressure is equal to zero. The pressure difference must be supplied by a gradient of the magnetic field.

Macroscopically, the balancing force is just the body force $j \times H$ dynes acting on each unit volume of the current carrying region, where $j$ is the current density. In this case the force is of course exerted only within the thin shell carrying the current. Within the conducting shell itself the field will fall from its value of $2I_0/a$ at the outside, to zero at the inner surface, while at the same time the particle density is assumed to rise from zero at the outside, to its central value, at the inner surface. If we assume that current is distributed uniformly through the layer, with thickness $\epsilon$, then at any point in the shell itself the field will have the value

$$H = 2I/r, \tag{14}$$

where $I$ is the total current within the radius $r$. Thus, if the thickness of the shell is $\epsilon$,

$$I(r) = I_0 \left( \frac{r + \epsilon - a}{\epsilon} \right), \qquad (a - \epsilon) < r < a.$$

Let $r = a - x$, $0 < x < \epsilon$, then

$$I(x) = I_0 \left( 1 - \frac{x}{\epsilon} \right). \tag{15}$$

$$H(x) = \frac{2I_0}{a-x}\left(1 - \frac{x}{\epsilon}\right) \approx \frac{2I_0}{a}\left(1 - \frac{x}{\epsilon}\right), \qquad \epsilon \ll a. \quad (16)$$

The current density $j$ is the total current divided by the area of the shell

$$j = \frac{I_0}{2\pi a\epsilon}, \qquad \epsilon \ll a. \quad (17)$$

The total (inward) force per unit area exerted by the shell is the integral of $j \times H$ through the shell. Here $j$ is perpendicular to $H$ so that

$$\int_0^\epsilon (jH)dx = \left(\frac{2I_0}{a}\right)\left(\frac{I_0}{2\pi a\epsilon}\right)\int_0^\epsilon \left(1 - \frac{x}{\epsilon}\right)dx$$

$$= I_0^2/(2\pi a^2). \quad (18)$$

This force must just balance the kinetic pressure of the plasma $P = (n_i + n_e)kT$

$$I_0^2/(2\pi a^2) = P = (n_i + n_e)kT, \quad (19)$$

i.e.,

$$I_0^2 = 2NkT, \quad (20)$$

where $N$ is the total number of particles per centimeter of length of the confined column of radius $a$. This relationship is useful in establishing the current necessary for pinch confinement. Note that the expression for $I_0$ is independent of the radius of the pinch.

### Magnetic Pressure

Remembering that $H_0 = 2I_0/a$, we find that (19) may be rewritten as

$$\lfloor H_0^2/8\pi \rfloor_{\text{outside}} = P_{\text{inside}}. \quad (21)$$

The role of the magnetic field appears here in its true light. The external magnetic field region balances the kinetic pressure on the region inside by a magnetic pressure of magnitude $H_0^2/8\pi$. ($H_0^2/8\pi$ is also the energy density in the magnetic field.) The same relations above can be simply rearranged to yield the result that within the layer

$$\nabla\left(\frac{H^2}{8\pi} + P\right) = 0,$$

i.e.,

$$(H^2/8\pi) + P = \text{constant} = (H_0^2/8\pi). \quad (22)$$

This relationship is valid where curvature of the magnetic field lines can be neglected, as in the present case. It has appeared frequently in the literature of plasma physics.[13],[14] It is important to note that the magnetic pressure exerted is a direct result of the interaction of plasma currents and the magnetic field, and that this

pressure results only where the plasma currents act to diminish the field, i.e., where the medium has properties which could loosely be called diamagnetic. The example chosen, where the currents producing the magnetic field are excluded from a conducting region, can only occur in a perfect conductor, but the illustration could also represent a situation which might persist for a short time in a plasma of finite electrical conductivity.

To illustrate the strength of magnetic field required for magnetic confinement let us return to the former example of a DD reaction carried out at 100 kev temperature and a deuteron particle density of $3 \times 10^{15}$ per cm³.

Assume that the electron temperature and ion temperature are approximately equal. As in the former example the plasma pressure is $P = (n_i + n_e)kT = 6 \times 10^{15} \times (1.6 \times 10^{-7}) = 10^9$ dynes/cm² or about 1000 atmospheres (15,000 psi). From this

$$H_0^2/8\pi = 10^9, \qquad H_0 = 1.6 \times 10^5 \text{ gauss.}$$

If the radius of the plasma cylinder is 10 cm, for example, (19) shows that the required current is about $8 \times 10^5$ emu, or $8 \times 10^6$ practical amperes.

Thus far it has not been necessary to discuss the specific microscopic mechanism by which the magnetic field effects containment of the plasma. It is clear that the qualitative effect in the example above is that ions and electrons attempting to escape radially from the boundaries will be turned around upon entering the region of strong magnetic field. The infinitesimal currents produced by the reflection of the charged particles from the boundary must integrate to a self-consistent set of macroscopic currents which produce the given magnetic field configuration. These concepts can serve as an introduction to the problem of *stability* of the confined plasma, which is of great importance.

### Pinch Instability—Longitudinal Confinement

Although it has been shown above that a state of pressure equilibrium can be established in pinch confinement, it has not been shown that a stable equilibrium is possible for the simple pinch. In matter of fact it is not. In a now classic paper Kruskal and Schwarzchild[15] have shown that a self-constricting current flowing in a plasma is unstable against "kinking" perturbations. That is to say, if the pinched column undergoes an infinitesimal localized lateral displacement, displacements of length greater than the column diameter will grow exponentially with time, until the column is disrupted. The situation is something like the instability of a long thin rotating shaft supported only at its ends. It arises physically from the fact that the magnetic field (and therefore the magnetic pressure) is greater near the concave side of a curved current carrying conductor than it is on the convex side. This can be

[13] L. Spitzer, "Physics of Fully Ionized Gases," Interscience Publishers, Inc., New York, N. Y.; 1956.
[14] H. Alvén, "Cosmical Electrodynamics," Oxford University Press, New York, N. Y.; 1950.

[15] M. Kruskal and M. Schwarzchild, "Some instabilities of a completely ionized plasma," *Proc. Roy. Soc.* (London), vol. A223, pp. 348–360; May, 1954.

seen from the illustration in Fig. 7, where it is seen that the magnetic field lines become crowded together on the concave side of the column and spread apart on the convex side. The instability predicted by Kruskal and Schwarzchild is a special case of a more general type of hydromagnetic instability. Instabilities of this type have the property that the time of growth by a factor $e$ of an unstable perturbation is about equal to the transit time of an ion going a distance equal to the "wavelength" of the perturbation. Thus short "wavelengths" grow most rapidly and at thermonuclear temperatures the predicted growth times are disturbingly short. (A 100-kev deuteron has a speed of about 300 centimeters per microsecond.) In order to convert the simple pinch effect into a means for producing controlled fusion reactions, ways would have to be found to circumvent this fundamental instability. One conceivable way is to establish the pinch within a time shorter than the instability growth time. If in such a case a high enough instantaneous reaction power were attained, a net energy yield might still be achieved.
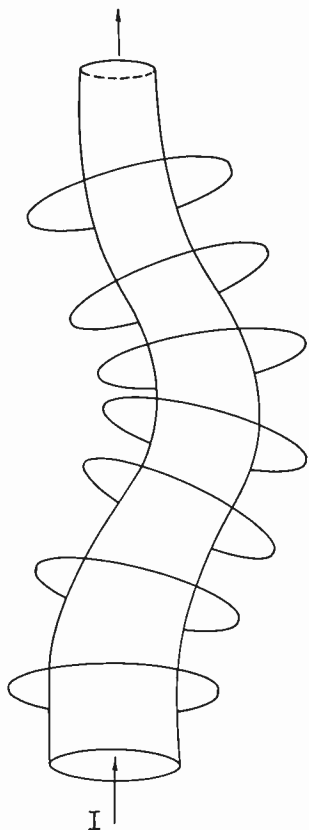


Fig. 7—"Kink" instability of the pinch effect.

To complete the picture of the simple (quasi-stationary) pinch effect as a possible confinement means, it should be mentioned that the field configuration shown in the example of Fig. 6 provides no effective means of longitudinal confinement of the plasma. In the experiments of Cousins and Ware,[10] some of the work of Tuck at Los Alamos and the work at Tufts and the University

of Southern California[12] toroidal induced discharges were used which provide in principle a way to reduce these losses, by closing the end of the discharge on itself.

## PHENOMENA IN A COMPLETELY IONIZED GAS

Problems in the physics of completely ionized gases have received considerable study in recent years. Although of obvious importance to controlled fusion reactions, any detailed treatment of this subject is outside the scope of this article and the reader is referred to the literature for further information.[13,14] However, some of the simple and fundamental features of this "fourth state of matter" will be reviewed in this section.

A plasma is governed in its gross behavior by three dominant processes: 1) long-range collective electrostatic or space-charge-like effects produced by the plasma itself; 2) short-range collisional interactions between individual particles of the plasma; and 3) interaction between the individual particles and externally applied electromagnetic fields. Strictly speaking, 1) and 3) are coupled effects, since the plasma may profoundly alter the character of an external applied field through collective effects. Nevertheless the separation suggested may often represent a reasonable approximation to the actual behavior.

### Charge Equality

Process 1) acts most strikingly in enforcing a state of charge equilibrium in a plasma. That is to say any tendency for the charge density to depart from neutrality will result in large electrostatic fields which oppose the change. To illustrate this effect it is only necessary to calculate the electrical field which would exist around a sphere of plasma with both electron and ion density equal to $10^{15}$ particles/cm$^3$ if the electrons of the plasma were somehow to be suddenly removed. If the radius of the sphere were 1 cm for example, Gauss' theorem gives for the field

$$E = \frac{Q}{r^2} = (4\pi/3) \times 10^{15} \times 4.8 \times 10^{-10}$$

$$= 2 \times 10^6 \text{ stat volts/cm} \quad \text{or} \quad 600 \text{ million volts/cm!}$$

Even a departure of only 1 part in $10^5$ from charge equality would give rise to a field of 6000 volts/cm near a sphere of plasma of 1 cm radius or 600,000 volts/cm near a sphere of 1 meter radius. Effective equality of electron and ion charge density must be assumed as a condition in any operating fusion reactor.

### Electrical Conductivity

Another obvious property possessed by a cloud of freely moving negative and positive charges is that of electrical conductivity. As in any conducting medium the numerical value of the conductivity depends inversely on the collision rates between the carriers and other neighboring particles. In a hot plasma (as opposed

to ordinary conductors) collision rates become smaller as temperature increases and thus the electrical resistivity *decreases* with increasing temperature. The theory of the resistivity of a fully ionized plasma[18] shows that it varies inversely with the temperature to the three-halves power. For a hydrogenic plasma the numerical value is given approximately by the expression ($T$ in kilovolts)

$$\rho_0 = \frac{3 \times 10^{-6}}{T^{3/2}} \text{ ohm-cm.} \qquad (23)$$

Thus at 100 kev the theoretical resistivity is about $3 \times 10^{-9}$ ohm-cm—less than 1 per cent of that of copper at room temperature. This expression strictly applies only when the effective dimensions of the plasma are large compared to a mean free collision path of electrons in the plasma. At low densities and high temperatures this condition may be far from satisfied. Nevertheless for other reasons the qualitative result of low resistivity remains valid in many cases of practical interest. It follows that in many situations a hot plasma behaves somewhat like a superconductor, *i.e.*, it can exclude or diminish externally generated magnetic fields (as in the example of the pinch) and it can act to "short out" charge accumulations or electric fields. These qualitative results are not particularly sensitive to the plasma density.

In the presence of a magnetic field, the theoretical steady state resistivity is slightly altered for conduction in a direction across the field lines and is essentially unaltered for conduction along the field lines. Transient effects may however be radically different, depending on the time scale.

## Plasma Oscillations

Another property of a fully ionized gas which owes its origin to space-charge phenomena and high effective conductivity is the phenomenon of electrostatic plasma oscillations. Simple forms of these oscillations which do not depend on a magnetic field being present for their existence were first predicted and observed by Tonks and Langmuir. These (longitudinal) oscillations of the plasma are associated with the fact that a small momentary average displacement of charges of one sign with respect to the other in the plasma will establish an electric field in such a direction as to oppose the displacement. Inertial effects of the displaced charges then provide a mechanism for periodic oscillations about the point of zero average charge displacement. The frequency of these oscillations is determined by the mass of the displaced charges and the "spring constant" of the electric field produced by the displacement. Plasma oscillations of this type may either be electronic or ionic in nature, *i.e.*, the frequency is determined by electronic or ionic masses. Ionic oscillations are much · lower in frequency than electron oscillations and are not as easy to identify. Considerable theoretical work has been done

in recent years on the subject of plasma oscillations, much of it by Soviet investigators.[16] The reader is referred to the literature for more complete details.[13,15,17,18]

It is important to note that the frequency of electronic plasma oscillations really represents a measure of the time of response of the electron gas of the plasma to a time varying electric field, whether it be the self-field resulting from a separation of charges or an externally applied field. Thus the plasma frequency appears ubiquitously in the descriptions of many phenomena in a plasma. Disregarding a correction for the effect of thermal velocities[16,17] one finds the electron plasma frequency is

$$\omega_p = [4\pi n_e e^2/m]^{1/2} \text{ radians/sec,} \qquad (24)$$

or

$$f_p = 9 \times 10^3 n_e^{1/2} \text{ cycles/sec.}$$

Thus at $n_e = 10^{15}$,

$$f_p = 2.7 \times 10^{11} \text{ cycles/sec.}$$

Transient electrical fields which are applied in a time much longer than $1/\omega_p$ are dominated by the electrical conductivity effects discussed earlier. If they occur in a shorter time, the inertial effects of the plasma dominate and it may have little influence. Thus it is that ordinary light passes freely through a plasma, whereas radio waves may be strongly reflected or attenuated. In matter of fact the effective dielectric constant of a plasma (in cases where collisional and magnetic field effects may be neglected) may be written as

$$K = 1 - \left(\frac{\omega_p}{\omega}\right)^2 = 1 - \frac{4\pi n_e e^2}{m\omega^2}, \qquad (25)$$

an expression no doubt familiar to students of classical dispersion theory. Here $\omega$ is the angular frequency of the electromagnetic wave propagated through the plasma.

The reciprocal of the index of refraction is

$$\nu = \frac{1}{\sqrt{K}} = \frac{1}{\left[1 - \left(\frac{\omega_p}{\omega}\right)^2\right]^{1/2}}, \qquad (26)$$

corresponding to phase velocities *greater* than the velocity of light. Unattenuated propagation of an electromagnetic disturbance can occur only if $\omega > \omega_p$, a situation resembling that in a waveguide.

## Hydromagnetic Waves

The introduction of a magnetic field has little effect on electromagnetic wave propagation in a plasma if both $\omega$ and $\omega_p$ are much greater than $\omega_c$, the electron

[16] L. Landau, "On the vibrations of the electronic plasma," *J. Phys. U.S.S.R.*, vol. 10, p. 25; 1946.
[17] D. Bohm and E. P. Gross, "Theory of plasma oscillations, A. Origin medium like behavior," *Phys. Rev.*, vol. 75; pp. 1851–1864; June, 1949.
[18] D. Gabor, "Wave theory of plasmas," *Proc. Roy. Soc. (London)*, vol. A213, pp. 73–86; June, 1952.

cyclotron frequency. Also if the electric vector of the electromagnetic wave is parallel to $H$, propagation of the wave as predicted by (26) is relatively unaffected, whatever the frequency. When $\omega$ or $\omega_p$ are comparable to or smaller than $\omega_c$, the situation is no longer simple, and complicated dispersion effects result.

In the presence of a magnetic field new types of waves are also possible which are of fundamental importance. These are the so-called magnetohydrodynamic or hydromagnetic waves described by Alfvén.[14] In their simpler forms these waves are transverse, propagate along magnetic lines of force, and may be thought of as similar to the transverse oscillations of loaded elastic bands (the field lines). The loading is of course supplied by the ions (and electrons) of the plasma and is related to a very important qualitative concept of the motion of plasma particles in the presence of a magnetic field. This concept is that for motions which occur sufficiently rapidly compared to interparticle collision frequencies (but not too rapidly compared to periods of rotation of the particles), charged particles act as though bound to magnetic lines of force (more precisely to certain special magnetic surfaces).[19] Since, as noted, magnetic field lines act very much like mutually repelling elastic bands, transverse vibrational or torsional waves may be propagated along the direction of the field lines at speeds determined by the mass loading per unit length (*i.e.*, the ion density) and a force constant determined by the strength of the magnetic fields. The speed of propagation of these waves is given by the Alfvén speed

$$v_A = \left(\frac{H^2}{4\pi\rho}\right)^{1/2}. \qquad (27)$$

$\rho$ is equal to $(n_i M_i + n_e m_e) \approx n_i M_i$. Squaring both sides of (27) and dividing by the square of the velocity of light one obtains

$$\left(\frac{v_A}{c}\right)^2 = \frac{2(H^2/8\pi)}{\rho c^2}. \qquad (28)$$

This expression is thus equal to the ratio of twice the magnetic energy density to the particle rest energy per cm³. In plasmas of practical interest for controlled fusion reactions this ratio is much less than 1, *i.e.*, the waves are much slower than light waves.[20]

One reason for describing the hydromagnetic oscillations is that they illustrate an important attribute of a hot plasma in a magnetic field, *i.e.*, that of "sticking" to magnetic lines of force (or at least to magnetic surfaces). Also, unstable modes of these same oscillations can be important in the general problem of the stability of a plasma confined by a magnetic field. For example,

(28) can be seen to represent the square of an index of refraction of the plasma for Alfvén waves, so that the reciprocal of the quantity $(v_A/c)^2$ itself resembles an effective dielectric constant of the medium for some types of hydromagnetic disturbances. This dielectric constant may be very large in cases of practical interest.

### Collision Processes

It has already been shown that the plasma of a controlled fusion reactor could not exist in equilibrium with its own radiation field. Furthermore it obviously must not exist in thermal equilibrium with surrounding material walls. Thus such a confined plasma may not even approach kinetic thermodynamic equilibrium in the usual sense. For example, the existence of a magnetically confined plasma must be considered as a nonequilibrium state, since confining electrodynamic forces may only exist if electrical currents flow within the plasma. Since a plasma is not a perfect conductor these confining currents will decay with time in an isolated plasma system, or may be maintained in a steady state only by a continuous input of energy, either kinetic or electromagnetic. Just as in the case of ordinary conductors, collisions provide the dissipative means and determine the time rate of approach to equilibrium. However, in contrast to ordinary conductors, collision rates in a plasma become slower as the temperature is increased, thus facilitating magnetic containment.

In a totally ionized gas, scattering, or collisional, processes are almost entirely ascribable to the Coulomb fields of the bare nuclei and free electrons of the plasma. Because of the well-known infinite effective range of a Coulomb field, it is usual to divide the particle Coulomb interactions into two general types, a division which has already been indicated. The combined effect of all particles beyond a certain cutoff or screening distance is lumped into a collective space-charge-like interaction where only gross charge effects are important. On the other hand, within this cutoff distance it is physically acceptable to discuss "collisions" as discrete independent events, even though in general many particles will lie within the unscreened region. The screening length ("Debye length") is determined in effect by the distance beyond which the surrounding electrons of the plasma can, by their collective motion, screen off the Coulomb field of a particular charge from that of another one moving by it. Thus the Debye length is closely related to the minimum response time of the plasma to a localized electrical transient, *i.e.*, the ordinary plasma frequency. This time is of course about equal to $1/\omega_p = (m/4\pi n_e e^2)^{1/2}$. The distance, $\lambda$, that the average electrons of the plasma can move in this time in attempting to screen the field is about equal to the mean electron velocity times the above time, *i.e.*,

$$\lambda = \left(\frac{\frac{3}{2}kT_e}{m}\right)^{1/2} \cdot \left(\frac{m}{4\pi n_e e^2}\right) \;=\; \left(\frac{kT_e}{(8/3)\pi n_e e^2}\right)^{1/2}. \qquad (29)$$

---

[19] W. A. Newcomb has constructed a comprehensive theoretical treatment of the *motion of magnetic lines*, which is reproduced in Princeton Univ. Observ. Tech. Rep. no. 1, 1955, and which rigorously treats this concept.

[20] F. de Hoffmann and E. Teller, "Magneto-hydrodynamic shocks," *Phys. Rev.*, vol. 80, pp. 692–703; November, 1950, have shown that in the limit of very high magnetic fields or low density Alfvén waves becomes ordinary light waves.

$\lambda$ is thus a crude estimate of the screening distance. It is more usual to use Debye's theoretical value (derived in connection with the theory of solids) which is slightly smaller:

$$\lambda_D = \left(\frac{kT_e}{4\pi n_e e^2}\right)^{1/2}. \qquad (30)$$

This quantity only appears in the argument of a logarithm in all quantities of physical interest, so that its exact numerical value is not important. In Fig. 8, $\lambda_D$ is
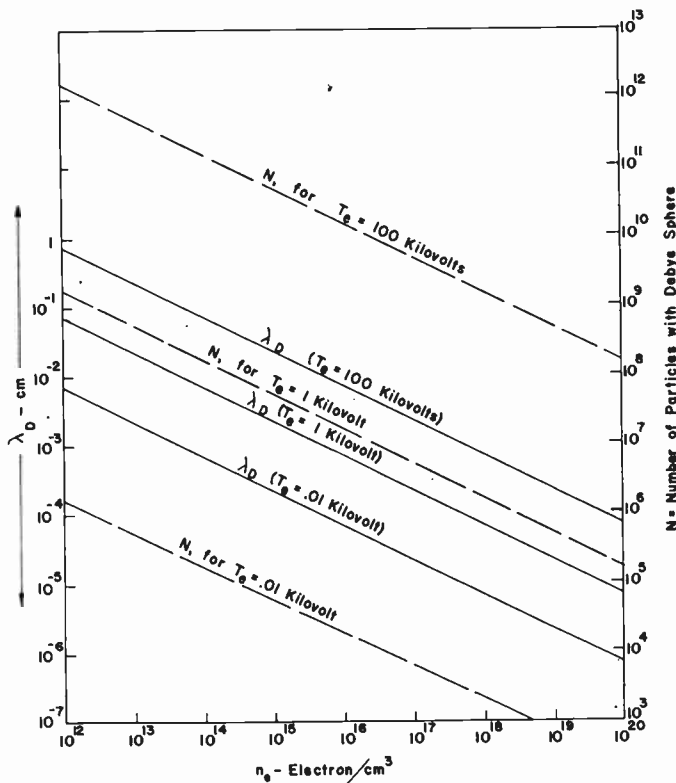


Fig. 8—Debye length and number of particles within a Debye sphere as a function of particle density.

plotted as a function of $n_e$ for various electron temperatures. Also plotted are the average number of charges lying within the Debye sphere (a sphere with radius equal to the Debye length). It can be seen that any particular particle may actually interact with many others in passing through a distance equal to the Debye length. For this reason it will turn out that the effect of "distant" collisions (far compared to nuclear dimensions, but still within a Debye length) is more important than close encounters in producing scattering of any particular charged particle as it passes through the plasma.

First consider close collisions. Now the collision cross section for a deflection of one charged particle by another is given by the classical Rutherford relation.[21] For a "large" deflection owing to a single close collision, an

[21] N. F. Mott and H. S. W. Massey, "Theory of Atomic Collisions," Oxford University Press, New York, N. Y.; 1949.

easy way to find the order of magnitude of the collision cross section is to find the distance of closest approach between two charged particles (at which the mutual Coulomb potential energy is about equal to the initial relative kinetic energy of the colliding particles). Particles which pass about this close together will evidently be scattered through a large angle, and the cross section for such a process will be about equal to the area of a disk with a radius equal to the above distance. Thus we set initial particle relative kinetic energy $W$ equal to the mutual electrostatic potential energy. For two singly charged particles then,

$$W = e^2/r_c,$$

with $r_c$ being the distance of closest approach of classical particles. Hence

$$\sigma_c \approx \pi r_c^2 = \pi e^4/W^2. \qquad (31)$$

Expressing $W$ in kilovolts, we find

$$\sigma_c \approx \frac{6 \times 10^{-20}}{W^2} \text{ cm}^2. \qquad (32)$$

The cross section $\sigma_c$ is seen to vary inversely with the square of the relative energy. At 100 kev relative energy it is about 6 barns, *i.e.*, about 160 times the DD reaction cross section at the same relative energy. At much higher energies the above relationship underestimates the cross section because of direct nuclear interaction effects.

To calculate the effect of distant collisions we may assume that the deflecting effect of each such collision is small and is random in its direction relative to the motion of the scattered particle. Now the incremental impulse given to a particle in such encounters is the mean force times the mean time of an encounter

$$F(\Delta t) = \Delta(Mv) = \Delta p, \qquad (33)$$

($\Delta p = $ impulse; $p = $ momentum). The force is just that arising from the moving charge interacting with the electric field of the "background" particle with which it is "colliding." Thus $F = Ee$, and $\Delta t$ is about equal to the time that the particle spends in going through a distance equal to its distance of closest approach as shown in Fig. 9. Thus $\Delta t \approx r/v$ and $R \approx e/r^2$. ($v = $ particle velocity.)
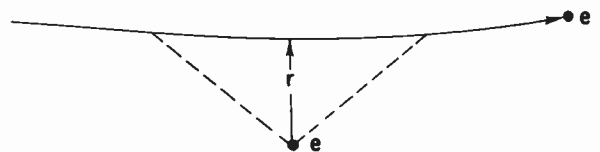


Fig. 9—Collision geometry for distant collisions.

$$\Delta p \approx Ee \cdot r/v = e^2/rv. \qquad (34)$$

Since the distant encounters produce momentum transfers which are essentially random in direction, the mean square momentum change will be proportional to the

momentum change per encounter. As the scattered particle moves a distance $L$ it "collides" with a number of equal mass particles lying in a shell at distance $r$ given by the particle density times the volume of such a shell: $dN_L = n(2\pi r L dr)$, thus for random encounters,

$$\langle (\Delta p)^2 \rangle_{Av} = (nL2\pi r dr) \cdot (e^2/rv)^2, \qquad (35)$$

we can find the total momentum change in path length $L$ by integrating over all collisions, *i.e.*, integrating over $r$ between the distance of closest and most distant approach.

$$\langle (\Delta p)^2 \rangle_{Av} = \frac{2\pi e^4}{v^2} nL \log(q) \text{ where } q = (r\max/r\min). \quad (36)$$

Let us divide both sides by the square of the initial momentum of the scattered particle $p^2 = (Mv)^2$.

$$\frac{\langle (\Delta p)^2 \rangle_{Av}}{p^2} = \frac{\pi e^4}{2(\tfrac{1}{2}Mv^2)^2} nL \log(q). \qquad (37)$$

When $\langle (\Delta p)^2 \rangle_{Av}$ has increased to the point where it is comparable to $p^2$, we can assume that the particle has been scattered through a large angle, such as 90° (or will have had its energy changed by a substantial amount). In other words, if $L$ is thought of as the mean free path for large angle scattering of a charged particle by distant collisions, then $L$ is found by setting $\langle (\Delta p)^2 \rangle_{Av}/p^2 = 1$

$$1 \approx \frac{\pi}{2} \frac{e^4}{(\tfrac{1}{2}Mv^2)^2} nL \log(q) = \frac{\pi}{2} \frac{e^4}{W^2} nL \log(q).$$

But $L = 1/n\sigma$ from the meaning of a cross section. Thus we find for the effective cross section for large angle scattering or substantial energy transfer due to distant collisions the expression

$$\sigma_d \approx \frac{\pi e^4}{W^2} \frac{\log(q)}{2}. \qquad (38)$$

It is seen that $\sigma_d$ is just $\sigma_c$ multiplied by the factor $\log(q)/2$. The factor $q$ is very large but bounded. It is given roughly by the ratio of the greatest distance of interaction between two colliding particles and their distance of closest approach. The former distance is evidently set by the Debye length. The latter is about equal to $r_c$ calculated above, if quantum effects are neglected. Thus $q \approx \lambda_D/r_c$ which is a number of order $10^9$ for typical plasma densities and temperatures. Thus $\log(q) \approx 20$, and is relatively insensitive to changes in plasma density and temperature. If we disregard corrections for relative energy of the background particles, we find $\sigma_d \approx 10\sigma_c$. Thus the accumulative effects of distant collisions are about an order of magnitude more effective in producing scattering of a charged particle than are close encounters. With energy $W$ in kilovolts,

$$\sigma_d \approx \frac{6 \times 10^{-19}}{W^2} \text{ cm}^2. \qquad (39)$$

This relationship may be expected to hold approximately for encounters between equal mass particles, or where the scattered particle is lighter than the scattering particles. Thus ion-ion, electron-electron, and electron-ion scattering effects are roughly estimable from $\sigma_d$ above. Detailed and more accurate computations of this type, for the case of stellar encounters, have been made by the astrophysicist Chandrasekhar.[22] Spitzer[18] has extended these calculations to the case of a fully ionized plasma.

One of the important consequences of the work of Chandrasekhar and of Spitzer is in the introduction of the notion of collisional *relaxation times* of a plasma. It is clear that the mean time for a large scattering in angle or change in momentum of a particle interacting with a background of other particles is a measure of the rate of attempted approach of a nonequilibrium plasma distribution to equilibrium.

Two separate cases are of particular interest. The first is concerned with the time for a particular energetic ion or electron (the "test particle") to be scattered in direction (or changed in energy) by encounters with other like particles.

The cross section $\sigma_d$ calculated above may be used to estimate these relaxation times.[23] These times are, respectively, $t_{ii}$, the mean collision time for an ion scattered by distant encounters with other ions, $t_{ee}$, a similar time for an electron scattered by electrons, and $t_{ei}$, the time for scattering of an electron by distant encounters with ions. Evidently $t_{ei} \approx t_{ee}$ if recoil effects are neglected. Thus (subscripts $i$ and $e$ refer to ions and electrons respectively)

$$t_{ii} = \frac{1}{n_i \sigma_d v_i}, \qquad (40)$$

$$t_{ee} = \frac{1}{n_e \sigma_d v_e} \approx t_{ei}. \qquad (41)$$

Because an electron has about 60 times the velocity of a deuteron of equal energy, $t_{ii} \approx 60 t_{ee}$ for comparable particle energies. By inserting the value of $\sigma_d$ above one may express the energy exchange times as a function of background particle density and scattered particle energy in kilovolts. Using the mass of the deuteron for $M$, so that $v_i = 3 \times 10^7 \sqrt{W}$ cm/sec, we have

$$t_{ii} = 6 \times 10^{10} \frac{W^{3/2}}{n_i} \text{ seconds,} \qquad (42)$$

$$t_{ee} = 10^9 \frac{W_e^{3/2}}{n_e}. \qquad (43)$$

Note that the relaxation times become *longer* at high energies. (These equations should be viewed only as estimates of the relaxation times, since differences of a

[22] S. Chandrasekhar, "Principles of Stellar Dynamics," University of Chicago Press, Chicago, Ill.; 1942.
[23] For more precise values the reader is referred to the original papers by Chandrasekhar and Spitzer.

factor of 2 or more with respect to accurately derived values are possible.)

From the equations it can be seen that, for example, at $n_i = 3 \times 10^{15}$ and $W = 150$ kev, $t_{ii} = 0.04$ second. (The corresponding mean free path is about $1 \times 10^7$ cm.) From Fig. 4, the mean reaction time for DD at 100 kilovolts kinetic temperature is seen to be 10 seconds or about 250 times the relaxation time of a deuteron at the mean particle energy appropriate to that temperature. This implies that at plasma temperatures of about 100 kilovolts a deuteron must on the average undergo a considerable but not extremely large number of effective scattering events before reacting. At lower energies this number is much larger.

The second case of interest pertains to the effect of collisions of an ion with background electrons. In this case two limiting situations are possible: 1) the ion energy is much less than the mean electron energy and 2) the ion energy lies above the mean electron energy. In case 1) the ion will gain energy by encounters with the more rapidly moving electrons at a rate essentially determined by the electron-ion collision rate and the mean energy transfer per collision. In 2), the ion will lose energy to the electrons, even though their mean velocity is greater than that of the ion.

The rate for 1) may be estimated by methods similar to those used to determine $\sigma_d$. Since the ions move very slowly compared to the average electron velocity, the problem is very much like the classical Brownian motion problem. As fast electrons pass in the vicinity of the ion they give it an incremental momentum kick of the kind described in the calculation of $\sigma_d$, except that here the mean electron velocity must be used.

$$\Delta p \approx e^2/r\bar{v}_e, \tag{44}$$

where $\bar{v}_e$ is the mean electron velocity.

The number of such collisions in time $\Delta t$ from electrons passing nearby, say within a ring of area $dA = 2\pi r dr$, is effectively proportional to the mean number of electrons per second passing through such a ring. From elementary kinetic theory this number can be set approximately equal to $3n_e\bar{v}_e dA$. Assuming that the encounters are random and integrating over $r$ one obtains

$$\langle(\Delta p)^2\rangle_{\text{Av}} = \frac{6\pi n_e e^4}{\bar{v}_e} \log (q)\Delta t. \tag{45}$$

Since $(1/2)m\bar{v}_e^2 \approx (3/2)kT_e$, and $\langle(\Delta p)^2\rangle_{\text{Av}} = 2M\langle\Delta W\rangle_{\text{Av}}$, substitution of these quantities in (45) yields for the rate of energy gain by an ion bombarded by energetic electrons ($W \ll (3/2)kT_e$, all units cgs)

$$\frac{dW}{dt} \approx 3\pi \log (q) \frac{n_e e^4}{(3m_e kT_e)^{1/2}} \left(\frac{m}{M}\right). \tag{46}$$

This expression shows that the rate of gain of energy of an ion is independent of its energy (as long as it is small compared to the mean electron energy) and that

it varies inversely with the square root of the electron temperature. This qualitative fact is of some interest in connection with the possibility of heating "cold" ions by collision with "hot" electrons. Since it is often easier to impart energy directly to the electrons of a plasma than to the ions, a possible heating cycle would be to heat electrons and then allow them to impart energy to the ions by collisions. This is likely feasible at energies of a few volts, but at fusion temperatures the energy transfer rate becomes slower and makes the heating less effective.

In the discussion of electron-ion collision effects, the other limiting case of interest is that where the ion energy is large compared to the mean electron energy. Here the result, not as easily obtainable by approximate methods, is that even though the mean electron velocity may be greater than the ion velocity, the energy transfer is in fact dominated by those few electrons of the distribution which have a lower velocity than the ion, and the effect of these collisions is always to reduce the ion energy until it reaches the mean electron energy. Spitzer and others have calculated exact expressions for energy transfer between ions and a Maxwellian electron distribution. A form is given below that encompasses both of the cases here considered.[24] This expression is for a hydrogenic plasma: (cgs units)

$$\frac{dW}{dt} = 4\pi\sqrt{2} \log (q) \frac{n_e e^4}{(\pi m kT_e)^{1/2}} \left(\frac{m}{M}\right)\left(1 - \frac{W}{\frac{3}{2}kT_e}\right). \tag{47}$$

When the ion energy $W$ is much less than $\frac{3}{2}kT_e$, this expression reduces to the same form as the approximate (46) above, apart from a numerical factor of $(3\pi/32)^{1/2}$ or 0.54. For this limiting case, inserting numerical values into (47), expressing $W$ and $kT_e$ in kilovolts, and setting $M$ equal to the mass of a deuteron, one obtains

$$\frac{dW}{dt} = 1.4 \times 10^{-12} \frac{n_e}{T_e^{1/2}} \text{ kilovolts/sec.} \tag{48}$$

At $n_e = 3 \times 10^{15}$ and $T_e = 0.1$ kilovolt (such as might be obtained in a high current gas discharge), a low energy ion will gain energy at a rate of $1.3 \times 10^4$ kilovolts/second so that a 1 volt ion may double its energy in about 0.1 $\mu$sec. However, at the same density and 100 kev the rate is 430 kilovolts/sec; a 10-kev ion will double its energy only after 0.023 sec.

When $W \gg kT_e$, (47) becomes (energy and temperature in kilovolts)

$$\frac{1}{W}\frac{dW}{dt} = -0.9 \times 10^{-12}\frac{n_e}{T_e^{3/2}}. \tag{49}$$

This may be integrated immediately to yield

$$W = W_0 e^{-t/\tau_e}, \tag{50}$$

where $\tau_e = 1.1 \times 10^{12}(T_e^{3/2}/n_e)$. Collisions with a "cold" electron distribution are seen to exponentially "damp"

[24] M. H. Johnson, private communication.

the energy of a fast ion with a time constant proportional to the three-halves power of the electron temperature, so that the ion approaches asymptotically to the energy $\frac{3}{2}kT_e$, in agreement with classical equipartition considerations. At $n_e = 3 \times 10^{15}$ and an electron temperature of 10 volts (typical of ordinary discharges), $\tau_e$ is only 0.3 microsecond or about $3 \times 10^{-7}$ times the mean reaction time for DD at 100 kilovolts temperature. This is of importance to certain problems discussed in a later section.

One of the interesting consequences of the general (47) is that at high plasma temperatures the "equilibrium" electron temperature may lag somewhat behind the ion temperature. If no appreciable energy is being fed into the plasma from reactions or external sources, the electron temperature will be determined by a balance between the rate of bremsstrahlung radiation and the rate of collisional energy transfer from the ions, found by combining (47) and (12). For a deuterium plasma the resulting temperature difference is given approximately by the equation (temperatures in kilovolts)

$$T_i - T_e = 4.4 \times 10^{-3} T_e^2. \tag{51}$$

Thus if $T_i = 100$ kev, $T_e$ would asymptotically approach 75 kev in the absence of other energy exchange processes.

*Particle Motions*

At fusion temperature the mean free path for collisions of ions and electrons in a reacting plasma will likely be very long, as already noted. In this case the motion of an ion or electron in the plasma through distances of the order of the apparatus dimensions will be largely dominated by electrodynamic forces. Each charged particle of the plasma will move nearly in accordance with the usual equation of motion of a particle in an electromagnetic field, where here the field is the resultant of the self-consistent superposition of externally generated fields and the space-charge and motional fields produced by the plasma itself. A sufficiently weak plasma may have a relatively minor effect on an externally generated field and a picture of its behavior may be deduced from the motion of individual particles in the applied field.

The equation of motion of a charged particle in an electromagnetic field is, in Gaussian units,

$$m \frac{dv}{dt} = e\left(E + \frac{v}{c} \times H\right). \tag{52}$$

Electric forces with a component parallel to the motion may produce a change in kinetic energy. Magnetic forces always act in a direction perpendicular to the motion and thus produce curvature but no change in energy.

If $E = 0$ and $H$ is constant in space and time, the motion consists of the superposition of an arbitrary constant velocity component along the magnetic lines and a rotation at the cyclotron angular frequency, $\omega_c = eH/mc$. In other words the orbits described are helices parallel to the direction of the field. Because of their opposite sign of charge, ions and electrons rotate in opposite directions, with frequencies in the ratio of $(m/M)$, respectively. For the same rotational energy, ions execute orbits which are $(M/m)^{1/2}$ times the diameter of the electron orbits (about 60 times in the case of deuterons). The product of magnetic field and ion orbit radius of curvature is for deuterons (rotational energy in kilovolts)

$$H\rho_c = 6.4 \times 10^3 \sqrt{W_\perp} \text{ gauss-cm.} \tag{53}$$

Also for a deuteron,

$$\omega_c = 4.8 \times 10^3 H \text{ radians/sec.} \tag{54}$$

When an electric field is present with a component perpendicular to the magnetic field the particle path consists of a helical motion and a superposed drift at constant velocity in a direction perpendicular both to the magnetic field and the transverse electric field component. When this drift velocity is small compared to the particle orbit velocity the actual motion may be thought of as arising from a transverse drift of the instantaneous center of rotation (guiding center) of the particle. This important drift velocity is given by

$$v_0 = c \frac{E \times H}{H^2} = c \frac{E}{H}, \tag{55}$$

if $E$ and $H$ are perpendicular. Note that its direction and magnitude are independent of the magnitude or sign of charge, or the mass of the particle.

If it is remembered that even in a static magnetic field a moving observer will see an electric field, the drift velocity $v_0$ may be readily deduced from the equation of motion. Assume that the actual particle motion consists of a uniform velocity $v_0' \ll c$ plus a variable part $v_1$, so that $v = v_0' + v_1$, *i.e.*, viewed from a coordinate system moving at velocity $v_0'$, the particles' velocity is $v_1$. In such a moving coordinate system there will appear in addition to the applied electric field a motional electric field $E_m = (v_0' \times H)/c$. In the moving frame therefore the equation of motion is

$$m \frac{dv_1}{dt} = e\left(E + E_m + \frac{v_1 \times H}{c}\right). \tag{56}$$

Pick for the velocity $v_0'$ that velocity where $E_m = -E$, so that $E + E_m = 0$. In this frame there is no electric field and the orbits are again helices. Picking the velocity

$$v_0' = v_0 = c \frac{(E \times H)}{H^2},$$

then one finds

$$E_m = \frac{(E \times H) \times H}{H^2} = -E,$$

since $H$ and $E$ are perpendicular.

The drift velocity $c(E/H)$ has another important meaning in connection with a magnetic field which varies with time. Suppose that a uniform magnetic field produced by a long solenoid is increasing with time. In this case an electric field appears which consists of circular lines of electric force centered on the axis of the solenoid. The magnitude of this field in the laboratory frame is found from the usual transformer equation

$$\oint E \cdot dl = \frac{-1}{c} \int \left(\frac{dH}{dt}\right) \cdot dA, \qquad (57)$$

*i.e.*, at a radius $r$

$$2\pi r E = \frac{-\pi r^2}{c} \frac{dH}{dt} . \qquad (58)$$

At any radius $r$ the drift velocity $v_0 = c(E/H)$ may be evaluated.

$$v_0 = dr/dt = c(E/H) = -(r/2)(dH/dt)(1/H), \quad (59)$$

*i.e.*,

$$(1/r)(dr/dt) = -\tfrac{1}{2}(1/H)(dH/dt). \qquad (60)$$

This can be integrated directly to give

$$\frac{r}{r_0} = \left(\frac{H_0}{H}\right)^{1/2}, \qquad (61)$$

or

$$\pi r^2 H = \pi r_0^2 H_0, \qquad (62)$$

showing that an inward motion of any part of the region at the local drift velocity $c(E/H)$ preserves flux between the moving point and the axis—*i.e.*, the guiding centers of particles moving at the local drift velocity stay on the surface of some collapsing flux tube of the magnetic field.

In the presence of magnetic field gradients more complicated motion of the particles will occur. When the gradients are small the motion can be predicted by simple relationships. The case of a magnetic field with a gradient perpendicular to the local direction of field lines is of particular interest. The qualitative effect of such a gradient is shown in Fig. 10. Where the field is strong, radii of curvature are smaller than average; where weak, they are larger. Thus a cycloidal "walk" results, with the guiding centers of positively and negatively charged particles moving oppositely and perpendicular to the direction of the gradient. In a plasma this results in a tendency for charge separation to occur, which in turn can produce local electric fields. The eventual result may be that the plasma drifts as a whole in the "downhill" direction of the transverse gradient. The details of the actual resultant plasma drift motion are best treated by the macroscopic equations of motion of the plasma, for example in the form discussed by Spitzer.[13] The velocity of the individual particle drifts

(if uninhibited by collective effects) is[14]

$$v_d = \frac{\rho_c v_\perp}{2} \frac{\nabla_\perp(H)}{H}, \qquad (63)$$

where $\Delta_\perp(H)$ is the gradient of the scalar magnitude of $H$ in a plane perpendicular to $H$, $\rho_c$ is the radius of curvature of the particle, and $v_\perp$ is its rotational velocity component. The expression is valid only if $v_d \ll v_\perp$, *i.e.* $(\rho_c/2)(\Delta_\perp(H)/H) \ll 1$. The latter condition merely implies that the fractional change in magnetic field strength across an orbit shall be small.
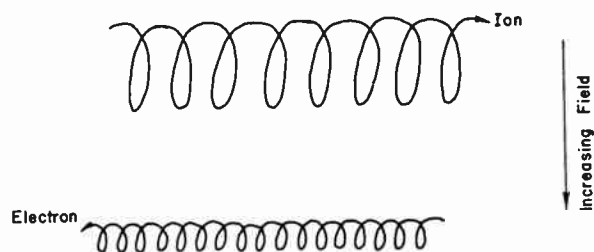


Fig. 10—Particle drifts in a magnetic field having a gradient transverse to the direction of the field.

Drifts which might be called "gyroscopic" in origin can arise from gravitational or centrifugal forces. These forces lead to drift motions perpendicular to the applied force and in opposite directions for ions and electrons. In a gravity field with a component $g_\perp$ perpendicular to the magnetic field the drift is [14]

$$v_g = \frac{g_\perp}{\omega}, \qquad (64)$$

$\omega_c$ is the cyclotron frequency of the particle in the magnetic field. For high magnetic fields this drift velocity is small.

If particles are moving along lines of force which are curved, a centrifugal acceleration will arise which will also result in a drift, calculable from (64) above by replacing $g_\perp$ by the centrifugal acceleration.[13] Suppose that the particle is executing a helical path along the curved field lines with a velocity component along the field lines of magnitude $v_{||}$. Then if the bundle of magnetic lines are curved with a radius of curvature $R$, the centrifugal acceleration is $v_{||}^2/R$, so that

$$v_c = v_{||}^2/R\omega_c. \qquad (65)$$

In contrast to gravitational drifts, this drift velocity may be substantial. Note that the guiding center drifts $v_d$, $v_g$, and $v_c$ are oppositely directed for ions and electrons and may give rise to charge separation and consequent electric field effects. In many such cases the end result may be that the plasma takes up a motion in the direction of weakest magnetic field—*i.e.*, it tends to be repelled from regions of strong magnetic field.

This apparent tendency for charged particles moving in a magnetic field to be repelled from regions of strong

magnetic field—*i.e.*, to behave diamagnetically—is also exhibited by particles moving in a magnetic field with a gradient parallel to the direction of the field lines. The repelling of charged cosmic-ray particles by the dipole magnetic field of the earth is a well-known effect. Fermi has applied similar concepts also to explain the origin of cosmic rays.[25] The repelling effect of a positive field gradient parallel to the magnetic field lines is most easily explained in terms of one of the so-called "adiabatic invariants" of the motion of a charged particle. Since particles moving in a static magnetic field spiral about lines of force and are acted on only by forces perpendicular to their motion, the angular momentum associated with their rotational motion about a line of force should be an approximate constant of the motion. Also, the rotational motion of a charge must generate a magnetic moment which will also be a constant of the motion. This can be shown from simple considerations.

The usual cyclotron equation for balance between centrifugal and magnetic forces on a circularly moving particle is

$$\frac{mv_\perp^2}{r} = \frac{Hev_\perp}{c}, \tag{66}$$

which can be rewritten as

$$\frac{\tfrac{1}{2}mv_\perp^2}{H} = \frac{e}{2mc}(mv_\perp r), \tag{67}$$

$mv_\perp r$ is the angular momentum of the rotational motion, call it $\alpha$; $\tfrac{1}{2}mv_\perp^2$ is the rotational energy of the particle; $e/2mc$ is a constant. Thus

$$\frac{\tfrac{1}{2}mv_\perp^2}{H} = \frac{W_\perp}{H} = \frac{e}{2mc}(\alpha) = \mu, \tag{68}$$

where $\mu$ is the magnetic moment in ergs/gauss. The constancy of $\alpha$ implies the constancy of $\mu$ and thus $W_\perp/H$ represents an adiabatic invariant of the motion. An "adiabatic invariant" is here to be understood as a quantity which is constant for slow changes in the magnetic field at the particle during a single rotation. $W_\perp/H$ has also been shown to be an invariant for magnetic fields varying slowly with time.[12],[26]

For charged particles moving freely in a static magnetic field another invariant of the motion is the total kinetic energy, *i.e.*, in moving along magnetic field lines the sum of rotational and translational energy should be a constant. If the magnetic field strengths at two different regions (1) and (2) encountered by the particle are $H_1$ and $H_2$ respectively, then conservation of energy requires that

$$W_\parallel(1) + W_\perp(1) = W_\parallel(2) + W_\perp(2). \tag{69}$$

The invariance of $\mu$ also implies that

[25] E. Fermi, "Galactic magnetic fields and the origin of cosmic radiation," *Astrophys. J.* vol. 119, pp. 1–6; January, 1954.
[26] G. Hellwig, "Über die bewegung geladeuer teilchen nin schwack veranderlichen magnetfelderm," *Z. Naturforsch.* vol. 10a, p. 508; July, 1955.

$$\frac{W_\perp(1)}{H_1} = \frac{W_\perp(2)}{H_2}, \tag{70}$$

whence

$$W_\parallel(2) = W_\parallel(1) - W_\perp(1)\left(\frac{H_2}{H_1} - 1\right), \tag{71}$$

if $H_2/H_1 > 1$, then $W_\parallel(2) < W_\parallel(1)$. Evidently the particle slows down its velocity component along the field lines in moving into a stronger field. If the initial $W_\parallel$ is not too great a particle may be turned around by a positive field gradient, which acts therefore somewhat like a potential. Since force is the negative gradient of potential, differentiation of (71) yields an approximate equation for axial motion of a particle spiraling along a magnetic line.

$$F_z = - \mu\frac{\partial H_z}{\partial z}, \tag{72}$$

where $z$ is parallel to the local direction of the magnetic field lines.

It is apparent from the equations above that particles which move with too small a component of rotational energy will not be turned around by the field.

The constancy of $\mu$ has another interesting corollary. Since $W_\perp/H = $ constant and $H\rho_c \sim W_\perp^{1/2}$ from (53), it follows that $\rho_c^2 H = $ constant, *i.e.*, the flux through the orbit circle is an approximate constant of the motion.

## Diffusion Across a Magnetic Field

Although of heuristic value, the picture of a plasma as composed of an assembly of independently moving charged particles subject only to simple adiabatic laws is an over-simplification. For many problems the macroscopic equations of the plasma should be used.[13] An example of a problem which is most simply analyzed by the macroscopic or hydrodynamical equations is the diffusion of charged particles across a strong magnetic field (*i.e.*, where the particle pressure is negligible compared to the magnetic pressure).

Microscopically, the diffusion of charged particles across a magnetic field is to be understood as a two-dimensional random walk of particles at a rate determined by their mutual collision frequency and with a step length roughly equal to the cyclotron radius. Thus, qualitatively, the diffusion rate will be lowered either, by 1) raising magnetic field (smaller orbit diameter), or 2) increasing temperature or reducing density (lower collision frequency).

In a quiescent plasma composed of electrons and identical ions the diffusion may originate either from electron-ion, ion-ion, or electron-electron collisions. Simon[27] has recently calculated the relative importance of these processes and has apparently resolved a paradox. The paradox is that a first-order treatment predicts that collisions between like particles lead to no net

[27] A. Simon, "Diffusion of like particles across a magnetic field," *Phys. Rev.*, vol. 100, pp. 1557–1559; December, 1955.

diffusion (since the center of mass remains fixed in such collisions).

The usually dominant term is the effect of electron-ion collisions for which the diffusion velocity is given by[18]

$$v_p = \frac{-c}{\sigma} \frac{\nabla P}{H^2} \text{ cm/sec}, \qquad (73)$$

where $\sigma$ is the electrical conductivity of the plasma in $cm^{-1}$ (emu) and $P$ is the plasma pressure. Expressed as a function of temperature in kilovolts for a deuterium plasma this is approximately

$$v_p \approx \frac{6.3 \times 10^3}{T^{3/2}} \frac{\nabla P}{H^2}. \qquad (74)$$

If $P$ is assumed to be some constant small fraction $\beta$ of the magnetic pressure term $H^2/8\pi$, and the characteristic distance associated with $\nabla P$ is denoted by $L$, then

$$v_p \approx \frac{1.5 \times 10^5}{T^{3/2}} \left(\frac{\beta}{L}\right) \text{ cm/sec}. \qquad (75)$$

For $T = 100$ kilovolts for example $v_p \approx 150(\beta/L)$ cm/sec corresponding to a very slow diffusion rate, and a correspondingly great reduction in heat transport across the field.

Simon[27] has shown theoretically that under certain circumstances diffusion owing to collisions between like particles can be of importance. However, in this case the diffusion rate varies as $H^{-4}$ and is not simply proportional to $\nabla P$.

The expressions above are derived under the assumption that transport of particles across a magnetic field arises only from mutual collisions. However, some study several years ago at the University of California Radiation Laboratory[28] of the diffusion of ions across the plasma of arc discharges in a magnetic field seemed to indicate that a more rapid diffusion than predicted above could occur. It was postulated by Bohm and co-workers at the Laboratory that random electrostatic fields arising from turbulent types of plasma oscillations were responsible for these excessive drifts. Bohm proposed an expression for the diffusion velocity; (temperature in kilovolts)

$$v_B = 6 \times 10^9 T_e \frac{\nabla n}{n} \cdot \frac{1}{H} \text{ cm/sec}. \qquad (76)$$

This depends directly on electron temperature and inversely only as the first power of magnetic field. It predicts a much higher diffusion loss rate than (74) above and correspondingly greater heat transport.

Some recent experimental work by Simon and Neidigh (unpublished) at ORNL seems to indicate that the Bohm diffusion mechanism, if it exists, may be associ-

ated with highly unsymmetric plasma systems. Their work, carried out in arc discharges with axial symmetry, seems to bear out the $1/H^2$ dependence of (74). However, the assumption of a nonturbulent plasma may be far from the truth in many situations of interest to controlled fusion research. An example is the instability of the pinch effect, mentioned in the preceding section.

### Compression of a Plasma

To conclude the discussion of phenomena in totally ionized media, mention should be made of the effects associated with compression of a plasma. Depending on the nature of the compression and its time-scale a plasma can act either as a one-, two-, or three-dimensional gas. This unusual circumstance arises because processes which are slow compared to the individual particle motions may still be rapid compared to the collisional relaxation times of the plasma particles. When this is the case the gas kinetic degrees of freedom of the plasma are effectively decoupled from each other. If such a plasma is confined and adiabatically compressed in one dimension only, then the energy in this degree of freedom will increase in accordance with the adiabatic law for a one-dimensional gas. For compression across a magnetic field two degrees of freedom are involved and the behavior is as a two-dimensional gas. In general, for a gas with number of degrees of freedom $f$, the variation of kinetic temperature $T$ is given by

$$T_f \sim n^{\gamma-1}, \qquad (77)$$

where $T_f$ is to be understood as the generalized kinetic temperature associated with the compression. Also,

$$\gamma = (2 + f)/f, \qquad (78)$$

and $n$ is the particle density.

For $f = 1$, $\gamma = 3$, and $T_1 \sim n^2$. This fact is noted by Fermi in his theory of the origin of cosmic rays,[25] where the acceleration mechanism he proposed arises from the repeated reflection of energetic particles from two approaching magnetic clouds at a variable distance $L$ from each other, so that $T_1 \sim 1/L^2$.

When the compression is two-dimensional, $f = 2$ and $\gamma = 2$, so that $T_2 \sim W_1 \sim n$. When the compression is adiabatic and is carried out in a time long compared to the relaxation times, $f = 3$, $\gamma = 5/3$, and $T \sim n^{2/3}$ as in an ordinary gas.

The existence of adiabatic compressional processes in a plasma confined by a magnetic field suggests strongly that such processes may be reversible and therefore could be useful for the extraction of energy from a reacting plasma by its expansion against the magnetic field. It is evident that one result of this expansion could be the generation of electrical energy in external coupled electrical circuits, possibly at very high thermodynamic efficiencies.

Nonadiabatic, *i.e.*, irreversible, shock-hydrodynamic effects can also occur in a plasma. Here the physical

[28] A. Guthrie and R. K. Wakerling, "The Characteristics of Electric Discharges," McGraw-Hill Book Company, Inc., New York, N. Y.; 1949.

situation is considerably complicated by the presence of a magnetic field and new types of phenomena can occur. De Hoffmann and Teller[29] have treated some of these, and new papers are beginning to appear on the subject. This facet of plasma physics is a new and interesting one, and is of relevance to some aspects of controlled fusion research, for example the transient pinch effect mentioned in a preceding section.

Summarizing, a plasma at thermonuclear temperatures may be qualitatively pictured as a mixture of two charged particle gases—the ions and the electrons. These two gases are only weakly coupled through collisions, in the small, but strongly coupled in the large through Coulomb forces. Depending on the time scale and nature of application of external electromagnetic forces, either gas may behave as a one-, two-, or three-dimensional ideal gas for compressional processes. Through exhibiting a high effective electrical conductivity, the plasma will tend to move in such a way as to minimize electric fields existing in frames of reference locally at rest with respect to the plasma. This latter property can be identified with a tendency for the plasma to preserve constant flux in its motion, and thus locally to "stick" to magnetic lines of force. As a consequence of this, the guiding centers of the individual particles of a sufficiently weak plasma will tend to move on the surface of flux tubes of a magnetic field, as long as the field varies slowly (but not too slowly) with time. Diffusion arising from collisional effects tends to destroy this situation, so that the diamagnetism of a plasma is of necessity a nonequilibrium phenomenon. In addition, the presence of induced particle drifts may give rise to charge separation and electric fields. These drift effects in turn may thus be self-perpetuating so that in particular cases the plasma may exhibit hydromagnetic instabilities.

### SCALING LAWS—POWER BALANCE

A moment's reflection will show that the phenomena of nuclear fusion and chemical fusion are similar in many ways. Both release energy through a combination or rearrangement of the reactive particles induced by close collisions. Both can be self-sustaining only under favorable physical circumstances; namely, attaining a minimum ignition temperature, exceeding a minimum quantity of fuel, and establishing a sufficiently low rate of energy loss to the surroundings to prevent quenching the reaction. If too small a fire is laid in the fireplace it will not catch; it is difficult to ignite a single piece of coal lying in a grate. These everyday observations about chemical fusion have their counterparts in the nuclear fusion reaction. In this section some general "scaling laws" which may apply to fusion reactors are discussed. Seen in their light, the first achievement of a detectable rate of thermonuclear reactions in a hot plasma might represent much less of a feat than the subsequent attainment of an energetically self-sustaining reaction.

[29] de Hoffmann and Teller, loc. cit.

The search for practical nuclear fusion power will not only depend on establishing theoretical operability, but also in showing that a self-sustaining reactor utilizing the proposed principles would not be too large to be constructed or require unattainable physical conditions for operation.

The achievement of a net power balance from a fusion reactor will depend on a favorable competition between the nuclear energy produced and two types of energy loss processes, "direct" and "indirect." Direct energy losses are those ascribable to escape of radiation or energetic particles from the reacting region. Indirect energy losses are those resulting from external inefficiencies connected with heating and confining the plasma. Failure to control direct losses results in quenching the reaction. Excessive indirect losses would prevent achievement of a closed cycle capable of producing a net power output, even though the internal reaction might be satisfactorily efficient. This latter situation resembles what would happen if the electrical generator on an automobile broke down while it was running. The internal combustion cycle of the automobile engine would continue only until the ignition battery ran down and the spark plugs ceased to fire.

### Scaling Laws for Power Output

To illustrate how power balance scaling relationships might be deduced for realistic schemes, let us make a simplified analysis of an unrealistic one—the stationary pinch example of a previous section. We shall suppose that the Kruskal-Schwarzchild instability does not exist and that the pinch is made so long that the losses to the end electrodes are unimportant. In steady state simple scaling relationships may be found for the nuclear power output. If the radius is $R$ cm, the fusion power per unit length for DD will be

$$p = \tfrac{1}{2}n_D{}^2\langle\sigma v\rangle_{Av}W\cdot\pi R^2. \tag{79}$$

Now the pressure is proportional to the square of the magnetic field, in accordance with (21), from which it follows that $n_D\sim H_0{}^2/T$. Neglecting the numerical constants, one finds the variation of $p$ with radius, temperature and magnetic field from (79) above to be

$$p \sim \left(\frac{\langle\sigma v\rangle_{Av}}{T^2}\right)R^2H_0{}^4. \tag{80}$$

Some conclusions can be drawn from the expression. Firstly, if losses were small, the most efficient operating temperature would be determined by the maximum point of the function $(\langle\sigma v\rangle_{Av}/T^2)$. This function has a broad maximum in the vicinity of 10 to 15 kev for both the DD and DT reactions. Since these temperatures are above the bremsstrahlung critical temperature, $T_c$, of the DT but below that for the DD reaction they represent a possible operating temperature region for the one, but not for the other. Secondly, the strong dependence

of power output on magnetic field is noteworthy, pointing up the advantages of the use of high confining fields, and the probable impracticability of operating any controlled fusion device at small fields. Since (80) was derived from the simple concept of magnetic pressure, similar relationships for the fusion power scaling laws can be expected to occur in the analysis of more complicated examples of magnetic confinement than the simple pinch discussed in this article.

Without a detailed analysis it is not possible to set down more complete expressions for the power loss scaling laws. A few results can, however, be found which are not especially dependent on the example.

### Direct Losses

It has already been shown that the ratio of bremsstrahlung loss to fusion power is independent of density and dimensions. Thus this part of the power balance is unaffected by scaling in physical dimensions or magnetic field (at constant temperature). Other direct losses may arise from the escape of particles out of the system by diffusion. Simple diffusion processes are characterized by a transport velocity which is proportional to the gradient of the density. For example, (74) and (76) of the preceding section for the rate of diffusion of particles across a magnetic field both depend on the gradient of the particle density, though having a different functional dependence on magnetic field.

As a simple example, consider systems which are long compared to their transverse dimensions, and assume that upon scaling in radial dimensions the relative spatial distribution of the plasma is preserved. In such a case the power loss per unit length through radial diffusion will be proportional 1) to the plasma boundary area per unit length, 2) the plasma diffusion velocity, 3) the particle density and, 4) the mean particle energy (proportional to the kinetic temperature).

$$p_L \sim R v_D n T. \tag{81}$$

If the radial diffusion arises from classical collisional effects, so that (74), applies this becomes

$$p_L \sim \frac{n^2 T^{1/2}}{H_0^2}, \tag{82}$$

independent of the radius (since gradient $n \sim n/R$ under a uniform dimensional scaling). $p_L$ is also seen to depend on $n^2$ as does $p$, since it also results from a binary process—collisional diffusion.

Eliminating the particle density through its dependence on confining magnetic field [see (21)] one obtains

$$p_L \sim \frac{H_0^2}{T^{3/2}}. \tag{83}$$

A kind of a figure of merit for the competition of nuclear power with direct losses other than radiation can be found by dividing $p$ by $p_L$

$$\frac{p}{p_L} \sim (\langle \sigma v \rangle_{Av} T^{-1/2}) R^2 H_0^2. \tag{84}$$

Where direct particle losses by collisional radial diffusion are important (84) shows that high temperature operation is favored, and that radial dimension and strength of the magnetic field are of equal and considerable importance in determining the attainment of a favorable power balance.

### Indirect Losses

The calculation of indirect losses is obviously more intricate. It would involve the assessment of the power loads in every auxiliary associated with the fusion reactor, as well as more immediate sources of energy loss such as joule heating of conductors (whether plasma or metallic). In a closed cycle requiring electrical energy to produce and sustain the fusion reaction, the efficiency of conversion of thermal energy derived from the fusion reaction to electrical energy could be an important factor.

One scaling law which should apply to any magnetic confinement scheme has to do with the power required to maintain a magnetic field. If a current density $j$ flows in a conductor (either the plasma of the pinch example, or the external conductors of other possible magnetic confinement methods), the power dissipated per unit volume of the conductor is equal to $\rho_0 j^2$ where $\rho_0$ is the resistivity. The magnetic field produced by any conductor system is proportional to $j$ times a linear dimension of the system. For simplicity consider a cylindrical system with a length great compared to its diameter. We may then write the scaling relationship for a magnetic field produced by the conductor system

$$H_0 \sim j R_c, \tag{85}$$

where $R_c$ is the radius of the conductor system. Thus

$$H_0^2 \sim j^2 R_c^2. \tag{86}$$

Now the total joule heat dissipated per unit length in the system is proportional to the unit volume dissipation rate $\rho_0 j^2$ times the volume per unit length of the system, *i.e.*,

$$p_j \sim \rho_0 j^2 R_c^2 \sim \rho_0 H_0^2,$$

independent of the diameter of the system. Thus for the ratio of $p/p_j$ there is obtained:

$$\frac{p}{p_j} \sim \left( \frac{\langle \sigma v \rangle_{Av}}{T^2} \right) \frac{1}{\rho_0} R^2 H_0^2. \tag{87}$$

Just as in the case of competition with the direct losses from particle diffusion, competition between fusion power output and joule heat losses in a long system is favored by increasing the radius or the magnetic field. It is interesting to note that in cases where volume joule heat dissipation, $\rho_0 j^2$, is the limiting factor in the design of a magnetic field coil, (85) shows that high magnetic fields are most easily obtained in large diameter systems.

*Other Scaling Factors*

To conclude discussion of scaling factors mention should be made of material and radiation problems. It has already been noted that heat transfer and other problems must eventually introduce limitations in the practical achievable power density of a fusion reactor. Attractive though it seems to dream of constructing a compact nuclear fusion engine about the size of an unabridged dictionary, limitations on power density introduced by the properties of materials would alone represent a serious barrier. In addition, all fusion fuels of present practical interest can undergo reactions which produce neutrons. Thus a fusion reactor producing any substantial amount of power would probably have to be surrounded by several feet of radiation shielding. This would make the dreamed-for miniature reactor somewhat less economically attractive even if it were technically feasible.

Another factor obviously important to the achievement of a favorable power balance is the choice of reacting fuel. Here the significant choices are as between the DD reaction and the DT reaction. The DT reaction exhibits a much larger reaction cross section than DD, as seen from Fig. 1. The total energy release per reaction is also substantially higher. However, among the various factors which must be taken into account in weighing the relative advantages of DT or DD as primary fuels, consideration should be given to the fact that most of the kinetic energy in the DT reaction appears in the neutrons. In the case of the DD reaction a large fraction of the reaction energy may eventually appear in the form of charged particles and might therefore be recoverable by direct conversion to electrical energy through interaction with confining fields.

## FRUITLESS APPROACHES

Controlled fusion research presents an unusual two-edged challenge to the physicist. While demanding great ingenuity in the formation of qualitative ideas to solve problems of heating and confinement, lack of diligent scrutiny of the *quantitative* aspects of the problem can render the most ingenious of plans completely worthless. By way of illustration of this, some simple examples of what appear to be fruitless approaches to controlled fusion will be discussed. It is often risky to attempt to prove general theorems about the impossibility of performing certain experiments. Nevertheless, confronted with the first and second laws of thermodynamics, few physicists today engage in a search for perpetual motion machines. (It is hoped that the construction of a practical fusion reactor does not fall into a similar category of endeavor!)

*Earnshaw's Theorem—"Electric Containment"*

The fuel of a fusion reactor is made up of charged particles and charged particles are influenced by electrostatic fields. Thus someone might suggest that a fusion reactor be constructed from a set of charged electrodes, forming an electrostatic "cage" for the plasma, which could thus be held free from contact with the material electrodes by the action of electrostatic fields.

Such a proposal fails on at least three counts, two qualitative and one quantitative. Firstly, Earnshaw's theorem from classical electrostatics shows that no position of stable equilibrium for even a single charged particle can be formed by the field of arbitrarily disposed charged conductors. Secondly, any electrostatic field which would represent a potential well for one sign of charge would represent a "hill" for the other sign. The third objection has to do with magnitude of the "pressure" which can be exerted by an electric field. Just as in the case of the magnetic field this pressure is limited by the energy density of the field, which for electric fields is $(E^2/8\pi)$ (stat volts/cm)$^2$ or ergs/cm$^3$. Choosing the former example of a plasma density of $6 \times 10^{15}$ particles per cm$^3$ and a temperature of 100 kev, one discovers that $E$ turns out to be $1.6 \times 10^5$ stat volts/cm or $4.8 \times 10^7$ volts/cm! This is an enormous field, and since the "electric pressure" varies as the square of the field there is not much hope in finding an operating point of the slightest practical interest, even if the previous objections did not apply.

*Target Bombardment*

It has already been mentioned that, although nuclear fusion reactions can be produced copiously in the laboratory by merely bombarding say, a deuterated target with a beam of high-energy deuterons, this does not offer a promising avenue to fusion power because of the low yield of reactions per incident deuteron. Most of the energy of the incoming particle is expended uselessly in ionization of the target atoms, which promptly radiate the lost energy by recapture of the electrons and by other similar processes. A frequently suggested improvement on this idea is to use the plasma of an ordinary gas discharge as the bombarded target, thereby eliminating the problem of ionization losses. Here the idea is qualitatively good, but fails quantitatively. Firstly, most ordinary gas discharge plasmas are not completely ionized and ionization and other losses still could predominate over reactions. Secondly, the electron temperature of most such discharges is of order 0.01 kev, usually established by processes occurring near the boundaries of the discharge. The expression (50) then shows that the mean time for energy loss from an incoming "hot" ion by collisions with the "cold" electrons would be about $10^9/n$ seconds. Thus if $n = 10^{15}$, for example the time for the ion energy to be substantially reduced is only $10^{-6}$ second, which is very short compared to the mean reaction time at these densities. The idea might have more merit if the electron temperature were substantially increased—however, this then becomes essentially the same problem as that of containing a hot self-reacting plasma, which if solved will also do the job.

## Colliding Beams

A variation on the theme of bombarding a target with an ion beam is the suggestion that reactions be produced by directing two ion beams at each other, obtaining reactions from the mutual collisions. Such a suggestion has the merit that the problem of getting high enough ion energies is apparently solved and, by confining the ion in a beam, collisions with walls are avoided, at least for a time. This scheme perishes from the numbers. A very intense ion beam might have a current density approaching 0.1 ampere per cm². The density of ions in such a beam is calculated by dividing the charges per second per cm² by the velocity. For 100-kev deuterons this yields a particle density of $2 \times 10^9$ ions/cm³. Application of (10) gives a reaction power density of $4 \times 10^{-11}$ watt/cm³!

Many other examples similar to those above could be cited. All illustrate the unusual importance of combining quantitative feasibility with qualitative operability as criteria in judging proposals for achieving controlled fusion reactions.

### PLASMA DIAGNOSTICS

One of the tantalizing things for the experimental physicist attempting to undertake an orderly study of phenomena going on in the midst of a hot plasma is the difficulty of performing measurements. If he were studying, say, the solid state properties of quartz, he would first order some quartz from a chemical supplier. He would then devise certain experimental tests and bring to bear in the course of the experiments a number of well-understood techniques, such as optical and electrical measurements or even acoustical methods. Moving the piece of quartz from one experimental setup to another, he could assemble data from which he could obtain a precise picture of the properties of the material quartz.

Contrast this situation with that of the physicist engaged in controlled fusion research. Firstly, he will find it rather difficult to order a liter or so of hot plasma from his supplier. The material he is to study must be manufactured in the course of the experiment. Secondly, he will find that the nearer he comes to success the fewer will be the number of feasible experimental techniques which may be applied to the problem. The best of these —the detection of neutrons from the fusion process— will properly only occur well along on the trail, and may even be misleading, since neutron producing fusion reactions can arise from entirely spurious processes. These considerations greatly intensify the need for sharpening of the observational and critical senses of the experimental physicist. Space does not permit an extensive discussion of the physics of plasma diagnostic measurements. Some example methods can be described however, which will illustrate the attacks involved. Types of measurements discussed are: 1) optical or spectroscopic, 2) electromagnetic interactions, and 3) nuclear reaction measurements.

## Spectroscopic Measurements

The controlled fusion problem is perhaps most nearly a direct descendant from the science of astrophysics. It is therefore natural that astrophysical measurement techniques be applied to study controlled fusion reactions. Until the advent of radio-astronomy, optical measurements were the only means of obtaining astrophysical data. However, in a successful fusion reactor the plasma must be essentially totally ionized, so that light from atomic processes is totally negligible. It would appear that spectroscopic measurements will be of minor use in the later stages of controlled fusion research. However, in the earlier stages of such research, where temperatures are still not up to the fusion point, and total ionization has not yet been achieved, spectroscopic measurements can give much useful information. An idea of the purity and composition of the plasma can be obtained by studying emission spectra. By studying the broadening of individual spectral lines, information on the plasma ion density and its temperature can be obtained. In a plasma the passage of ions near to each other results in localized electric fields which in turn produce a fluctuating Stark broadening[30] and the amount of this broadening is related to the plasma density. Also, motion of an emitting ion relative to the observer will give rise to a Doppler shift of the emitted spectral line. If light is being received from a number of such ions and their direction of motion is random, an additional broadening of the line will result. This may be distinguished from that owing to Stark broadening and may be used to determine the ion temperature. Ordered or gross turbulent motions of the plasma may also produce Doppler shifts. Thus care must be used in interpreting this kind of spectroscopic data, especially since emitting atoms may comprise a nonrepresentative sample of the plasma.

## Electromagnetic Measurements—Microwaves

Measurement of electromagnetic effects produced by the plasma is probably the most important class of plasma diagnostic techniques. Some of these are derived from the older field of gas discharge physics, to which fusion research owes a considerable debt. Especially in recent years electromagnetic measurements of gas discharge phenomena have reached a point of considerable refinement. The brilliant experimental and theoretical work of Allis and Brown and their collaborators at Massachusetts Institute of Technology in using microwave techniques to analyze discharges has an application in fusion research. In the M.I.T. work plasma properties are usually deduced from the detuning of cavity resonators containing the discharge plasma. Such a detuning effect is to be understood most simply in terms of the simple dielectric constant of a plasma defined in (25)

[30] J. Holtsmark, "Broadening of spectral lines—part II," *Physik. Z.*, vol. 25, pp. 73–84; February, 1924.

$$K = 1 - (4\pi n e^2/m\omega^2) = 1 - (f_p{}^2/f^2)$$
$$= 1 - 8.1 \times 10^7 (n/f^2). \qquad (88)$$

Since the M.I.T. group found it convenient to use cavities operating at a wavelength of 10 cm ($f = 3000$ mc) in their work, (88) shows that most of their measurements had to be restricted to electron densities below about $10^{11}$ particles per cm³. Larger size experimental chambers would have required even lower operating densities.

Though it can be seen that because of size and density limitations microwave cavity techniques cannot often be used in controlled fusion research, a related microwave technique can be employed. By using millimeter wave techniques the density at which $K$ becomes negative (and thus $\sqrt{K}$ becomes imaginary) can be increased considerably. Then by constructing a microwave interferometer in which an extended path through the plasma forms one leg of the interferometer, the electron density can be measured and plotted as a function of time or other variables. The state of the millimeter microwave art still does not permit covering the whole spectrum of interesting plasma densities by this method by much useful information can nevertheless be obtained.

### Electromagnetic Measurements—Induction Effects and Probes

It has already been noted that a plasma possesses properties which might be described as "diamagnetic" and that its high electrical conductivity may often result in a "freezing-in" of lines of magnetic force within the plasma. For this reason the presence of a confined plasma and its internal electrical currents can profoundly alter the nature of magnetic fields existing near the plasma. Motion of the plasma can then induce voltages in external circuits which can in turn be used to infer something about the motion of the plasma body and possibly its density and temperature. For example, in the pinch effect of a preceding section, if the pinch becomes smaller or larger as a function of time, this will be evidenced by an effect on the external circuit.

Consider a linear pinch discharge as in Fig. 6. The voltage which will be measured from one end to the other will be the sum of resistive and inductive effects.

$$V = RI + \frac{d}{dt}(LI). \qquad (89)$$

Let us suppose that resistive effects are small. Then

$$V = \frac{d}{dt}(LI) = L\frac{dI}{dt} + I\frac{dL}{dt}. \qquad (90)$$

$V$, $I$, and $dI/dt$ can be measured externally. There remains an equation which can be solved for $L$ as a function of time. But in this case the geometry is known, so that $L$ is a known function of the radius of the pinched

discharge (for current flowing on the surface, as in the example). Thus from externally measured voltages and currents the dimensions of a transient pinched discharge can be determined as a function of time. More detailed information can be obtained about the pinch from pickup loops appropriately placed in the field of the pinch currents.

Information about electrostatic fields surrounding the plasma can sometimes be gleaned by the use of electric probes, as in classical gas discharge physics. Here however the situation is not as clean, and the need for minimizing physical contact between the confined plasma and the probe make this technique of limited applicability.

### Nuclear Measurements

Measurement of the neutrons resulting from fusion reactions in a plasma would certainly be a satisfying way to diagnose the condition of the plasma. Under the proper circumstances such measurements could give information on the plasma temperature, its absolute density, and its spatial distribution. If the plasma is composed of deuterium and is at a temperature below about 10 kilovolts, the temperature dependence of the reaction rate would make neutron yield a useful thermometer. A clear-cut variation of the neutron yield with the square of the plasma density could strongly suggest that the neutrons were of thermonuclear origin, whereas a linear dependence on density would probably indicate that neutrons were merely being produced by target bombardment. This and other pitfalls await careless interpretation of neutron production in controlled fusion experiments. For example, in some cases, motional or other electric fields could produce acceleration of ions in such a way as to give rise to a neutron yield which is not of thermonuclear origin.

The general problem of plasma diagnostics is one of the most important and challenging aspects of controlled fusion research. Unfortunately it is most difficult in the early stages of the research and will become easier only when the end is in sight.

### CONCLUSION

This article is an attempt to present some of the physical background and practical problems of a research field which is yet in its infancy. The intent has not been to impress readers with the difficulty and complexity of controlled fusion research but rather to assemble in one place some of the important pieces of a jig-saw puzzle yet to be put together. The challenge of the field to the scientific mind should be tremendous, and the goal to be won is of the highest worth. It is the firm belief of many of the physicists actively engaged in controlled fusion research in this country that all of the scientific and technological problems of controlled fusion will be mastered—perhaps in the next few years. Several different and novel approaches to the problem, involv-

ing both transient and steady-state methods, are being studied in this country. No doubt similarly promising research is being carried out in other laboratories throughout the world.

Out of the pursuit of this problem a new and fertile field of experimental and theoretical physics is arising. From a thorough understanding of the physics of ultra-high temperature plasmas and their interaction with electromagnetic fields one can hope not only for the achievement of controlled fusion power, but also, as a result of this increased knowledge of nature, there will no doubt arise new and important applications to other fields of science and technology.

That an early success in achieving a self-sustaining controlled fusion reaction would lead to economically competitive fusion power in the near future is highly unlikely. Still, in the fusion reaction are implicit new dimensions—those of power obtained, possibly by direct electrical conversion, from an inexpensive, safe, and virtually inexhaustible fuel. These possibilities will surely someday play a dominant role in shaping the world of the future.

# CORRECTION

The following correction to "Methods for Measuring Piezoelectric, Elastic, and Dielectric Coefficients of Crystals and Ceramics," by W. P. Mason and H. Jaffe, which appeared in the June, 1954 issue of PROCEEDINGS, has been called to the attention of the authors by P. J. C. Lewington.

On page 929, (15) in column one should be

$$k_r^2 = \frac{\Delta f}{f_r}\left[\frac{R_1^2 - (1 - \sigma^2)}{1 + \sigma}\right]\left[1 - \frac{\Delta f}{f_R}\left(\frac{R_1 - 1 - \sigma^2}{1 + \sigma}\right)\right].$$

# The Measurement and Specification of Nonlinear Amplitude Response Characteristics in Television*

STEPHEN DOBA, JR.†, SENIOR MEMBER, IRE

*Summary*—With the advent of color television the need for extreme linearity of over-all systems has become apparent. The general problem of the measurement and specification of nonlinear characteristics in television is considered. A function, called the gradient gain of a system, is presented. At a point in the picture for which the gradient gain is positive there is an enhancement (expansion) of the tone rendition. Similarly there is a reduction (compression) in tone rendition when the gradient gain is negative. The gradient gain is a function of the characteristics of the camera tube, picture tube, and purely electronic devices. These characteristics have been defined and methods of measurements suggested in a way which follows most simply from the physical and physiological factors involved. The characteristics of electronic amplifiers are defined so that the extension to color television is readily achieved. The terms "Differential Gain" and "Differential Phase," which have been adopted by the IRE, are defined so that they are applicable to electronic devices, such as amplifiers, used for either monochrome or color television.

## INTRODUCTION

WITH THE advent of color television the need for extreme linearity of over-all systems has become apparent. This necessity follows directly from the fact that in a color system nonlinearities produce chromaticity errors which are more detrimental than tone value errors.

We propose to consider here the general problem of the measurement and specification of nonlinear characteristics in television systems. The devices to be considered include camera tubes, reproducing cathode-ray tubes as well as electrical amplifiers, radio transmitters, in short, all the component elements in an over-all system.

This listing of the elements in a complex system brings to the fore the basic problem of how to measure the nonlinearity of each component so that the over-all performance may be obtained as the sum of the individual measurements. A subsidiary problem is one of defining terms and measurements in a way which follows most simply the physical and physiological factors involved.

## BRIGHTNESS VALUES

The fundamental starting point in a consideration of the factors influencing the measurement and specification of nonlinear characteristics in television is a recognition of the approximate validity of the Weber-Fechner law of sensation: over a wide range of luminance values human vision recognizes a luminance change which is a fixed percentage of the luminance. As a practical detail this means that in the optical portions of the system the changes in log luminance ($\Delta \log B = \Delta B/B$) should be taken as the basic quantity.

Fortunately this quantity can be measured by the use of neutral density filters. At a camera pickup device the use of such a filter interposed between the device and the light source produces a known value of $\Delta B/B$ for any $B$ once the density of the filter is known. At a reproducing device it is possible to cause a match in $\Delta B/B$ produced by either the filter or by a suitable signal. This method is described by M. W. Baldwin.[1]

## GRADIENT

The approach to the problem can be stated as follows: If we are given a percentage change ($\Delta I/I$) in scene luminance, the quantity that is important in the reproduced picture is again the change in log luminance, ($\Delta B/B$). For small changes this can be expressed as the relationship

$$\frac{\Delta B}{B} = g \, \frac{\Delta I}{I} \tag{1}$$

where $g$ is a function of $I$.

In the over-all system the problem is one of defining and measuring the quantity $g$ in a meaningful manner. B. M. Oliver has called the $g$ the "gradient."[2] Others have suggested the use of "gamma." It is not intended to enter into a discussion of this definition, but merely to point out that in a complex system the quantity $g$ is determined by a large number of devices.

If the reproduced luminance ($B$) is a function of the scene luminance ($I$), then for small changes we have the relationship

$$\Delta B = \Delta I \, \frac{dB}{dI} \cdot \tag{2}$$

At any suitable point in the electrical system we will have a voltage $e$ as a function of $I$, and similarly $B$ will

[1] M. W. Baldwin, "Measurement method for picture tubes," *Electronics*, pp. 104–105; November, 1949.
[2] B. M. Oliver, "Tone rendition in television," PROC. IRE, vol. 38, pp. 1288–1305; November, 1950.

be a function of $e$.[3] Hence (2) may be rewritten as

$$\Delta B = \Delta I \frac{dB}{de} \frac{de}{dI}. \tag{3}$$

If we divide both sides of this equation by $B$ and multiply and divide the right-hand side by $I$, we have

$$\frac{\Delta B}{B} = \frac{dB}{Bde} \frac{Ide}{dI} \frac{\Delta I}{I}. \tag{4}$$

Eq. (4), it may be seen, bears a resemblance to (1) in that $\Delta B/B$ is related to $\Delta I/I$. The first factor on the right ($dB/Bde$) has been called by Baldwin the "sensitivity" of the receiving tube and is a quantity which he has demonstrated can be measured readily.

Its use, however, has one drawback in that the sensitivity is not normalized. A useful quantity should be normalized and dimensionless in a linear system. We propose to remedy this by introducing the maximum value $E$ of the signal corresponding to the maximum luminance in the picture. Eq. (4) thus becomes

$$\frac{\Delta B}{B} = \frac{EdB}{Bde} \frac{Ide}{EdI} \frac{\Delta I}{I}. \tag{5}$$

As a final step let us introduce an amplifier which may be nonlinear. The camera output voltage $e_1$ is the amplifier input. The amplifier output voltage $e_2$ is the receiving picture tube input. Eq. (5) is then transformed to

$$\frac{\Delta B}{B} = \frac{E_2 dB}{Bde_2} \frac{E_1 de_2}{E_2 de_1} \frac{Ide_1}{E_1 dI} \frac{\Delta I}{I}. \tag{6}$$

The form of (6) strongly suggests the input-output relationship of a system exemplified by a number of amplifiers in tandem

$$E_{out} = A_1 A_2 A_3 \cdots A_n E_{IN} \tag{7}$$

where $A_1$, $A_2$, $A_3$, $\cdots$, $A_n$ represent the amplification factors of the various amplifiers. Where a number of devices such as these are connected in tandem, it is customary to deal with logarithmic units (such as db), so that one need use only addition of effects rather than multiplication.

## Gradient Gain

By obvious extension we may use rather than a numeric $g$, a logarithmic term derived from it. We suggest for this the name "gradient gain." In terms of (6) this would be expressed by

$$G = 20 \log_{10} \left( \frac{\Delta B/B}{\Delta I/I} \right) = 20 \log_{10} g$$

$$= 20 \log_{10} \frac{E_2 dB}{Bde_2} + 20 \log_{10} \frac{E_1 de_2}{E_2 de_1}$$

$$+ 20 \log_{10} \frac{Ide_1}{E_1 dI} \tag{8}$$

---

[3] This discussion assumes either dc coupling throughout or that the dc relationships between $e$, $B$, and $I$ are maintained by some means as clampers.

where $G$ (like $g$) is a function of $I$.

It may be well at this point to examine (8) term by term to reveal its meaning.

1) $\Delta B/B$ is the desired expression for the percentage change in luminance of the reproduced picture. Similarly $\Delta I/I$ is the desired expression for the percentage change in luminance of the original scene.

2) When $\Delta B/B$ is greater than $\Delta I/I$, there is an enhancement of tone rendition. In this the gradient gain is positive. When $\Delta B/B$ is less than $\Delta I/I$ there is a reduction in tone rendition and the gradient gain is negative.

3) The quantity $E_2 dB/Bde_2$ is a function of the receiving tube and the maximum signal at the receiving tube. This quantity may be measured (following Baldwin) by changing the picture tube signal a small amount $de_2$ while simultaneously interposing a neutral density filter in the field of an illuminometer. $de_2$ is then varied until there is no luminance change as indicated by the illuminometer.

It is proposed to call $20 \log_{10} E_2 dB/Bde_2$ the differential sensitivity of the picture tube, and the method of its measurement as described above. As an example of its use, consider the characteristic of a picture tube defined by the power-law relationship

$$B = B_0 \left( \frac{e_2}{E_2} \right)^n \tag{9}$$

where $B_0$ is the maximum luminance, $e_2$ the signal volts measured from cutoff, and $E_2$ the maximum signal volts measured from cutoff. Fig. 1 shows luminance characteristics for several values of $n$. Picture tube characteristics usually fall between the curves of $n = 2$ and $n = 3$. The differential sensitivity is then given by

$$20 \log_{10} \left( \frac{E_2 dB}{Bde_2} \right) = 20 \log_{10} n - \frac{1}{n} 20 \log_{10} \left( \frac{B}{B_0} \right). \tag{10}$$

Hence, as shown on Fig. 2, the slope of the plot of differential sensitivity vs log brightness is $-(1/n)$. This is, of course, similar to the results given by Baldwin, and indicates that picture tubes may be specified just as readily by the differential sensitivity as by the luminance-voltage characteristic.

4) The quantity $Ide_1/E_1 dI$ is similarly a function of the camera tube and the maximum output signal. The reciprocal of this quantity, namely $E_1 dI/Ide_1$, becomes identical in form with that for the differential sensitivity of the picture tube. Apart from an interchange between input and output functions it may be seen that the two expressions are similar. We propose therefore that the quantity $-20 \log_{10} E_1 dI/Ide_1$ be called the differential sensitivity of the camera pickup tube.

This quantity may be measured by interposing a neutral filter of known density between the light source and the camera tube and measuring the corresponding change in output voltage.

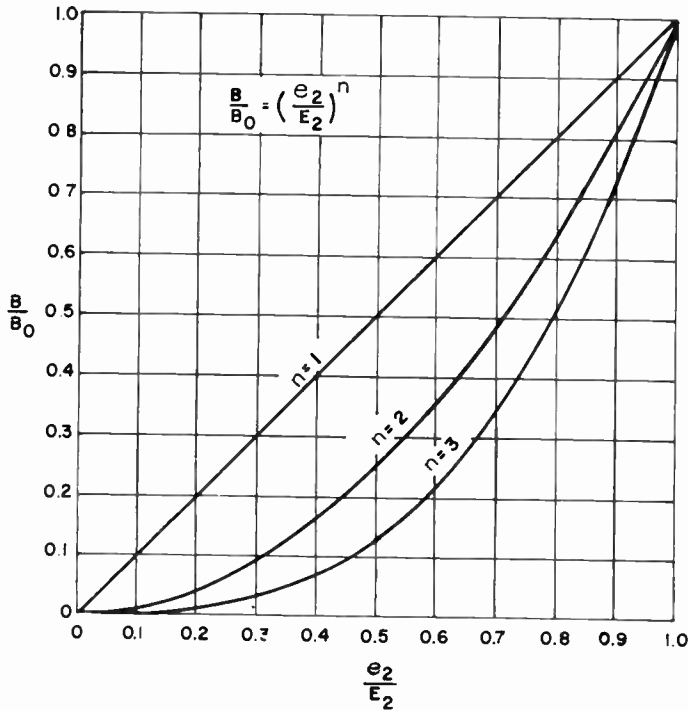As an example of the use of this quantity consider the
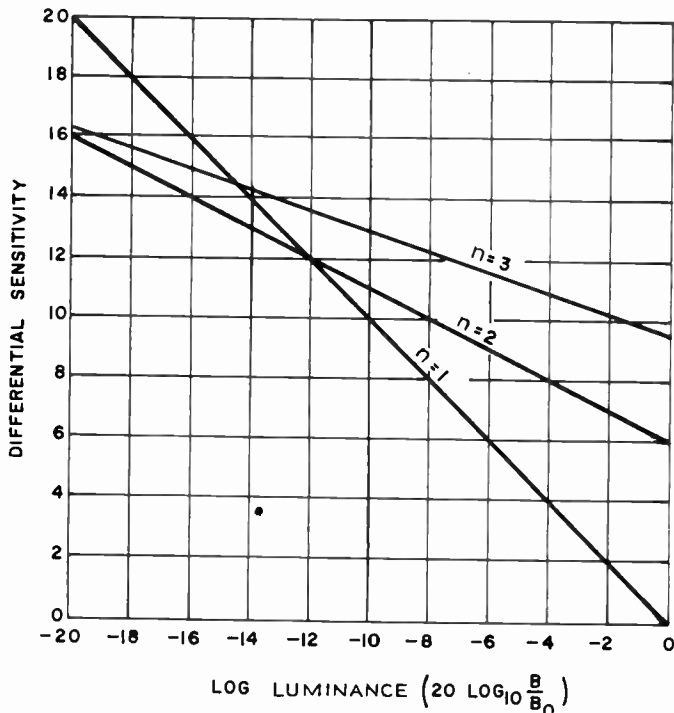
Fig. 1—Luminance characteristics.



Fig. 2—Picture tube differential sensitivity vs log luminance.

characteristic of a camera tube defined by a power-law relationship:

$$e_1 = E_1 \left( \frac{I}{I_0} \right)^m \tag{11}$$

where $I_0$ is the maximum scene luminance, $e_1$ is the change in output signal volts, and $E_1$ is the maximum change in output signal volts. We then have

$$\frac{de_1}{dI} = mE_1 \frac{I^{m-1}}{I_{0_m}} \tag{12}$$

and the differential sensitivity is then given by

$$-20 \log_{10} \frac{E_1 dI}{I de_1} = 20 \log_{10} m + m \, 20 \log_{10} \frac{I}{I_0} \cdot \tag{13}$$

The plot of differential sensitivity vs log luminance is again a straight line and has a slope of $m$.

5) The quantity $20 \log_{10} E_1 de_2 / E_2 de_1$ is a function of the amplifier and the maximum signal. It may be seen to be the difference between the small-signal gain $20 \log_{10} de_2 / de_1$ and the maximum-signal gain $20 \log_{10} E_2 / E_1$. Another way of expressing the quantity is by means of the ratio of the slope $(de_2/de_1)$ at any point of the input-output curve, to the average slope $E_2/E_1$.

## DIFFERENTIAL GAIN

The measurement of this quantity can readily be achieved wlth the desired accuracy. One obvious way is by the superposition of a small-amplitude high-frequency wave upon a large saw-tooth, stepped wave, sine wave, or other low-frequency signal which sweeps over the whole signal range. The quantity desired is the difference in gains for the high-frequency signal between any two points in the low-frequency cycle. As used here one of these points is always the point for which the high-frequency gain is the average (with respect to the input voltage), but there are applications where the average is neither readily determinable nor necessary. This quantity has been called ,the differential gain of the amplifier. It should be emphasized that for nonlinear amplifiers the differential gain will be zero at only one value of input.

As an example of the use of this concept consider an amplifier which is supposedly linear but has an input-output characteristic given by the expression

$$\frac{e_2}{E_2} = \frac{\dfrac{e_1}{E_1} + a \left( \dfrac{e_1}{E_1} \right)^3}{1 + a} \cdot \tag{14}$$

The form of this was chosen to make obvious the relation that $e_2 = E_2$ when $e_1 = E_1$ and that $e_2 = 0$ when $e_1 = 0$. Fig. 3 shows the input-output characteristics for $a = 0$ (ideal case) and $a = 0.1$. The differential gain is then given by

$$20 \log_{10} \frac{E_1 de_2}{E_2 de_1} = 20 \log_{10} \frac{1 + 3a \left( \dfrac{e_1}{E_1} \right)^2}{1 + a} \cdot \tag{15}$$

From (15) and Fig. 4 we see that the differential gain varies from a maximum negative value when $e_1 = 0$ to a maximum positive value when $e_1 = E_1$. If we recall that in (8) the differential gain is added to the differential
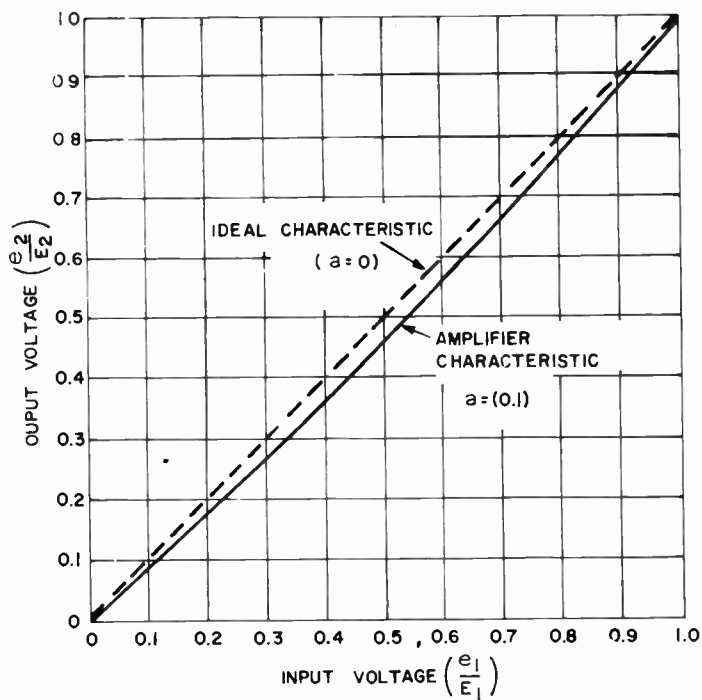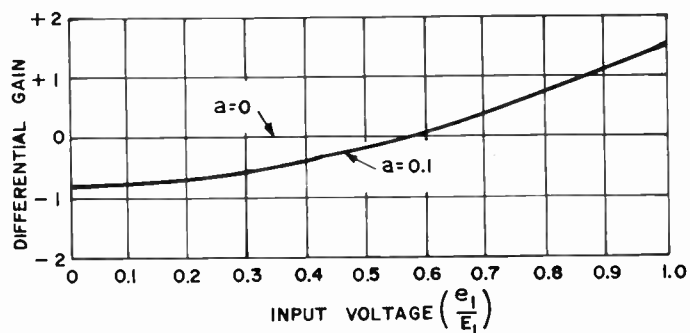
Fig. 3—Amplifier characteristics.



Fig. 4—Differential gain characteristic.

sensitivities it will be apparent that a differential gain less than zero implies relative compression or reduction in tone rendition, while a differential gain greater than zero implies relative expansion or enhancement in tone rendition.

As a further example of the use of this concept, consider an amplifier designed as a "gamma corrector" with a characteristic given by

$$\frac{e_2}{E_2} = \left(\frac{e_1}{E_1}\right)^\gamma. \tag{16}$$

The differential gain of this is given by

$$20 \log_{10} \frac{E_1 de_2}{E_2 de_1} = 20 \log_{10} \gamma \left(\frac{e_1}{E_1}\right)^{\gamma-1} = 20 \log_{10} \gamma \left(\frac{e_2/E_2}{e_1/E_1}\right)$$

$$= 20 \log_{10} \gamma + 20 \log_{10} \frac{e_2/E_2}{e_1/E_1}. \tag{17}$$

The last form shown in (17) yields a simple expression for the exponent $\gamma$. If we recognize

$$20 \log_{10} \left(\frac{e_2/E_2}{e_1/E_1}\right)$$

as the relative signal gain, then $20 \log_{10} \gamma$ is given by the difference between the differential gain and the relative signal gain. The advantage of this method of representation rather than merely the log-log plot of (16) lies in the fact that the latter is quite sensitive to errors in the value of zero assigned to $e_1$ for small values of $e_1$. Eq. (17) is much less sensitive in this respect and hence its use should yield greater accuracy in measurement. It should be stressed that in all these considerations an attempt has been made to measure voltages as changes only in voltage because of the difficulty of assigning an accurate zero in most cases. Similarly, in measuring luminance, critical measurements should be made on percentage changes only because these are the ones most readily made. As an extension of this philosophy the actual luminance-voltage characteristic need not be measured or used.

## GRADIENT GAIN OF A SYSTEM

The gradient gain of a system can now be expressed in terms of these quantities as

$$G = S_R + S_C + G_1 + G_2 + \cdots + G_n \tag{18}$$

where $S_R$ is the differential sensitivity of the receiving tube, $S_C$ the differential sensitivity of the camera device, and $G_1, G_2, \cdots, G_n$ are the differential gains of the various electrical devices involved. When $G$ is zero for all values of input, the system is distortionless. For instance, this may be achieved when the camera and picture tubes have opposite differential sensitivity characteristics and all the amplifiers have zero differential gain over their operating range. Fig. 5 shows an example of the individual and combined differential effects of a camera, a picture tube, and a gamma correcting amplifier on the gradient gain of a system.

## SIGNAL-TO-NOISE RATIO

An example of the further usefulness of these concepts is offered in the specification of signal-to-noise ratios. Consider a point in the electrical circuit where an interfering or extraneous signal $\Delta e$ will produce a variation $\Delta B$ in luminance in accordance with the relationship

$$\Delta B = \Delta e \frac{dB}{de}. \tag{19}$$

In introducing the maximum signal $E$ it is possible by the use of (19) to express the signal-to-noise ratio by

$$\frac{E}{\Delta e} = \frac{B}{\Delta B} \left(\frac{E dB}{B de}\right). \tag{20}$$

If we introduce an amplifier between the noise source and the picture tube, with the input voltage designated $e_1$ and the output voltage $e_2$, (20) is replaced by
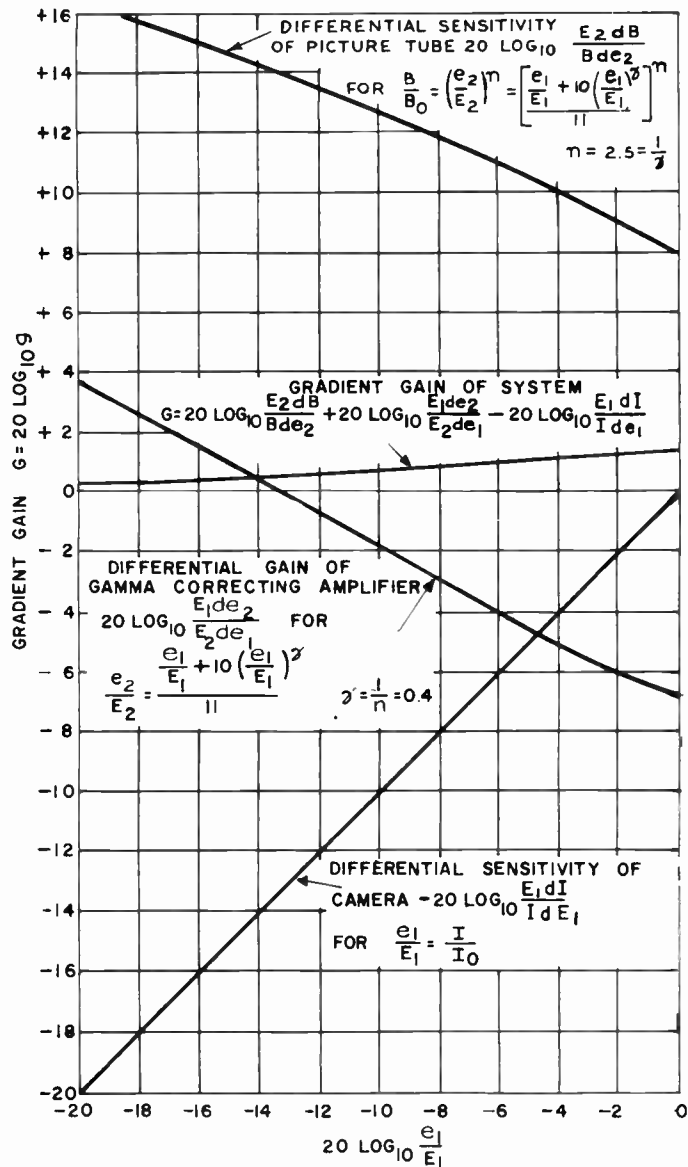
Fig. 5—Gradient gain of a system.

$$\frac{E_1}{\Delta e_1} = \left(\frac{B}{\Delta B}\right)\left(\frac{E_1 de_2}{E_2 de_1}\right)\left(\frac{E_2 dB}{B de_2}\right). \tag{21}$$

In terms of our definitions above, this is seen to be expressible as

$$20 \log_{10} E_1/\Delta e_1 = 20 \log_{10} \frac{B}{\Delta B} + G_a + S_R. \tag{22}$$

These factors, of course, must all be evaluated at the same point. $G_a$ is the differential gain of the amplifier, and $S_R$ is the differential sensitivity of the receiving tube. Eq. (22) hence expresses the signal-to-noise ratio in db as the sum of three terms, the first of which can be determined from suitable visual data, while the others are obtainable from the devices involved.

An extension of these concepts is also felt to be necessary because of the development of the NTSC Color

Television System. Since in this system the television signal is composed of a luminance signal plus a 3.58-mc carrier modulated in amplitude and phase by the color information, cross modulation between these components is possible. It may be seen that this signal is quite similar to the one proposed for the measurement of differential gain. The variation in gain to the high-frequency signal measures now not only the compression of the low-frequency (luminance) signal, but more importantly measures the variation in amplitude of the color signal due to nonlinearities of the electrical system. Hence the concept of differential gain can be adopted directly. It should be emphasized that this variation in gain is due to the modulation of the amplitude of the chrominance signal by the luminance signal.

## DIFFERENTIAL PHASE

Since a similar modulation effect can obtain for the phase of the chrominance signal, an obvious extension is to include the phase of the signal as well. This effect has been designated by the expression differential phase.

It is felt that this extension of the concept is a useful and necessary one. By analogy with the complex gain of amplifiers, the differential gain of amplifiers may be also considered complex. This however, restricts the method of measurement and definition to one which permits of complex measurements.

## DEFINITIONS

The definitions which have been adopted by the IRE are:[4]

### Differential Gain

In a Video Transmission System "differential gain" is the difference in the gain of the system in decibels for a small high-frequency sine-wave signal at two stated levels of a low-frequency signal on which it is superimposed.

### Differential Phase

In a Video Transmission System "differential phase" is the difference in phase shift through the system for a small high-frequency sine-wave signal at two stated levels of a low-frequency signal on which it is superimposed.

## ACKNOWLEDGMENT

It is with a great deal of pleasure that acknowledgment is made of the contributions of H. P. Kelly toward the preparation of this paper. These extended from the composition of figures to critical comments and suggestions including the use of the term "differential phase." Acknowledgment is also made of the careful review and comments by M. W. Baldwin, Jr. and J. R. Hefele.

[4] In these definitions level means a specified position on an amplitude scale applied to a signal waveform.

# Transient Analysis of Coaxial Cables Considering Skin Effect*

R. L. WIGINGTON† AND N. S. NAHMAN‡, ASSOCIATE MEMBER, IRE

*Summary*—A transient analysis of coaxial cables is made by considering the skin effect of the center conductor as the distorting element. Generalized curves are presented by which the response of any length of coaxial cable can be predicted if one point on the attenuation vs frequency curve is known. An experimental check on the analysis is made by comparing measurements and prediction of the responses of several different coaxial cables.

## INTRODUCTION

IN A STUDY of oscilloscope systems for use in observing voltage waveforms of the duration of a few millimicroseconds (1 m$\mu$s = $10^{-9}$ sec), the problem of the distortion of waveforms by the high frequency loss of coaxial cable was encountered. Elementary consideration of the problem indicated a degradation of fast rise times (1 m$\mu$s or less) due to greater attenuation of the high-frequency components of the signal.

In polyethylene dielectric coaxial cables, the conductance loss is extremely small. Polyethylene has a dissipation factor of 0.0031 at 3000 mc[1] and less at lower frequencies. Likewise, in air dielectric cables the conductance loss is even less. Therefore, the major portion of high-frequency loss could not be blamed on leakage conductance. The other source of loss in coaxial cable is the series resistance of the center conductor. For analysis the skin effect of the outer conductor was considered to be lumped with the skin effect of the center conductor increasing it slightly. Using empirical data to evaluate the skin effect constant achieves this directly. Ordinary analysis of transmission lines ignore this resistance as being negligible. However, at frequencies at which the skin effect of conductors becomes significant, the analysis must include its effects, both as series resistance and inductance.

In this analysis, a transmission line is treated as a four-pole network. With the aid of an approximation which is good at high frequencies, an analysis including skin effect and neglecting dielectric effects can be made. All calculations are in mks units.

## POSSIBLE APPLICATIONS

Before proceeding with the analytical details of the problem, a few words about the engineering applications would be indicative of the role which skin effect distortion in coaxial cables may play in contemplated and future systems using fast transients.

The origination of this problem was in the design of an oscilloscope system for observing very fast rise times, 1 m$\mu$s or less. In triggered oscilloscope systems a signal delay path (usually a simulated line or a coaxial cable) is necessary to allow time for the trigger circuits to detect the pulse to be observed and to start the sweep. The delay of this path is 50 m$\mu$s or longer in present systems. As shown in this paper, the distortion in this amount of coaxial cable is very serious for millimicrosecond transients. Therefore, along with the other limitations of oscilloscope systems (such as rise time of the signal amplifiers, writing speed, and vertical sensibility), the distortion due to the signal delay cable must be considered. Perhaps a knowledge of the form of this distortion will enable the extension of the range of oscilloscope systems which are limited by the signal delay distortion.

If preserving the rise times in fast pulse circuits is in any way critical to the proper operation of the circuitry, one must begin to consider the skin effect distortion in 10-mc prf circuits for long cable runs, and in 100-mc prf circuits, the distortion would be troublesome even in short cable lengths. The practice of using special small size coaxial cable to conserve space results in greater attenuation per unit length than for larger cable of the same characteristic impedance, and thus, also makes the skin effect distortion greater.

Another example of a problem in which the analysis may be very useful is in the analysis of regenerative pulse generators, a circuit which is essentially a loop consisting of an amplifier and a delay circuit.[2] For practical, high rep-rate pulse generation, the delay circuit is usually a coaxial cable. The pulse shape obtained is a composite of the characteristics of the cable and of the amplifier.

In short, for any electronic circuit application using coaxial cables as transmission media to provide either time delay or transmission of millimicrosecond pulses, the effects of skin effect distortion must be considered.

## ANALYSIS

For a transmission line of length, $l$, terminated in its characteristic impedance, $Z_0$, and with propagation constant, $\gamma$, the following relation exists between input ($E_1$) and the output ($E_2$) voltages as functions of complex frequency:[3]

---

[1] "Reference Data for Radio Engineers," Federal Telephone and Radio Corp., 3rd ed., p. 51.

[2] C. C. Cutler, "The regenerative pulse generator," PROC. IRE, vol. 43, pp. 140–148; February, 1955.
[3] The complex variable is the Laplace Transform variable $p$. Eqs. (1) and (2) comprise the Laplace Transform equations of the system differential equations.

$$E_2 = e^{-\gamma l}E_1 \tag{1}$$

where in general

$$\gamma = \sqrt{(R + pL)(G + pC)} \tag{2a}$$

$$Z_0 = \sqrt{\frac{R + pL}{G + pC}}. \tag{2b}$$

For high frequencies (skin depth small with respect to conductor radius), the skin effect impedance of a round wire is:[4]

$$Z_s = K\sqrt{p} \tag{3a}$$

and

$$K = \frac{1}{2\pi r}\sqrt{\frac{\mu}{\sigma}} \tag{3b}$$

where $r$ is conductor radius, $\mu$ is the permeability and $\sigma$ is the conductivity of the wire.

At high frequencies the series resistance of a wire is expressed by the skin effect equation. Since an increase in inductance is also caused by skin effect, it is treated as an impedance rather than as a resistance. Therefore, replacing $R$ in (2) by $Z_s$ and neglecting dielectric leakage ($G = 0$), (2) becomes

$$\gamma = \sqrt{(K\sqrt{p} + pL)pC} \tag{4a}$$

$$Z_0 = \sqrt{\frac{K\sqrt{p} + pL}{pC}}. \tag{4b}$$

The transfer function of a length of line is then:

$$\frac{E_2}{E_1} = e^{-\gamma l} = e^{-l\sqrt{p^2 LC + pCK\sqrt{p}}}. \tag{5}$$

The inverse Laplace Transform of the transfer function (5) is the impulse response of the section of line. For simplification, the following approximation was made. Expanding the square root in the exponent of (5) by the binomial expansion, one obtains

$$\gamma(p) = (p^2 LC + p^{3/2}CK)^{1/2}$$

$$= p\sqrt{LC} + \frac{Kp^{1/2}}{2}\sqrt{\frac{C}{L}} + \frac{1}{2}\sum_{n=2}^{\infty}(-1)^{n-1}$$

$$\cdot\left(\frac{1\cdot 3\cdots(2n-3)}{2^{n-1}n!}\right)\frac{K^n}{L^{n-1}}\sqrt{\frac{C}{L}}\,p^{1-n/2}. \tag{6}$$

The first term of (6) is the delay term and the remaining terms describe the waveform distortion. The series is an alternating convergent series (for $p^2 LC > p^{3/2}CK$). Approximating it by the second term of (6), the $p^{1/2}$ term, results in an error less than the next term, the $p^0$ term. The ratio of these two terms will be used as a measure of validity of applying this approximation to specific examples.

S. Ramo and J. R. Whinnery, "Fields and Waves in Modern Radio," John Wiley and Sons, Inc., New York, N. Y.; 1944.

$$A \equiv \left|\frac{p^0 \text{ term}}{p^{1/2} \text{ term}}\right| = \left|\frac{\frac{K^2}{8L}\sqrt{\frac{C}{L}}}{\frac{Kp^{1/2}}{2}\sqrt{\frac{C}{L}}}\right| = \left|\frac{K}{4Lp^{1/2}}\right|$$

$$= \frac{K}{4L\sqrt{2\pi f}}. \tag{7}$$

Using the first two terms of (6) in (5) and letting $R_0 = \sqrt{L/C}$, $T = \sqrt{LC}$, results in

$$\frac{E_2}{E_1} = e^{-l(pT + (K/2R_0)p^{1/2})}. \tag{8}$$

The exp $(-ltp)$ is simply a delay term so that the inverse transform of (8) is the inverse transform of exp $(-lkp^{1/2}/2R_0)$ delayed an amount $lT$. The latter exponential is a common transform and is listed in ordinary Laplace Transform tables.[5] Its inverse giving the impulse response is:

$$g(t) = \alpha x^{-3/2}e^{-\beta/x} \qquad x \geqq 0$$
$$= 0 \qquad\qquad x < 0 \tag{9}$$

where

$$\alpha = \frac{lK}{4R_0\sqrt{\pi}}, \beta = \left(\frac{lK}{4R_0}\right)^2, \text{ and } x = t - Tl.$$

Of greater utility in studying the distortion of fast rise times by skin effect are the step response and the response to a linear rise. The step response can be obtained by finding the inverse transform of $1/p$ times the transfer function. As before, the transform $1/p$ exp $(-lkp^{1/2}/2R_0)$ is listed in tables.[5] Therefore in terms of $x$ and $\beta$ as defined above, the step response is:

$$h(t) = \text{cerf }\sqrt{\frac{\beta}{x}} \qquad x \geqq 0$$
$$= 0 \qquad\qquad x < 0. \tag{10}$$

cerf $(y)$ is the "complementary error function of $y$."

The linear rise referred to previously is defined specifically as the following, and it will be referred to as a ramp input.

$$F(t) = 0 \qquad t < 0$$
$$= t/a \qquad 0 \leqq t \leqq a$$
$$= 1 \qquad t > a.$$

The response to $F(t)$, called $f(t)$, is given by the convolution of $F(t)$ with the impulse response of the line, $g(t)$.

$$f(t) = \int_0^t F(t - \tau)g(\tau)d\tau.$$

This integral reduces to the following special cases:

[5] S. Goldman, "Transformation Calculus and Electrical Transients," Prentice-Hall, Inc., New York, N. Y., p. 423; 1949.

*Case I:* $0 < t \leq Tl$   $f(t) = 0$ since $g(\tau) = 0$ for $\tau < Tl$

*Case II:* $Tl \leq t \leq Tl + a$

$$f(t) = \int_0^x \left(\frac{x - \tau}{a}\right) \tau^{-3/2} e^{-\beta/\tau} d\tau \qquad x = t - Tl$$

*Case III:* $t > Tl + a$

$$f(t) = \int_0^{x-a} \tau^{-3/2} e^{-\beta/\tau} d\tau$$

$$+ \int_{x-a}^x \left(\frac{x - \tau}{a}\right) \tau^{-3/2} e^{-\beta/\tau} d\tau, \qquad x = t - Tl.$$

Note that Case II is contained in Case III providing that the integrands are limited to positive values of $\tau$ only for Case II.

Considering Case III only and evaluating with the aid of the identity derived in Appendix I, one obtains

$$f(t) = \text{cerf} \sqrt{\frac{\beta}{x-a}} + \frac{x}{a}\left(\text{cerf} \sqrt{\frac{\beta}{x}} - \text{cerf} \sqrt{\frac{\beta}{x-a}}\right) \quad (11)$$

$$- \frac{1}{a} \int_{x-a}^x \tau \alpha \tau^{-3/2} e^{-\beta/\tau} d\tau.$$

Integrating the last term of (11) by parts one obtains

$$\frac{1}{a} \int_{x-a}^x \tau \alpha \tau^{-3/2} e^{-\beta/\tau} d\tau$$

$$= \frac{x}{a} \text{cerf} \sqrt{\frac{\beta}{x}} - \frac{x-a}{a} \text{cerf} \sqrt{\frac{\beta}{x-a}}$$

$$- \frac{1}{a} \int_{x-a}^x \text{cerf} \sqrt{\frac{\beta}{\tau}} d\tau. \quad (12)$$

Observing that the first two terms of (12) cancel the corresponding terms of (11), the function $f(t)$ is simply,

$$f(t) = \frac{1}{a} \int_{x-a}^x \text{cerf} \sqrt{\frac{\beta}{\tau}} d\tau \qquad \begin{aligned} x &\geq 0 \\ x &= t - Tl \end{aligned} \quad (13)$$

with the understanding that for $x < a$ the lower limit is zero.

As verification, one may note that the limit of the ramp response as "$a$" approaches zero is simply the step response. Also, as $x$ gets large, the function approaches unity; physical interpretation of the function required that this be true.

## EVALUATION OF CONSTANTS

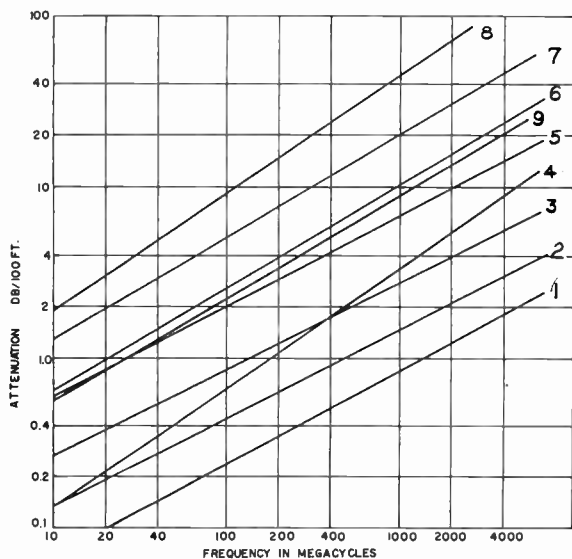Using the first two terms of (6), the propagation constant is approximately

$$\gamma(p) = \rho T + \frac{K}{2R_0} p^{1/2}$$

$$\gamma(j\omega) = \frac{K}{2R_0}\sqrt{\frac{\omega}{2}} + j\left(\omega T + \frac{K}{2R_0}\sqrt{\frac{\omega}{2}}\right).$$

The real part of $\gamma(j\omega)$ is the attenuation constant of the transmission line, for the purposes of the analysis, called $C(f)$.

$$C(f) = \frac{K\sqrt{\pi f}}{2R_0} \text{ nepers/meter.} \quad (14)$$

Any coaxial cable whose attenuation constant obeys the above law will have a straight line relation of slope one-half between the logarithm of the attenuation constant and the logarithm of the frequency. The majority of types of coaxial cable have very nearly this characteristic (see Fig. 1). The ratio of $C(f)$ to $\sqrt{f}$ from (14) is therefore a constant for each type of cable and can be calculated from the attenuation characteristic of the cable.



1) Styroflex $1\frac{5}{8}$ inches     6) General Radio—874A2
2) Styroflex $\frac{7}{8}$ inch     7) RG—58 A/u
3) Styroflex $\frac{1}{2}$ inch     8) RG—38, 39, 40/u
4) RG—19, 20/u     9) RG—8/u
5) RG—63/u

References:
1), 2), 3)—Brochure of Phelps-Dodge Copper Products Corp.
4), 5), 7), 8), 9)—"Reference Data for Radio Engineers," Federal Telephone and Radio Corp., 3rd ed.
6)—Catalog N, General Radio Co.

Fig. 1—Attenuation vs frequency characteristics for common coaxial cables.

In this way, the value of $K$, and subsequently of $\beta$, can be evaluated for each case as follows:

$$\beta = \left(\frac{lK}{4R_0}\right)^2 = \left(\frac{l}{4R_0} \frac{2R_0 C(f_0)}{\sqrt{\pi f_0}}\right)^2 = \left(\frac{lC(f_0)}{2\sqrt{\pi f_0}}\right)^2 \quad (15)$$

where $f_0$ is the frequency chosen to evaluate $\beta$. For convenience in calculation let $l = T_l/T$ where $T_l$ is the time length of the cable and $T = \sqrt{LC}$ is the delay per unit length.

$$\beta = \left(\frac{T_l C(f_0)}{2T\sqrt{\pi f_0}}\right)^2. \quad (16)$$

## RESISTIVE TERMINATION

The analysis assumes that the transmission line is terminated in its characteristic impedance which is given in (4b). However, in the ordinary circuit, a purely resistive termination of value $R_0 = \sqrt{L/C}$ would be used. To see at what frequencies $R_0$ would be a good approximation for $Z_0$, the following comparison of actual $Z_0$ with $R_0$ is made.

From (4b)

$$Z_0 = \sqrt{\frac{pL + K\sqrt{p}}{pC}} = \left(R_0^2 + \frac{K}{C\sqrt{p}}\right)^{1/2}$$

$$= R_0 + \frac{K}{2R_0C\sqrt{p}} - \frac{K^2}{8R_0^3C^3p} + \cdots . \quad (17)$$

The fractional deviation of $Z_0$ from $R_0$ as a function of $p$ is less than the second term of (17) divided by $R_0$. The smallness of the magnitude of this fraction indicates the closeness of approximation.

$$\left|\frac{Z_0(p) - R_0}{R_0}\right| < \left|\frac{K}{4R_0^2C\sqrt{p}}\right| = \frac{K}{4R_0^2C\sqrt{2\pi f}}. \quad (18)$$

Since $R_0^2C = L$ then (18) is the same as (7). Thus, $A$, the validity constant calculated previously is also an expression of the departure of $Z_0$ from $R_0$.

## GENERALIZATION OF THEORY

In order to present curves with which any transient problem involving skin effect distortion of rise times could be solved, the theory is generalized. First, the assumption is made that any rising function can be approximated sufficiently closely for engineering analysis by a series of a few straight line segments. The response to any function can then be obtained from the sum of the responses to the ramp functions used for approximation. A generalized ramp response is then the function to be plotted.

Recalling from the analysis the three basic functions,

Impulse response $\equiv g(t)$

$$= g(x + Tl) = \sqrt{\frac{\beta}{\pi}}\, x^{-3/2} e^{-\beta/x} \quad (9)$$

Step response $\equiv f(t) = f(x + Tl) = \operatorname{cerf}\sqrt{\frac{\beta}{x}} \quad (10)$

Ramp response $\equiv h(t)$

$$= h(x + Tl) = \frac{1}{a}\int_{x-a}^{x} \operatorname{cerf}\sqrt{\frac{\beta}{\tau}}\, d\tau \quad (13)$$

$$x \geqq 0, \text{ all cases,}$$

the problem is to generalize them so that $\beta$, the constant which is determined by the specific case, does not appear in the functions, but only in the scales to which the responses are plotted.

As the first step, the transformation $x = \beta\rho$ is used in (9). The resulting function of $\rho$ is[6]

$$g_0(\rho) = \frac{\rho^{-3/2} e^{-1/\rho}}{\beta\sqrt{\pi}} \qquad \rho \geqq 0 \quad (19)$$

or

$$\beta g_0(\rho) = \frac{\rho^{-3/2} e^{-1/\rho}}{\sqrt{\pi}} \qquad \rho \geqq 0. \quad (20)$$

To apply the normalized impulse response (20) as plotted in Fig. 2 to a specific case, the $\beta$ is calculated from (15) or (16) using physical data. The horizontal scale is then multiplied by $\beta$ and the vertical scale divided by $\beta$ to obtain the impulse response $g(x + Tl)$ vs $x$.
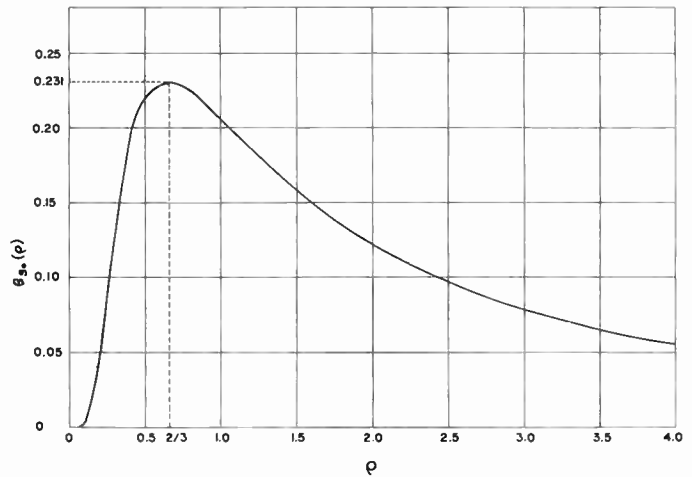


Fig. 2—Normalized impulse response,

$$\beta g(\rho) = \frac{\rho^{-1/\rho}\rho^{-3/2}}{\sqrt{\pi}}.$$

Performing the same transformation in (10), a normalized step response is obtained.

$$h_0(\rho) = \operatorname{cerf}\sqrt{\frac{1}{\rho}} \qquad \rho \geqq 0. \quad (21)$$

To obtain $h(x + Tl)$ vs $x$ the horizontal scale is multiplied by the proper $\beta$.

Likewise, performing the same operation on (13), the normalized ramp response is obtained.

$$f_0(\rho) = \frac{1}{a'}\int_{\rho-a'}^{\rho} \operatorname{cerf}\sqrt{\frac{1}{\rho}}\, d\rho \qquad \rho \geqq 0 \quad (22)$$

where $a' = a/\beta$.

This represents a family of curves (Figs. 3, 4, and 5) with $a'$ as the parameter. Practical utilization of them again requires only a time scale multiplication of magnitude $\beta$. Thus, the response of a particular piece of coaxial cable is obtained for a series of ramp inputs with 0–100 per cent rise times of $a'\beta$. For $a' = 0$ the step re-

[6] This transformation is simple; however much confusion can arise if one does not state and visualize the problem. This is particularly true with respect to obtaining (22). See Appendix II for details.
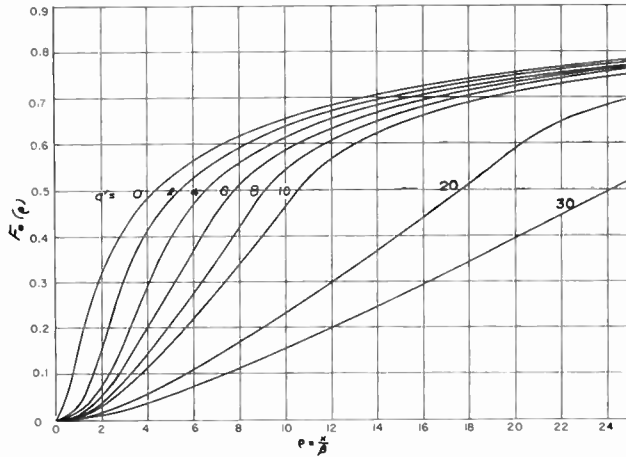
Fig. 3—Normalized ramp responses,

$$f_0(\rho) = \frac{1}{a'} \int_{\rho-a'}^{\rho} \operatorname{cerf} \sqrt{\frac{1}{\rho}}\, d\rho.$$

$$a' = \frac{a}{\beta}.$$
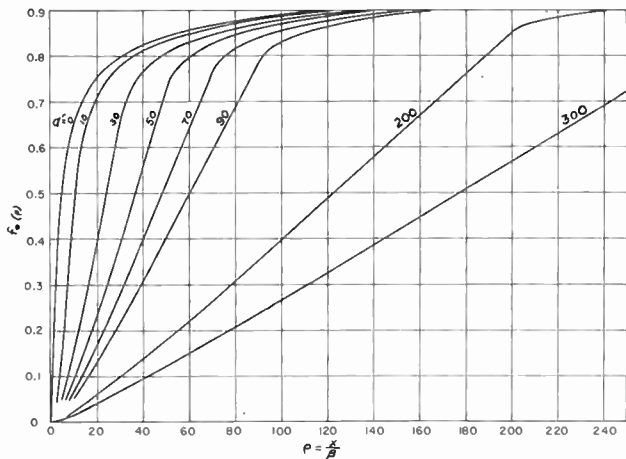


Fig. 4—Normalized ramp responses,

$$f_0(\rho) = \frac{1}{a'} \int_{\rho-a'}^{\rho} \operatorname{cerf} \sqrt{\frac{1}{\rho}}\, d\rho.$$

$$a' = \frac{a}{\beta}.$$

sponse (21) is obtained. The ramps corresponding to $a'$ larger than the largest one plotted are relatively undistorted.

## EXPERIMENTAL VERIFICATION

The experimental verification of the analysis which has been presented required the use of an extremely wide-band oscilloscope. Facilities which were available at the Naval Research Laboratory were used to obtain the transient response of eight pieces of coaxial cable.[7] Two time lengths of each of four types of cable, namely, RG-8/U, RG-58/AU, General Radio-874A2, and $\frac{7}{8}$-inch-diameter Styroflex, were tested. The signal applied to

Fig. 5—Normalized ramp responses,

$$f_0(\rho) = \frac{1}{a'} \int_{\rho-a'}^{\rho} \operatorname{cerf} \sqrt{\frac{1}{\rho}}\, d\rho.$$
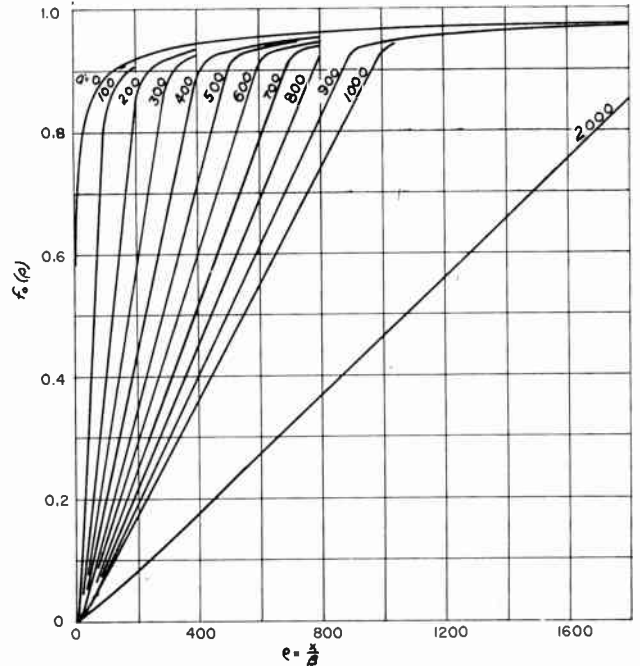
$$a' = \frac{a}{\beta}.$$

the cables was approximated by five ramp functions, and the response was calculated and compared with the observed response for each case.

## EXPERIMENTAL SYSTEM

Fig. 6 shows the cable comparison test circuit employing the NRL TW-10 traveling-wave cathode-ray tubes as the indicating instrument. The TW-10 has a bandwidth well in excess of 2000 mc, which should be sufficient for displaying rise times of the order of 0.1 m$\mu$s.
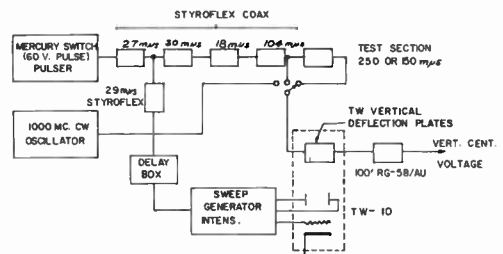


Fig. 6—Cable comparison test circuit.

The test pulse was generated by a mercury contact relay pulser giving a 60-volt pulse, 45 m$\mu$s wide and having a rise time of 0.25 m$\mu$s. Some signal delay (179 m$\mu$s of $\frac{7}{8}$-inch Styroflex) was required to allow time for operation of the sweep and intensifier circuits of the crt. The pulse observed at the end of the 179-m$\mu$s delay was called the standard pulse. Cable test sections of either 150 or 250 m$\mu$s were added, and the response

of the added sections to the standard pulse, as well as the standard pulse itself, were recorded photographically. Time reference was added to each photograph by applying a 1000-mc sine wave to the crt and taking double exposures.

## ANALYSIS OF DATA

Data was taken from the photographs using the sine wave as the time reference and the maximum amplitude of the standard pulse as the amplitude reference.

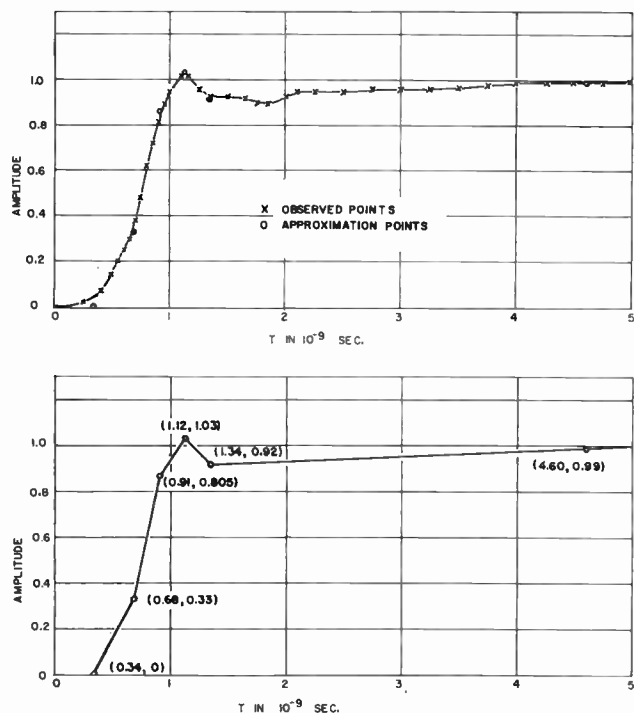The rise of the standard pulse (Fig. 7) was approx-



Fig. 7—Standard pulse and linear approximation.

imated by five straight line-segments as specified in the following Table I.

segment, were calculated from the general curves in Figs. 3, 4, and 5.

The general curves consider ramp responses for ramps of amplitude unity; therefore, it was necessary to correct the amplitudes as listed in Table I. Points (in time) for calculation were preselected so that when the ramp responses were shifted according to the correct $t_0$ (listed in Table I) addition of ordinates would give the response to the standard pulse. The calculated responses as compared to the observed responses are given in Figs. 8–11 (next page).

In all cases no attempt was made to keep track of the zero time position of the transients. No information as to the time at which the transient first departed from zero amplitude after passing through a test section with respect to the time at which the transient "entered" the test section could be obtained. This difficulty is the same as is always met in relating physical transient data to mathematical prediction. The mathematician can define exactly a time before which the system is quiescent. However, the engineer must define the beginning of a transient as the time at which the waveform reaches same measurable value.

For comparison of calculation and observation, therefore, the curves were shifted in time relative to each other so the leading edges most nearly coincided at the region of steepest slope.

## EXPERIMENTAL RESULTS AND DEPARTURES FROM THEORY

From the comparisons of Figs. 8–11, one may conclude that in the coaxial cables considered the major cause of distortion of fact rise time transients is the skin effect. Each type of cable seems to have its own characteristic departure from the predicted response. During this study the causes of some of the departure has become apparent.

First, the analysis involves an approximation in taking the inverse transform of the transfer function as

TABLE I
ANALYSIS OF STANDARD PULSE

| Line Segment | End Points of Segments (10⁻⁹ second, Amplitude) | Amplitude | 0–100 Per Cent Rise Time | $t_0$ |
|---|---|---|---|---|
| 1 | (0.34, 0); (0.68, 0.33) | 0.330 | $0.34 \times 10^{-9}$ second | 0 |
| 2 | (0.68, 0.33); (0.91, 0.805) | 0.535 | 0.23 | $0.34 \times 10^{-9}$ second |
| 3 | (0.91, 0.805); (1.12, 1.03) | 0.165 | 0.21 | 0.57 |
| 4 | (1.12, 1.03); (1.34, 0.92) | −0.110 | 0.22 | 0.78 |
| 5 | (1.34, 0.92); (5.00, 1.00) | 0.080 | 3.66 | 1.00 |

The approximation to the standard pulse is then a succession of ramp functions having rise times and amplitudes as specified above and each starting at the appropriate $t_0$.

The $\beta$ and appropriate values for $a'$ for each case were calculated from (16) and $a' = a/\beta$ [see (22)]. Considering now each example (*i.e.*, 150-mμs delay of ⅞-inch Styroflex), five ramp responses, one for each approximation

expressed in the validity constant $A$ (7). The $A$ for each case is indicated on the graphs (Figs. 8–11). As yet no quantitative measure has been developed to determine limits of error due to a particular value of $A$. However, the values of $A$ in the examples considered are believed to be sufficiently small as to cause negligible error in the time range plotted. One may note that in the propagation constant $\gamma(p)$ (6) the first term ignored is a con-
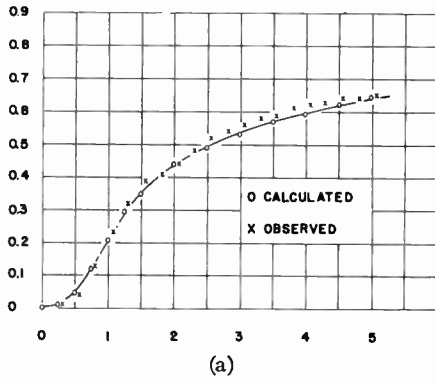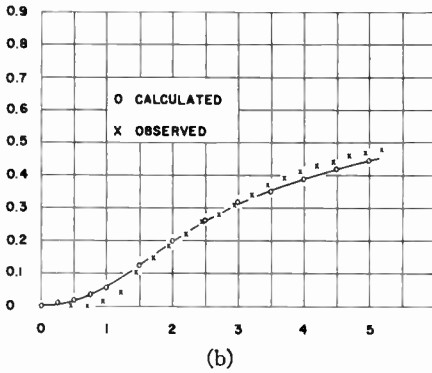
Fig. 8—Response of RG—58 A/u; $A$ (100 mc) = 00.56.
(a) 150 mμs of cable—$\beta = 4.50 \times 10^{-10}$ second.
(b) 250 mμs of cable—$\beta = 1.25 \times 10^{-9}$ second.



Fig. 9—Response of GR—874A2. $A$ (100 mc) = 0.0027.
(a) 150 mμs of cable—$\beta = 1.02 \times 10^{-10}$ second.
(b) 250 mμs of cable—$\beta = 2.83 \times 10^{-10}$ second.
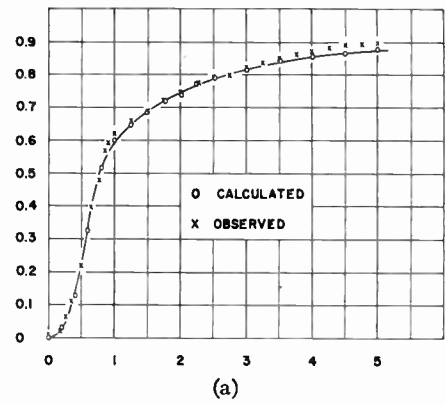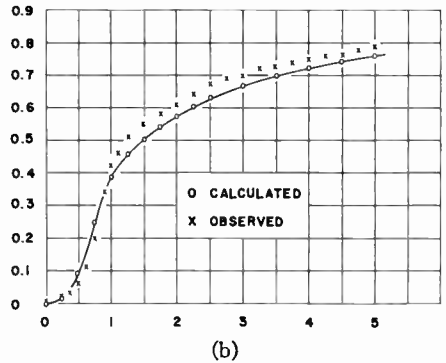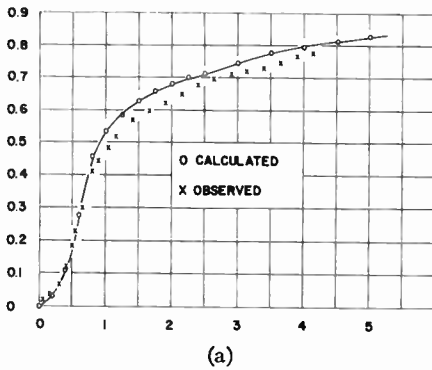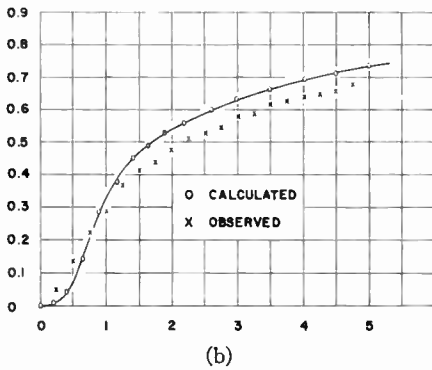


Fig. 10—Response of RG—8/u. $A$ (100 mc) = 0.0024.
(a) 150 mμs of cable—$\beta = 8.14 \times 10^{-11}$ second.
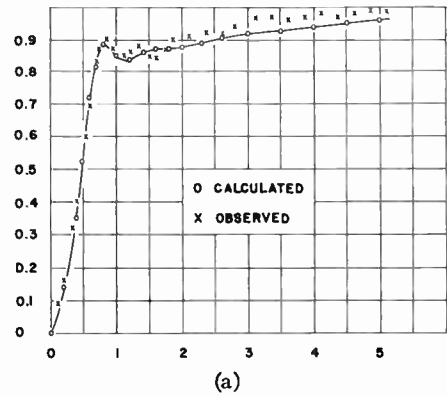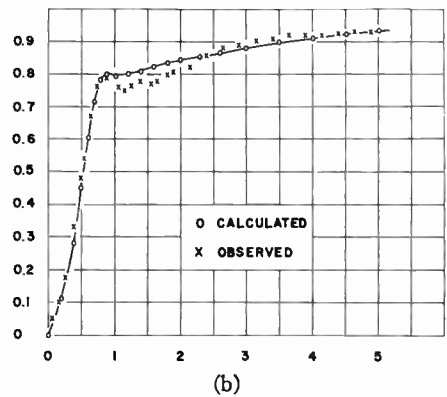(b) 250 mμs of cable—$\beta = 2.26 \times 10^{-10}$ second.



Fig. 11—Response of ⅛ inch Styroflex. $A$ (100 mc) = 0.00057.
(a) 150 mμs of cable—$\beta = 4.57 \times 10^{-12}$ second.
(b) 250 mμs of cable—$\beta = 1.27 \times 10^{-11}$ second.

stant ($p^0$ term) which adds nothing to the distortion and only insignificantly affects the amplitude.

The analysis assumes a $f^{0.5}$ law for the variation of attenuation with frequency [see (3) and (4)]. This is very nearly true for Styroflex cable. However, other cables have a somewhat greater exponent, GR-874 being as high as 0.6. A more elaborate analysis using $f^m$, $0 \leq m \leq 0.5$, has been made; however its usefulness is questionable since it cannot be directly related to the real physical problem. A realistic approach is to search for a second distorting factor such as dielectric loss which in this study was assumed to be negligible. Dielectric loss should be greater for GR-874 and other polyethylene dielectric cables than for Styroflex, although still it should not be the major distorting mechanism. Work on this phase of the problem is continuing.

Useful engineering results may be obtained even though the $f^{0.5}$ law is not followed exactly by the cable. The choice of the frequency at which $\beta$ is evaluated (16) then becomes important. The frequency chosen in this study was $f_0 = 1000$ mc because the components of most importance were in the region of 1000 mc (considering a logarithmic frequency scale).

The bandwidth of the TW-10 was considered to be sufficient not to distort appreciably the response. The 10–90 per cent rise time of the standard pulse is 0.5 m$\mu$s. Approximately 700–900 mc of bandwidth (to the 3-db points) is needed to pass such a rise. The designers of the TW-10 oscilloscope system have established that the 3-db point of the deflection structure is well in excess of 2000 mc although no detailed data of deflection as a function of frequency is available. The ringing which is evident in some of the responses is probably due to the slight impedance discontinuities in the system.

Another possible source of error is in the nonlinearity of the crt deflection as a function of input amplitude. Checking this possibility showed that the crt deflection was within approximately 2 per cent of being linear. A slight curvature of the field of view (sometimes called "pin-cushion effect") made transcription of amplitude data difficult for time values of 3 to 5 m$\mu$s after the beginning of each response. Errors of up to 4 per cent (positive) may arise from this cause.

The RG-8 flexible connection between the TW-10 and the waveform to be observed (not explicitly shown in Fig. 6) does introduce appreciable distortion in the crt display; however, it does not invalidate the technique used to check the analysis.

Referring to Fig. 6, let the waveform entering the test section be represented by $F_1(p)$.[8] Let the transfer function of the 15-m$\mu$s RG-8 connecting cable be $G_1(p)$. Also let $F_1'(p)$ represent the waveform observed on the CRT (the standard pulse) when the test section is not included. Then, $F_1'(p) = F_1(p)G_1(p)$. Now let $G_2(p)$ be the transfer function of the test section of cable. Then,

$F_2(p)$ which represents the waveform observed on the CRT when the test section is included is given by

$$F_2(p) = F_1(p)G_2(p)G_1(p)$$
$$= F_1(p)G_1(p)G_2(p) = F_1'(p)G(p)$$

since transfer functions of passive networks are commutative.

In words, what this means is that the distorting element, $G_1(p)$ having been present both in observation of the input and output of the test section allows isolation of the characteristics of the test section alone. This is the basis for all comparison type measurement techniques. For accuracy, the distortion due to $G_1(p)$ must be of the same order of magnitude or preferably less than that due to $G_2(p)$. It is less in all cases.

## Conclusion

The analysis as described is a first order theory for the transient response of coaxial cables. As presented, it is useful in engineering problems involving milli-microsecond transients, however, later refinements in the theory may permit greater accuracy for cables in which dielectric loss is an appreciable factor.

## Appendix I

The following identity was useful in the analysis.

$$I(x) \equiv \int_0^x \sqrt{\frac{\beta}{\pi}} \, \tau^{-3/2} e^{-\beta/\tau} d\tau = \text{cerf} \sqrt{\frac{\beta}{x}} .$$

It may be verified by using Laplace Transformation operational theorems.[9] Letting $L$ indicate the operation of taking the Laplace Transform and $L^{-1}$ the inverse,

$$L[I(x)] = \frac{1}{p} L\left[\sqrt{\frac{\beta}{x}} \, x^{-3/2} e^{-\beta/x}\right] = \frac{1}{p} e^{-2\sqrt{\beta p}}$$

$$I(x) = L^{-1}L[I(x)] = L^{-1}\left[\frac{1}{p} e^{-2\sqrt{\beta p}}\right] = \text{cerf} \sqrt{\frac{\beta}{x}}.$$

This inverse has been listed.[5]

Since a function which is expressed as a definite integral with a variable in the limits is a function only of the limits, then

$$I(x - a) = \int_0^{x-a} \sqrt{\frac{\beta}{\pi}} \, \tau^{-3/2} e^{-\beta/\tau} d\tau = \text{cerf} \sqrt{\frac{\beta}{x - a}}.$$

## Appendix II

The normalization of (9), (10), and (13) to obtain (19), (21), and (22) is performed as follows. Consider first (9) and (10).

$$g(x + Tl) = \sqrt{\frac{\beta}{\pi}} \, x^{-3/2} e^{-\beta/x} \qquad x \geq 0 \qquad (9)$$

$$h(x + Tl) = \text{cerf} \sqrt{\frac{\beta}{x}} \qquad x \geq 0. \qquad (10)$$

---

[8] These expressions are given in complex variable form as Laplace transforms of the time functions.

[9] C. R. Wylie, "Advanced Engineering Mathematics," McGraw-Hill Book Co., Inc., New York, N. Y.; 1951.

Let $x = \beta\rho$

$$g(\beta\rho + Tl) = \sqrt{\frac{\beta}{\pi}}(\beta\rho)^{-3/2}e^{-1/\rho} = \frac{\rho^{-3/2}e^{-1/\rho}}{\beta\sqrt{\pi}}$$

$$h(\beta\rho + Tl) = \text{cerf}\sqrt{\frac{1}{\rho}}.$$

As written above, the functions $g$ and $h$ are still plotted on the $x$ time scale although $x$ does not appear in the expressions. Changing the time scale to the dimensionless $\rho$ ($\beta$ has the dimensions of time) new functions $g_0(\rho)$ and $h_0(\rho)$ are obtained.

$$g_0(\rho) = \frac{\rho^{-3/2}e^{-1/\rho}}{\beta\sqrt{\pi}} \qquad \rho \geqq 0 \qquad (19)$$

$$h_0(\rho) = \text{cerf}\sqrt{\frac{1}{\rho}} \qquad \rho \geqq 0. \qquad (21)$$

For plotting, (19) is changed to

$$\beta g_0(\rho) = \frac{\rho^{-3/2}e^{-1/\rho}}{\sqrt{\pi}} \qquad \rho \geqq 0. \qquad (20)$$

Note that in the transformation the shape of the functions were preserved, and in order to plot the functions $g(x+Tl)$ and $h(x+Tl)$ for any particular physical case the horizontal scale is altered by the factor $\beta$ for that case. In (20) the vertical scale must also be altered by the factor $\beta$.

Considering (13), more care must be used in the change of time scales.

$$f(x + Tl) = \frac{1}{a}\int_{x-a}^{x}\text{cerf}\sqrt{\frac{\beta}{\tau}}\,d\tau \qquad x \geqq 0. \qquad (13)$$

In the above, change the scale on the dummy variable by the substitution $t = \beta\rho$. A corresponding change of scale must be made in the limits by dividing by $\beta$.

$$f(x + Tl) = \frac{1}{a}\int_{(x-a)/\beta}^{x/\beta}\text{cerf}\sqrt{\frac{1}{\rho}}\,\beta d\rho.$$

The function is now set up for normalization by letting $x = \beta\rho$ and plotting the resulting function $f_0(\rho) \equiv f(\beta\rho + Tl)$ vs $\rho$

$$f_0(\rho) = f(\beta\rho + Tl) = \frac{\beta}{a}\int_{(\beta\rho-a)/\beta}^{\beta\rho/\beta}\text{cerf}\sqrt{\frac{1}{\rho}}\,d\rho.$$

Finally, letting $a' = a/\beta$,

$$f_0(\rho) = \frac{1}{a'}\int_{\rho-a'}^{\rho}\text{cerf}\sqrt{\frac{1}{\rho}}\,d\rho \qquad \rho \geqq 0. \qquad (22)$$

### ACKNOWLEDGMENT

# CORRECTION

The editors wish to point out the following correction to "SSB Performance as a Function of Carrier Strength," by William L. Firestone, which appeared on pages 1839–1848 of the December, 1956 issue of PROCEEDINGS. On page 1843, the illustrations in the first column identified as Fig. 10 and Fig. 11 should be transposed.

# The Serrodyne Frequency Translator*

RAYMOND C. CUMMING†, SENIOR MEMBER, IRE

*Summary*—A serrodyne frequency translator will translate, or shift, the frequency of a signal in a nearly ideal manner. Linear sawtooth modulation of a transit-time device, such as a traveling-wave tube or klystron, is employed to effect the translation. The power output at the translated frequency is practically equal to the capability of the same device operating as an ordinary amplifier. Furthermore, very little power is produced in undesired intermodulation frequency components.

Frequency translations of a microwave signal ranging from subaudio frequencies to 57 megacycles have been accomplished. The translation loss for a translation of about 30 megacycles was less than 1 decibel; moreover, each undesired frequency component was suppressed at least 20 decibels. The serrodyne method can be used at nonmicrowave frequencies and with larger fractional frequency shifts, provided modulable transit-time devices having certain properties can be devised.

An analysis is given of serrodyne performance as limited by a number of important practical factors. A general spectrum analysis is included which is in a convenient form for a variety of problems in combined AM and PPM (or combined AM and PM) when arbitrary modulating waveforms are used.

## INTRODUCTION

THE SERRODYNE[1] is a frequency translator which employs linear sawtooth modulation of transit time. Its performance closely approaches that of an ideal frequency translator.

An ideal frequency translator, given an input signal of a certain frequency, will produce an output signal whose frequency is shifted by some desired amount from that of the input. A maximum of power will be produced at the desired output frequency, and no power will be produced at other, undesired, frequencies.

In general, the input signal may consist of a band of frequencies instead of a single frequency component. Frequency translation ordinarily involves a modulation process; the frequency of the modulation is numerically equal to the translation frequency or to a submultiple thereof.

Some often-used synonyms for the name frequency translator include: Frequency shifter,[2] frequency converter,[3] single-sideband modulator,[4] and synchrodyne.[5] Frequency translators are used in microwave relay sta-

tions,[6,7] in measuring apparatus,[8] in superheterodyne receivers, and in other important systems applications.

Diemer and Knol[3] have shown that continuous linear variation of phase will produce ideal frequency conversion, or translation. They gave, however, no adequate means for realizing the method. Giacoletto[9,10] suggested that linear sawtooth modulation of the drift field in a filamentary transistor would produce a realization of the Diemer and Knol idea. Independently, the author successfully used linear sawtooth modulation of the beam voltage of klystrons and traveling-wave tubes (twt) to produce frequency translation of microwave signals. Frequency shifts ranging from subaudio frequencies to 57 mc were imparted to a microwave signal by use of an S-band twt.[11] Larger frequency shifts can be produced, apparently, by extension of the techniques employed.

Fig. 1 shows the measured output spectrum of a T-301 twt modulated by a 28.7 mc sawtooth. In the case shown, the desired output frequency component is the first lower-side frequency (labeled $n = -1$). Notice that the amplitude of the desired component is within 1 db of the capability of the same twt operating as an amplifier. The undesired components are each suppressed by a factor of at least 20 db. The circuit diagram of the sawtooth generator and its connections to the electron gun of the twt are given in Fig. 2.

In the past, microwave frequency translation has been accomplished by means of a *sinusoidal* modulating waveform.[2,5-7] The theoretical maximum power at the desired output frequency, using sinusoidal modulation, is 4.7 db less than the output power obtainable from the same tube operating as an amplifier. Furthermore the output spectrum is symmetrical; hence, several undesired components are present and comparable in amplitude with the desired component.

[1] The name "serrodyne" is derived from the Latin word "serra," meaning "saw" or "sawtooth."
[2] W. V. Bray, "The traveling-wave valve as a microwave phase-modulator and frequency shifter," *Proc. IEE*, vol. 99, pt. III, pp. 15–10; January, 1952.
[3] G. Diemer and K. S. Knol, "Frequency conversion by phase variation," *Philips Res. Rep.*, vol. 4, pp. 161–167; June, 1949.
[4] J. Cacheris, "Microwave single-sideband modulator using ferrites," *Proc. IRE*, vol. 42, pp. 1242–1247; August, 1954.
[5] A. E. Harrison, "Klystron Tubes," McGraw-Hill Book Co., Inc., New York, N. Y., 1st ed.; 1947.

[6] V. Learned, "The klystron mixer applied to tv relaying," *Proc. IRE*, vol. 38, pp. 1033–1035; September, 1950.
[7] W. W. Siekanowitz, "A developmental medium-power traveling-wave tube for relay service in the 2,000 mc region," *Proc. IRE*, vol. 42, pp. 1091–1097; July, 1954.
[8] P. D. Lacy and G. W. C. Mathers, "New twt amplifiers with provision for simulating special microwave signals," *Hewlett-Packard J.*, vol. 7, pp. 1–4; January, 1956.
[9] L. J. Giacoletto, "Semiconductor Frequency Converter," U. S. Patent 2,701,302; February 1, 1955.
[10] This method is limited by the restriction that the transit time through the filamentary transistor must be short compared with the period of the modulation; otherwise, severe distortion of the sawtooth modulating waveform will be produced because of the averaging which will occur during the transit time. Furthermore, the analysis given by Giacoletto, being based on phase modulation theory rather than on considerations of transit time, fails to show the above limitation and also yields an incorrect value for the optimum amplitude of sawtooth modulation.
[11] The Stanford University T-301, which is essentially the same as the Huggins HA-2A, was the tube used.
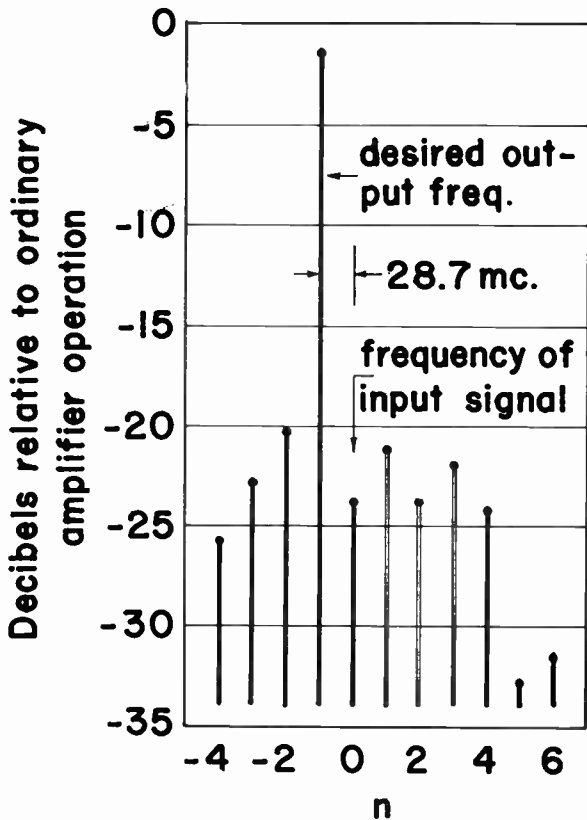
Fig. 1—Output spectrum from a traveling-wave tube operating as a serrodyne. The sawtooth modulation has a fundamental frequency of 28.7 mc. Zero db is the output of the same twt operating as an ordinary amplifier with the same input signal power.
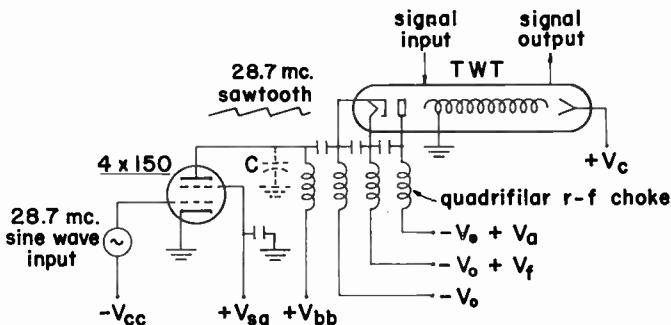


Fig. 2—Basic circuit of 28.7-mc sawtooth generator modulating the electron-gun potential of a traveling-wave tube. Stray capacitance C is charged linearly through the multifilar rf choke and discharged quickly by the Class-C-biased 4X150 tube.

Frequency translation systems using sinusoidal modulation normally employ a filter to eliminate the undesired frequency components. The serrodyne, on the other hand, using sawtooth transit-time modulation (TTM), suppresses the undesired components to a large degree by virtue of the modulation process itself; thus the need for a filter is eliminated or reduced drastically. This reduced requirement on output filtering can be very important, for example, in systems in which the signal frequency varies rapidly over a band.

The purpose of this paper is to give a physical description of serrodyne operation and to derive theoretical bounds on the quality of performance obtainable. The analysis applies to presently developed or foreseeable modulable transit-time devices like the klystron, the twt, the ferrite phase shifter,[12] the gas-discharge phase shifter, the filamentary transistor, and the time-variable electrical delay network. In such devices, only a limited deviation of transit time can be produced by modulation. As a consequence, the ratio of translation frequency to signal frequency will be small compared with unity and hence will be so assumed in much of the analysis.

## Description of Serrodyne Operation

A *transit-time device* is a device in which an output cycle appears with a certain time delay after the occurrence of the corresponding input cycle. The time delay is termed the *transit time* for that cycle. Let us assume that the effect of each input cycle is propagated to the output by means of a physical disturbance (for example, in a klystron, the disturbances consist of electron bunches; in a twt, of combined bunches and traveling waves). Such a transit-time device is illustrated schematically in Fig. 3.
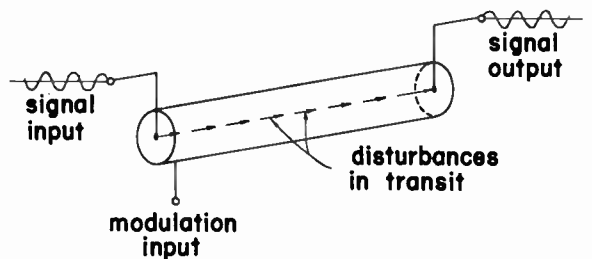


Fig. 3—Schematic drawing of a modulable transit-time device.

Obviously, if the transit time is constant (no ttm applied) and the input frequency is constant also, the spacing of disturbances arriving at the output terminal will be uniform, and the output frequency will be equal to that of the input. Modulation of the transit time will alter the spacing of disturbances at the output terminal. In particular, if the modulation consists of a *linear* increase of transit time with time, the disturbances will arrive at a lower, but uniform, frequency. Similarly a linear *decrease* of transit time will give a higher, uniform, output frequency. The action may be thought of as a sort of Doppler effect.

In practice, of course, the linear increase or decrease of transit time cannot be continued indefinitely. Therefore, a periodic linear sawtooth waveform is used.[13] In

[12] R. F. Soohoo, "Ferrite micfowave phaseshifters," 1956 IRE Convention Record, pt. 5, pp. 84–98.

[13] Periodicity is not a necessary requirement on the modulating waveform. The necessary requirements, it will be seen, determine only the slope of the linear portion and the amplitude of the discontinuity; a nonperiodic sawtooth-type wave can satisfy these requirements. However, for simplicity, only periodic modulating waveforms will be considered in this paper.
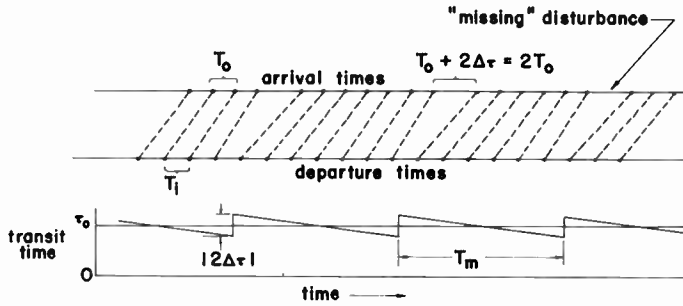
Fig. 4—Departure and arrival times of disturbances. The transit time is modulated as shown, by a periodic linear sawtooth. Input frequency $= 1/T_i$; modulation frequency $= 1/T_m$; output frequency $= 1/T_o = 1/T_i + 1/T_m$.
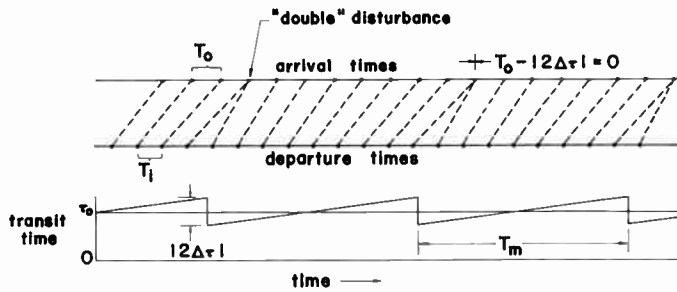


Fig. 5—Similar to Fig. 4, but for a modulating sawtooth having opposite slope. Output frequency $= 1/T_o = 1/T_i - 1/T_m$.

Figs. 4 and 5 are shown the departure and arrival times of disturbances when periodic linear sawtooth modulation is applied to the transit time.

The use of periodic instead of continuous modulation brings up a question of phase coherence. The signal output during every modulation period must be in phase with that of every other modulation period, in order to give a large output power at a single output frequency.

The phase coherence problem can be clarified by a study of the intervals between disturbance arrival times, shown in Figs. 4 and 5. During the linear portion of the sawtooth wave, the intervals between arrival times is equal to $T_o$, the desired output period (reciprocal of the desired output frequency). At the flyback of the sawtooth wave, a discontinuity of magnitude $|2\Delta\tau|$ is introduced in the transit time; it follows that the interval between arrival times of adjacent disturbances on opposite sides of the discontinuity will be equal to $T_o + 2\Delta\tau$. If the linear variation of transit time is in the negative direction (*i.e.*, transit times progressively smaller), then the sign of the discontinuity $2\Delta\tau$ will be taken as positive, and vice versa.

Phase coherence will be assured if the magnitude of the discontinuity $|2\Delta\tau|$ is set equal to the desired output period $T_o$, or to an integral multiple thereof. In Figs. 4 and 5, $|2\Delta\tau|$ has been equated to $T_o$, and the resulting phase coherence is apparent. Fortunately, as will be shown later, the value of $2\Delta\tau$ for optimum performance is not critical. If the equation

$$2\Delta\tau = \bar{n}T \quad (1)$$

where

$\bar{n} = $ a positive or negative integer[14]

is satisfied only approximately, good performance will result.

If the period of the input signal is $T_i$, and the rate of change of transit time is $\dot{\tau}$, and the modulating wave is sampled at the disturbance departure times, it is easy to verify that the period $T_o$ of the output signal will be given by

$$T_o = T_i(1 + \dot{\tau}). \quad (2)$$

With periodic linear sawtooth modulation of transit time, $\dot{\tau}$ is a constant, given by the following equation, which is in accord with the above-mentioned sign convention on $\Delta\tau$:

$$\dot{\tau} = -\frac{2\Delta\tau}{T_m}, \quad (3)$$

where $\Delta\tau$ is the amplitude and $T_m$ is the period of the modulation. Elimination of $\dot{\tau}$ and $\Delta\tau$ between (1), (2), and (3) gives the following relation between input, output, and modulation frequencies:

$$\frac{1}{T_o} = \frac{1}{T_i} + \bar{n}\,\frac{1}{T_m}. \quad (4)$$

From this it can be seen that the translation frequency is $\bar{n}/T_m$, an integral multiple of the modulation frequency $1/T_m$.

Notice in Figs. 4 and 5 that, when phase coherence exists, the output is perfectly periodic, except for a number (equal to $|\bar{n}|$) of either "missing" or else "overlapping" disturbances in each modulation period. These interruptions in periodicity are caused by the discontinuities in the sawtooth wave. Their effect on serrodyne performance, along with a number of other effects, will now be analyzed.

## ANALYSIS OF SERRODYNE PERFORMANCE

In general, a modulation process produces a large number of intermodulation frequency components. The separation between the component frequencies is numerically equal to the fundamental frequency of the modulation.[15]

If the goal of the modulation process is to effect frequency translation, only one of the intermodulation components will be desired in the output; it is called the desired output frequency; all other components are undesired output frequencies. The relative strengths of the desired and undesired frequency components can be specified by the *suppression* and the *translation loss*, defined in decibels as follows:

---

[14] It will be shown shortly (4) that $\bar{n}$ is the order of the intermodulation side frequency taken as the desired output frequency. In practice, $\bar{n}$ is usually given a value of either plus unity of minus unity.

[15] For a somewhat more accurate specification of the component frequencies, see (30) of Appendix I.

db suppression of a particular undesired frequency component

$$= 20 \log \frac{\text{output amplitude for ordinary amplifier operation}}{\text{amplitude of the particular undesired frequency component}}, \qquad (5)$$

$$\text{db translation loss} = 20 \log \frac{\text{output amplitude for ordinary amplifier operation}}{\text{amplitude of desired output frequency component}}. \qquad (6)$$

For example, in the output spectrum shown in Fig. 1, the suppression is about 20 db and the loss is about 1 db. (When the term "suppression" is not qualified by the specification of a particular frequency component, the suppression of the larger undesired component adjacent to the desired component will be meant.)

When the analytical results are presented, it will be seen that the translation loss of a serrodyne is usually of the order of 1 db, or less, and hence is usually negligible in practice. Therefore, the performance figure of most practical interest will be the suppression.

The following items constitute the principal degrading factors to serrodyne performance. They will presently be treated one by one.

1) The "missing" or "overlapping" disturbances mentioned in connection with Figs. 4 and 5.

2) The finite flyback time of realizable approximations to a perfect linear sawtooth waveform.

3) Nonlinearity in the sawtooth.

4) AM which often accompanies the application of TTM to a device.

5) Bandwidth. That is, variations of the input or modulation frequency (in the presence of a fixed amplitude of TTM).

6) Nonoptimum amplitude of sawtooth modulation.

### Method of Analysis

The mathematical analysis is based on a model of a transit-time device like that shown in Fig. 3. A modulating wave is applied to the device to vary the time required for disturbances to make the transit from input terminal to output terminal. It is assumed that the modulating wave is sampled instantaneously at the moment a disturbance embarks from the input terminal.

For example, in a twt or klystron amplifier, variation of the beam voltage will modulate the injection velocity of the electron beam. An electron bunch, once formed, will travel at a constant velocity, approximately equal to the injection velocity of the electrons in the bunch. Thus, the velocity, and hence the transit time, of the bunch are determined at the instant the bunch leaves the input terminal.

On the other hand, in some modulable transit-time devices, such as the ferrite phase shifter, the gas-discharge phase shifter, the filamentary transistor, and the time-variable electrical delay network, sampling of the modulating wave will not occur instantaneously. Instead, each disturbance will experience a varying value of modulation throughout its transit. The analysis to follow tacitly assumes that for such devices the transit time is short compared to the period of the modulation. Under this assumption the sampling becomes essentially instantaneous.

Another general assumption is that the disturbances arriving at the output terminal have a fixed *shape*, independent of the modulation. Ordinarily some variation of disturbance shape is to be expected when modulation is applied. However, since the main consequence of a change in disturbance shape is a corresponding change in gain, the variation of disturbance shape is essentially an AM effect. This AM effect can be taken into account later by the application of the proper waveform of AM in the final results.

The mathematical model described above is identical to that for combined AM and PPM with periodic sampling.[16] Appendixes I through VI make use of existing AM and PPM analysis as the basis for deriving the results which follow.

### Results of Analysis

The theoretical performance of a serrodyne is given below in terms of the suppression and translation loss, defined by (5) and (6). The effects which limit serrodyne performance are considered to exist one at a time, all other effects being assumed negligible in each case. This procedure is realistic, for in practical situations one of the effects usually dominates the others.

*"Missing" or "Overlapping" Disturbances:* This phenomenon was discussed in connection with Figs. 4 and 5. It is intimately related to the fact that we are dealing with TTM and not PM; analyses based on PM methods ignore the effect completely.

Appendix II shows that serrodyne performance, in the presence of only the "missing" or "overlapping" disturbance effect, will be given by

$$\text{db suppression} = 20 \log \frac{\omega}{|\omega_t|}, \qquad (7)$$

$$\text{db translation loss} = 0, \qquad (8)$$

where $\omega$ is the frequency of the input signal, and $\omega_t$ is the translation frequency. If the device is operating as a frequency multiplier as well as a frequency translator, replace $\omega$ in (7) by $\bar{m}\omega$, where $\bar{m}$ is an integer equal to the multiplication factor.

[16] S. H. Moss, "Frequency analysis of modulated pulses," *Phil. Mag.*, vol. 39, pp. 663–691; September, 1948.

For the example shown in Fig. 1, the ratio $\omega/|\omega_t|$ is approximately (3000 mc)/(30 mc), or 100, giving a db suppression value from (7) of 40 db. Since the suppression actually obtained was only 20 db, the performance in the case illustrated was evidently limited chiefly by considerations other than the "missing" or "overlapping" disturbance effect.

*Finite Flyback Time:* It is impossible to generate a sawtooth wave having zero flyback time. If the duration of the flyback is a fraction $F$ of a modulation period, Appendix III shows that serrodyne performance will be

$$\text{db suppression} = -20 \log F, \qquad (9)$$

$$\text{db translation loss} = -20 \log (1 - F). \qquad (10)$$

For the example illustrated in Fig. 1, the sawtooth was known to have a fractional flyback time $F$ of about one-tenth. This value of $F$, substituted in (9) and (10), closely checks the values of 20-db suppression and 1-db loss, which are apparent in Fig. 1.

*Nonlinearity in the Sawtooth Wave:* As a specific example of nonlinearity, we shall consider the square-root variation of electron velocity with beam voltage which exists in tubes like the klystron and twt. In such tubes, the disturbance transit time varies approximately inversely with the square root of the beam voltage. If we assume that linear modulation is applied to the beam voltage, the resulting variation of transit time will be nonlinear, and Appendix IV shows that serrodyne performance will be

$$\text{db suppression} = 20 \log \frac{2\pi N}{3\bar{n}^2}, \qquad (11)$$

$$\text{db translation loss} = -20 \log \left(1 - \frac{\pi^2}{40} \frac{\bar{n}^4}{N^2}\right), \qquad (12)$$

where $N$ is the length of the transit-time device expressed in wavelengths at the desired output frequency and at the disturbance velocity in the absence of modulation; and $\bar{n}$ is a positive or negative integer equal to the order of the intermodulation side frequency which is taken as the desired output frequency (for example, $\bar{n} = -1$ in the spectrum given in Fig. 1).

In a typical twt, the value of $N$ may be, for example, about 30. For such a tube, operating with linear sawtooth modulation of beam voltage, (11) and (12) show that the suppression should be about 36 db and the loss approximately 0 db.

*Incidental Amplitude Modulation:* In many modulable transit-time devices, the gain and transit time are not independent. Thus, modulation of the transit time will necessarily cause some gain variation, or AM. In a twt, for example, modulation of the beam voltage about the synchronous value (the value giving maximum gain) will produce the desired TTM, but will obviously produce some undesired AM, as well.

For some important devices (the klystron and twt are examples), the variation of amplitude gain with instantaneous modulation amplitude is a parabolic function. An analysis is given in Appendix V of serrodyne performance when such a device is used. The results are given in terms of an easily measurable parameter $D_\theta/\theta^2$, which denotes the sensitivity of gain to variations in transit time. The following two paragraphs comprise a brief digression to define the parameter $D_\theta/\theta^2$.

The quantity $D_\theta$ is the decibel gain reduction which results when the transit time is changed, from the value giving maximum gain, by a small amount $\delta\tau$. The angle $\theta$ is the phase shift of a signal of the desired output frequency which would be produced by a time delay $\delta\tau$; that is, $\theta = \bar{\Omega}\delta\tau$, where $\bar{\Omega}$ is the desired output radian frequency (*i.e.*, $\bar{\Omega} = 2\pi/T_o$).

The values of $D_\theta$ and $\theta$ can be measured for a klystron or twt as follows: Apply a signal of frequency $\bar{\Omega}$ to the input terminal of the klystron or twt. Adjust the beam voltage to give maximum gain. Change the beam voltage a small amount, and measure the resulting power reduction and phase shift at the output terminal. The power reduction, expressed in decibels, is $D_\theta$, and the phase shift is $\theta$. The ratio $D_\theta/\theta^2$ will be independent of the magnitude of the beam voltage change, provided the gain vs transit time relation is parabolic.

Appendix V gives the following results:

$$\text{db suppression} = -20 \log \left(0.229\bar{n}^2 \frac{D_\theta}{\theta^2}\right), \qquad (13)$$

$$\text{db translation loss} = -20 \log \left(1 - 0.377\bar{n}^2 \frac{D_\theta}{\theta^2}\right). \qquad (14)$$

It can be shown[17] that the value of $D_\theta/\theta^2$ for a klystron is

$$\frac{D_\theta}{\theta^2} = \frac{2.36}{N^2} \qquad (15)$$

and for a twt, approximately

$$\frac{D_\theta}{\theta^2} = \frac{0.39}{NC} \qquad (16)$$

where $C$ is the twt gain parameter defined by Pierce.[18] Since $N$ is always very much greater than $C$, these equations show that the effect of AM will be much less pronounced in a klystron than in a twt. This is to be expected, since there is no velocity-synchronism requirement in a klystron.

*Bandwidth:* It was pointed out in connection with (1) that the optimum peak-to-peak amplitude of sawtooth ttm is numerically equal to the reciprocal of the desired output frequency (or to an integral multiple thereof). The question may arise: What will happen to serrodyne performance if the desired output frequency is varied (by changing the frequency of the input signal or of the

[17] R. C. Cumming, "Theory of the Serrodyne at Low Translation Frequencies," Stanford Univ., (Stanford Electronics Labs.), Tech. Rep. 262-2; May 1, 1956.
[18] J. R. Pierce, "Traveling-Wave Tubes," D. Van Nostrand Co., Inc., New York, N. Y., 1950.

modulation) in the presence of a fixed amplitude of saw-tooth TTM? The answer is derived in Appendix VI in terms of the total fractional bandwidth $B$ of the output signal. The suppression and loss for operation at the extreme edges of the band are shown to be

$$\text{db suppression} = 20 \log \left| \frac{\pi(|\bar{n}| B/2 - 1)}{\sin \bar{n}\pi B/2} \right|, \quad (17)$$

$$\text{db translation loss} = 20 \log \left| \frac{\bar{n}\pi B/2}{\sin \bar{n}\pi B/2} \right| . \quad (18)$$

*Nonoptimum Amplitude of Sawtooth TTM:* The analysis of this condition is identical to that discussed above for bandwidth. This time the output frequency is held constant and the amplitude of sawtooth TTM is changed. The performance can be expressed as

$$\text{db suppression} = 20 \log \left| \frac{\pi(|\bar{n}| X - 1)}{\sin \bar{n}\pi X} \right|, \quad (19)$$

$$\text{db translation loss} = 20 \log \left| \frac{\bar{n}\pi X}{\sin \bar{n}\pi X} \right|, \quad (20)$$

where $X$ is the fractional discrepancy in sawtooth amplitude from the optimum.

### DISCUSSION

Eqs. (7) through (20) give measures of serrodyne performance as limited by any of six effects, each considered in the absence of the others. When one calculates these six performance limitations for a practical situation, he may find that two (or more) of the limitations are comparable in magnitude. In such a case, analysis shows that the performance will be degraded somewhat from that computed considering only one effect at a time.

For example, consider a twt having a gain-sensitivity parameter $D_\theta/\theta^2$ of 0.13; this tube is to be modulated by a sawtooth wave which has a fractional flyback time $F$ of 0.03. Eq. (13) shows that consideration of the AM effect alone gives a suppression of about 30.5 db. Similarly, (9) gives a suppression of about 30.4 db. when the effect of finite flyback time is considered alone. The general analysis in Appendix I can be used to show that if both effects exist simultaneously, the suppression will be reduced to about 27.2 db.

Nothing has been said so far about the amplitude of beam-voltage modulation required for serrodyne operation of a klystron or twt. For a klystron, the modulation amplitude follows easily from a consideration of (1) and of the square-root variation of electron velocity (and disturbance velocity) with beam voltage. The result for a klystron is

$$\frac{\Delta V}{V_0} = \frac{|\bar{n}|}{N} \quad (21)$$

where $\Delta V$ is the peak amplitude of beam-voltage modulation required, $V_0$ is the average value of the beam voltage, $N$ is the length of the klystron drift tube ex-

pressed in wavelengths at the desired output frequency when propagation is at the average electron velocity, and $\bar{n}$ is an integer equal to the order of the intermodulation side frequency chosen as the desired output frequency.

For a modulated twt, the variation in disturbance velocity (*i.e.,* in phase velocity of the waves) is not as large as the variation in electron velocity; this is because the wave velocity is dependent on the properties of the slow-wave circuit (which are of course unmodulated), as well as on the velocity of the electron beam. It can be shown[19] that the sensitivity of transit time to beam voltage is in fact about 50 per cent less for a typical twt than for a klystron; thus, for a typical twt, the peak fractional beam-voltage modulation $\Delta V/V_0$ required for serrodyne operation is approximately

$$\frac{\Delta V}{V_0} = \frac{3}{2} \frac{|\bar{n}|}{N} \quad (22)$$

in which $N$, the distance between input and output terminals expressed in wavelengths at the desired output frequency, is reckoned using the average phase velocity of the waves, which is approximately equal to the average electron velocity.

In a twt, the effect of AM is considerably greater than the effect of the square-root variation of electron velocity with beam voltage, independent of $\bar{n}$ and $N$. This can be seen by consideration of (11), (13), and (16), assuming $C < 0.19$, as is practically always the case. The opposite is true for a klystron: The square root variation is more important than the AM, independent of $\bar{n}$. Eqs. (11), (13), and (15) can be used to show this fact, assuming $N > 1.1$, as is usually the case.

In practice a value of unity is usually taken for the magnitude of $\bar{n}$, the order of intermodulation side frequency taken as the desired output frequency. This is done for the following reasons.

1) The required modulation voltage is minimized, as shown by (21) and (22).

2) The amplitude of incidental AM is obviously minimized by minimizing the modulation voltage.

3) The spacing between undesired spectral lines is maximized. This follows from the fact that the fundamental modulation frequency required for a given translation of frequency is inversely proportional to $|\bar{n}|$.

4) The performance for operation over a band of frequencies is maximized, as shown by (17) and (18).

5) The effects of nonoptimum modulation amplitude are minimized, as shown by (19) and (20).

6) At first glance, it might seem that $F$, the fractional flyback time, could be reduced by doubling $|\bar{n}|$ and thereby halving the modulation frequency. This does not follow, however, since in the usual types

[19] R. C. Cumming, "Frequency Translation by Modulation of Transit-Time Devices," Stanford Univ. (Appl. Electronics Lab.), Tech. Rep. 39 (Task 32); August 1, 1955.

of sawtooth generator, it is the *rate* of flyback which is a constant, fixed by available components. Therefore, since the amplitude and the period of the required sawtooth wave are both proportional to $|\bar{n}|$, the value of $F$ will be unchanged by an increase in the value of $|\bar{n}|$.

It may be well to point out that modulation applied to the helix potential of a twt may cause standing waves at modulation frequency to be set up on the helix. Hence, for modulation frequencies exceeding the order of a few megacycles, the beam-voltage modulation should be applied to the cathode (or better, to the entire electron gun) and not to the helix.[20]

In practical applications of the serrodyne system, only small fractional frequency shifts have been accomplished. If it is desired to translate a signal downward in frequency by a large relative amount, as for example in a superheterodyne receiver application, an important problem must be solved: A transit-time device must be found in which the transit time can be varied at least one period *at the desired output frequency*. To produce a large fractional frequency translation in a downward direction, this means that a variation of a large number of periods at the input frequency must be available.

The problem of sawtooth waveform generation will not be considered in detail in this paper. Obviously, the problem becomes more difficult as the sawtooth frequency is increased. The circuit shown in Fig. 2 has been used successfully for generating sawtooth voltage waves having fundamental frequencies up to 57 mc. This circuit uses constant-current charging of a capacitance and rapid discharging by a short pulse from a tetrode tube biased well below cutoff. Another possible approach is to synthesize the sawtooth by a summation of a number of harmonics having carefully adjusted relative amplitudes and phases.

The serrodyne has been defined and described here in terms of devices which can produce TTM with periodic sampling,[16] or an approximation thereto. It should be pointed out that some other types of time modulation exist which are suitable for serrodyne use. Two examples are PM, and TTM with synchronous sampling.[16] True PM, defined as modulation of the argument of a periodic function, can be produced by certain special devices.[21,22] These devices, when sawtooth modulated to a peak-to-peak deviation of $2\pi$ radians, should produce good serrodyne operation.

Finally, for workers with other modulation problems, some aspects of the general spectrum analysis given in Appendix I may be of interest. The results are in a very

convenient form for application to systems having combined AM and PPM (or TTM) by arbitrary modulating waveforms. Furthermore, it can be shown[19] that the results are applicable to combined AM and PM as well, provided the parameter $r$ is taken as the peak phase deviation. For AM waveforms which can be expressed in terms of the time- or angle-modulating waves, it is quite convenient to use (38), which expresses the spectrum in terms of the results for the time or angle modulation alone.

## APPENDIX I

### GENERAL SPECTRUM ANALYSIS

The following analysis treats combined TTM and AM which is produced by periodic waveforms that are arbitrary except for rationally related fundamental frequencies.

Let the shape of a single disturbance (amplitude as a function of time) centered at the time origin be represented by $f(t)$. Then the amplitude of the output $i(t)$ from a transit-time device can be written

$$i(t) = \sum_{k=-\infty}^{\infty} A_k f(t - t_k) \tag{23}$$

where $A_k$ and $t_k$ are the amplitude and arrival time of the $k$th disturbance, respectively.

Assuming the input signal to be periodic of radian frequency $\omega$, we can express the departure time $d_k$ of the $k$th disturbance by

$$d_k = \frac{2\pi k}{\omega}. \tag{24}$$

The arrival time $t_k$ is given by the sum of the departure time $d_k$ and the transit time $\tau_k$

$$t_k = d_k + \tau_k. \tag{25}$$

Let the transit time and amplitude be modulated by arbitrary periodic functions $\Delta\tau S(\omega_r t)$ and $qH(\omega_a t)$, respectively; in which:

$S(\xi)$ and $H(\xi)$ are periodic functions having period $2\pi$, amplitude unity, and arbitrary waveform (examples are the functions $\sin \xi$, $\cos \xi$, and the function saw $\xi$, which is illustrated in Fig. 6 below).
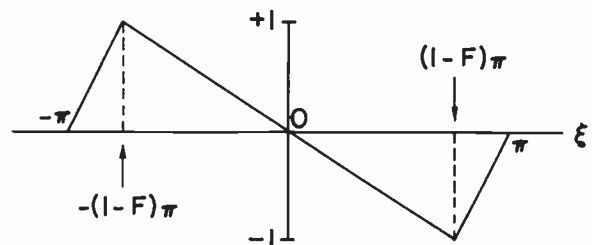


Fig. 6—One cycle of the periodic function saw $\xi$.

$\omega_r$    is the fundamental radian frequency of the TTM.
$\omega_a$    is the fundamental radian frequency of the AM.
$\Delta\tau$    is the peak deviation of transit time.
$q$      is the AM depth-of-modulation factor.

[20] R. C. Cumming, "Resonance Effects in Serrodyne Modulation of Traveling-Wave Tubes," Stanford Univ. (Stanford Electronics Labs.), Tech. Rep. 262-1; May 1, 1956.

[21] R. E. Shelby, "A cathode-ray frequency modulation generator," *Electronics*, vol. 13, pp. 14–18; February, 1940.

[22] R. Adler, "A new system of frequency modulation," PROC. IRE, vol. 35, pp. 25–31; January, 1947.

Thus, the modulated transit-time function $\tau_k$ and amplitude function $A_k$ can be written as follows, assuming sampling of the modulating waveforms to occur at the instants of departure $d_k$

$$\tau_k = \tau_0 + \Delta\tau S(\omega_r d_k) \qquad (26)$$

$$\frac{A_k}{A_0} = 1 + qH(\omega_a d_k) \qquad (27)$$

where $\tau_0$ and $A_0$ are the values of transit time and amplitude in the absence of modulation.

For simplicity, and with no loss of generality for the serrodyne application, we assume the fundamental frequencies of TTM and AM to be rationally related

$$\frac{\omega_r}{\omega_a} = \frac{N}{P} \qquad (28)$$

where the ratio of integers $N/P$ is in lowest terms (*i.e.*, $N$ and $P$ have no integer factors in common).

It can be shown by an extension[19] of the work of Moss[16] that the spectrum analysis of $i(t)$, which we desire, can be written[23]

$$i(t) = \operatorname{Re} \sum_n \sum_\Omega M_n(q, \Omega\Delta\tau)g(\Omega)\epsilon^{j\Omega t} \qquad (29)$$

in which the frequency components $\Omega$ are given by

$$\Omega = m\omega + n\omega_m > 0 \qquad (30)$$

where $m$ and $n$ are positive or negative integers which are consistent with the inequality sign; and $\omega_m$ is the fundamental radian frequency of the combined TTM and AM, given by

$$\omega_m = \frac{\omega_r}{N} = \frac{\omega_a}{P} \qquad (31)$$

and $g(\Omega)$ is a complex factor which depends on the disturbance shape, and is simply $\omega/\pi$ times the Fourier integral of the shape of an individual disturbance

$$g(\Omega) = \frac{\omega}{\pi} A_0 \int_{-\infty}^{\infty} f(t)\epsilon^{-j\Omega t}dt \qquad (32)$$

and $M_n(q, r)$ is a complex factor which depends on the waveforms of TTM and AM, and is given by the following integral, which is easily evaluated for the waveforms of interest

$$M_n(q, r) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [1 + qH(M\xi)]\epsilon^{-j[rS(N\xi)+n\xi]}d\xi. \qquad (33)$$

The parameter $r$ is a modulation index, which, as indicated in (29), is given by

$$r = \Omega\Delta\tau. \qquad (34)$$

The fact that the modulation index is different for each spectral component $\Omega$ is characteristic of TTM employing periodic sampling.[16]

[23] Re means "real part of."

This completes the most general spectrum analysis. The remainder of this appendix presents a number of simplifications, definitions, and evaluations of integrals which will be helpful in Appendixes II through VI, which analyze serrodyne performance.

1) In the case of TTM only (no AM present), $q=0$, and we have

$$M_n(0, r) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \epsilon^{-j[rS(\xi)+n\xi]}d\xi. \qquad (35)$$

2) In some practical cases, the AM waveform can be written as a power series in the TTM waveform

$$qH(\xi) = q_1 S(\xi) + q_2 S^2(\xi) + \cdots \qquad (36)$$

where $q_1, q_2, \cdots$, are real constants. In such cases $\omega_a$ is equal to $\omega_r$, and the function $M_n(q, r)$, as defined by (33), becomes

$$M_n(q, r) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [1 + q_1 S(\xi) + q_2 S^2(\xi) + \cdots ] \\ \cdot\epsilon^{-j[rS(\xi)+n\xi]}d\xi. \qquad (37)$$

Successive differentiation with respect to $r$ of (35), and substitution into (37), gives the useful relation

$$M_n(q, r) = M_n(0, r) + jq_1 M_n'(0, r) \\ - q_2 M_n''(0, r) + \cdots \qquad (38)$$

where

$$M_n'(0, r) = \frac{\partial}{\partial r} M_n(0, r), \text{ etc.} \qquad (39)$$

Eq. (38) expresses the function $M_n(q, r)$, which applies for combined TTM and AM, in terms of the function $M_n(0, r)$, which is for TTM alone. An example of the use of (38) is given in Appendix V.

3) In Appendixes II through VI it is assumed that the ratio $\omega_m/\omega$ is either irrational or small (or both). If $\omega_m/\omega$ is irrational, then the summation over $n$ in (29) reduces to a single term because no two combinations $m, n$ will give the same value for $\Omega$ in (30). If $\omega_m/\omega$ is small, then only one term in the summation over $n$ in (29) will be of appreciable magnitude. In either case (29) reduces to

$$i(t) = \operatorname{Re} \sum_\Omega M_n(q, \Omega\Delta\tau)g(\Omega)\epsilon^{j\Omega t} \qquad (40)$$

from which it is evident that the complex amplitude $I(\Omega)$ of the frequency component $\Omega$ is

$$I(\Omega) = M_n(q, \Omega\Delta\tau)g(\Omega). \qquad (41)$$

Let us normalize $I(\Omega)$ with respect to the amplitude of the component which would exist at the frequency $m\omega$ in the absence of modulation. This normalized amplitude is designated $\hat{I}(\Omega)$ and is defined by

$$\hat{I}(\Omega) = \frac{I(\Omega)}{I(m\omega)_{\text{no mod.}}} \qquad (42)$$

By equating $\Delta\tau$ and $n$ to zero in (41), (30), (32), and (35), we find that the denominator of (42) is equal to $g(m\omega)$. Often the components $\Omega$ of interest are in the near vicinity of $m\omega$ (*i.e.*, small fractional frequency translation with the serrodyne, for example). For such cases, the approximation

$$g(\Omega) \approx g(m\omega) \tag{43}$$

is valid, and the normalized amplitude $\hat{I}(\Omega)$ of the component at $\Omega$ becomes simply

$$I(\Omega) \approx M_n(q, \Omega\Delta\tau). \tag{44}$$

4) Since this paper is concerned with frequency translation, one of the components $\Omega$ will be the desired output frequency. We shall designate the desired output frequency and quantities related to the desired output frequency by use of bars over their algebraic symbols. Thus, we denote the desired output frequency by $\bar{\Omega}$. And we define integers $\bar{m}$ and $\bar{n}$ in accord with (30)

$$\bar{\Omega} = \bar{m}\omega + \bar{n}\omega_m. \tag{45}$$

Similarly the modulation index for the desired output frequency is written $\bar{r}$ and is given by (34)

$$\bar{r} = \bar{\Omega}\Delta\tau. \tag{46}$$

5) The definitions of suppression and loss, defined by (5) and (6), can now be expressed mathematically by use of (44). Recall that when the term "suppression" is not qualified by the specification of a particular frequency component, the suppression of the larger undesired component adjacent to the desired component is meant.

$$\text{db suppression} = -20\log \left| M_{\bar{n}\pm1}[q, (\bar{\Omega} \pm \omega_m)\Delta\tau] \right|, \tag{47}$$

$$\text{db translation loss} = -20\log \left| M_{\bar{n}}(q, \bar{\Omega}\Delta\tau) \right|. \tag{48}$$

In cases in which the modulation frequency is low compared with the output frequency, (47) becomes, to a good approximation

$$\text{db suppression} \approx -20\log \left| M_{\bar{n}\pm1}(q, \bar{\Omega}\Delta\tau) \right|. \tag{49}$$

6) A periodic linear sawtooth function saw $\xi$ having a fractional flyback time $F$ and a period $2\pi$ is illustrated in Fig. 6. Mathematically, saw $\xi$ can be defined as an odd function of $\xi$, given for $\xi$ between 0 and $\pi$ by

$$\text{saw } \xi = \frac{-\xi}{\pi(1 - F)}, \quad 0 < \xi < \pi(1 - F)$$

$$= \frac{\xi - \pi}{\pi F}, \qquad \pi(1 - F) < \xi < \pi. \tag{50}$$

The function $M_n(0, r)$ for TTM by the waveform saw $\xi$ follows from a straightforward integration of (35)

$$M_n(0, r) = \frac{r \sin\left[r - n\pi(1 - F)\right]}{(r + n\pi F)\left[r - n\pi(1 - F)\right]}. \tag{51}$$

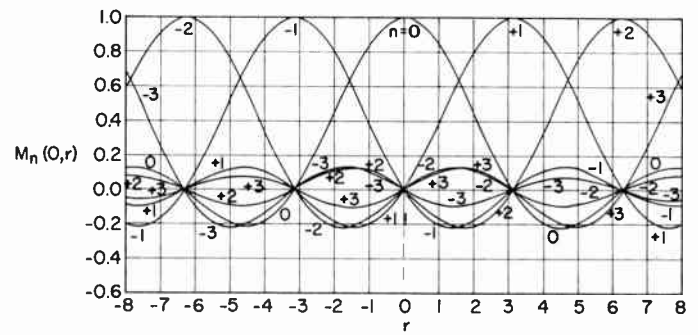This function is plotted in Figs. 7 and 8 for $F = 0$ and $F = 0.20$, respectively.



Fig. 7—Amplitudes of several intermodulation side frequencies as functions of the modulation index $r(=\Omega\Delta\tau)$. These curves are for ttm by an ideal linear sawtooth having zero flyback time. Near-ideal frequency translation occurs for $|r|$ in the vicinity of $\pi$, or a multiple thereof.
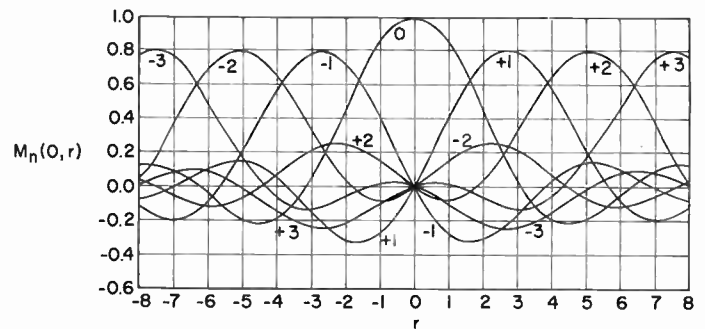


Fig. 8—Similar to Fig. 7, but for a sawtooth having a fractional flyback time $F$ of 0.20.

In order to obtain a maximum amplitude at the desired output frequency, we must maximize the above function for $n = \bar{n}$. Let $\bar{r}(\text{opt.})$ designate the value of $r$ which maximizes the function $M_{\bar{n}}(0, r)$. It can be shown[19] that to a good approximation

$$\bar{r}(\text{opt.}) = \bar{n}\pi(1 - F). \tag{52}$$

Combining (46) and (52), we find that the optimum amplitude of sawtooth TTM is given by

$$\Delta\tau(\text{opt.}) = \frac{\bar{n}\pi(1 - F)}{\bar{\Omega}}. \tag{53}$$

This reduces to (1) if $F$ is zero, since, by definition, $\bar{\Omega} = 2\pi/T_o$.

## Appendix II

### "Missing" or "Overlapping" Disturbances

In this appendix we analyze ideal serrodyne operation, as was described in connection with Figs. 4 and 5. Ideal serrodyne operation is specified by setting: $S(\xi) = \text{saw } \xi$ (*i.e.*, linear sawtooth TTM), $F = 0$ (*i.e.*, zero flyback time), $q = 0$ (*i.e.*, no AM present), and $\Delta\tau = \Delta\tau(\text{opt.})$.

Eqs. (51) and (53) reduce to, respectively

$$M_n(0, r) = \frac{\sin(r - n\pi)}{r - n\pi}, \tag{54}$$

$$\Delta\tau(\text{opt.}) = \frac{\bar{n}\pi}{\bar{\Omega}}. \tag{55}$$

Substitution of (55) into (54) and simplification yields

$$M_{\bar{n}}[0, \bar{\Omega}\Delta\tau(\text{opt.})] = 1 \tag{56}$$

and

$$M_{\bar{n}\pm1}[0, (\bar{\Omega} \pm \omega_m)\Delta\tau(\text{opt.})] = \frac{\sin\dfrac{\pi\omega_t}{\bar{\Omega}}}{\pi - \dfrac{\pi\omega_t}{\bar{\Omega}}} \tag{57}$$

in which $\omega_t$ is the translation frequency, defined by

$$\omega_t = \bar{n}\omega_m. \tag{58}$$

If the fractional translation of frequency is small ($|\omega_t| \ll \bar{\Omega}$), then (57) reduces to approximately

$$M_{\bar{n}\pm1}[0, (\bar{\Omega} \pm \omega_m)\Delta\tau(\text{opt.})] \approx \frac{\omega_t}{\bar{m}\omega}. \tag{59}$$

Eqs. (59) and (56) yield (7) and (8) when substituted into (47) and (48), respectively.

## Appendix III

### Finite Flyback Time

In this appendix we assume $F$ nonzero and use (49) instead of (47) to define the suppression.

Substitution of (53) into (51) and simplification gives

$$M_{\bar{n}}[0, \bar{\Omega}\Delta\tau(\text{opt.})] = 1 - F \tag{60}$$

and

$$M_{\bar{n}\pm1}[0, \bar{\Omega}\Delta\tau(\text{opt.})] = \frac{\bar{n}\sin[\pi(1 - F)]}{(\bar{n} \pm F)\pi}. \tag{61}$$

Eq. (61) reduces to the following, under the assumption $F \ll 1$

$$M_{\bar{n}\pm1}[0, \bar{\Omega}\Delta\tau(\text{opt.})] \approx F. \tag{62}$$

Eqs. (62) and (60), when substituted into (49) and (48), yield (9) and (10), respectively.

## Appendix IV

### Minus-One-Half-Power Variation of Transit Time with Modulation Voltage

In electronic devices like the klystron and twt, the electron injection velocity $v$ is related to the voltage $V$ between cathode and drift tube or helix by the conservation of energy equation

$$\frac{1}{2}mv^2 = eV \tag{63}$$

in which $e/m$ is the ratio of electron charge to mass. It follows that the transit time $\tau$ of an electron through a drift tube or helix of length $L$ will be

$$\tau = \frac{L}{\sqrt{\dfrac{2eV}{m}}}. \tag{64}$$

It can be shown[19] that the transit times of disturbances (*i.e.*, electron bunches in a klystron; waves and bunches in a twt) also follow approximately (64).

Let us assume that sawtooth modulation with zero flyback time is applied to the electron-accelerating voltage $V$

$$V = V_0\left(1 + \frac{\Delta V}{V_0}\text{ saw }\xi\right) \tag{65}$$

where $V_0$ is the value of $V$ in the absence of modulation, $\Delta V$ is the peak deviation of $V$, and $\xi$ is the instantaneous phase of the sawtooth modulation.

Combination of (64) and (65) and retention of the first three terms of a power-series expansion gives

$$\tau = \tau_0 - \frac{1}{2}\tau_0\frac{\Delta V}{V_0}\text{ saw }\xi + \frac{3}{8}\tau_0\left(\frac{\Delta V}{V_0}\right)^2(\text{saw }\xi)^2 \tag{66}$$

in which

$$\tau_0 = \frac{L}{\sqrt{\dfrac{2eV_0}{m}}}. \tag{67}$$

A close approximation to the peak deviation of transit time $\Delta\tau$ can be obtained by considering the first two terms of (66) at a time when saw $\xi$ is equal to unity

$$\Delta\tau = -\frac{1}{2}\tau_0\frac{\Delta V}{V_0}. \tag{68}$$

Eq. (66) can be written

$$\tau = \tau_0 + \Delta\tau S(\xi) \tag{69}$$

where

$$S(\xi) = \text{saw }\xi + \frac{3}{2}\frac{\Delta\tau}{\tau_0}(\text{saw }\xi)^2. \tag{70}$$

The value of $\Delta\tau(\text{opt.})$ is given with negligible error by (53)

$$\Delta\tau(\text{opt.}) = \bar{n}\pi/\bar{\Omega} \tag{71}$$

Substitution of (70) and (71) into (35) yields

$$M_{\bar{n}}[0, \bar{\Omega}\Delta\tau(\text{opt.})] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\epsilon^{-j(3/4\pi)(\bar{n}^2/N)\xi^2}d\xi \tag{72}$$

$$M_{\bar{n}\pm1}[0, \bar{\Omega}\Delta\tau(\text{opt.})] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\epsilon^{-j\xi((3/4\pi)(\bar{n}^2/N)\xi\pm1)}d\xi \tag{73}$$

where

$$N = \frac{\bar{\Omega}\tau_0}{2\pi}. \tag{74}$$

The approximation

$$\epsilon^{-j\sigma\xi^2} \approx 1 - j\sigma\xi^2 - \tfrac{1}{2}\sigma^2\xi^4 \tag{75}$$

in which $\sigma$ is the small quantity

$$\sigma = \frac{3}{4\pi}\frac{\bar{n}^2}{N} \tag{76}$$

is useful in evaluating (72) and (73), and gives the following results

$$\left| M_{\bar{n}}[0, \bar{\Omega}\Delta\tau(\mathrm{opt.})] \right| = 1 - \frac{\pi^2\bar{n}^4}{40N^2} \tag{77}$$

$$\left| M_{\bar{n}\pm1}[0, \bar{\Omega}\Delta\tau(\mathrm{opt.})] \right| = \frac{3}{2\pi}\frac{\bar{n}^2}{N}. \tag{78}$$

Substitution of (78) and (77) into (49) and (48) gives (11) and (12), respectively.

## Appendix V

### Incidental Amplitude Modulation

We assume in this appendix that the transit time $\tau$ varies linearly and the output amplitude $A$ varies parabolically with modulation voltage. These assumptions equated to (26) and (27) can be written

$$\tau = \tau_0 + \tau_0 P_1 \frac{V - V_0}{V_0} = \tau_0 + \Delta\tau S(\xi) \tag{79}$$

$$\frac{A}{A_0} = 1 + K_2\left(\frac{V - V_0}{V_0}\right)^2 = 1 + qH(\xi) \tag{80}$$

where

$$\xi = \omega_m t \tag{81}$$

and $P_1$ and $K_2$ are real constants ($K_2 < 0$).

It follows from (79) that the peak deviations of transit time and modulation voltage, $\Delta\tau$ and $\Delta V$, respectively, will be related by

$$\frac{\Delta\tau}{\tau_0} = P_1 \frac{\Delta V}{V_0}. \tag{82}$$

Elimination of $(V - V_0)/V_0$ and $\Delta\tau/(\tau_0 P_1)$ between (79), (80), and (82) yields

$$qH(\xi) = K_2\left(\frac{\Delta V}{V_0}\right)^2 [S(\xi)]^2. \tag{83}$$

Comparison of (36) and (83) shows

$$q_2 = K_2\left(\frac{\Delta V}{V_0}\right)^2 \tag{84}$$

$$q_1 = q_3 = q_4 = \cdots = 0. \tag{85}$$

This paragraph is a digression to derive $K_2$ in terms of the more convenient parameter $D_\theta/\theta^2$. By definition, $\theta$ is the phase change in a signal of frequency $\bar{\Omega}$ which would result from a time delay $\tau - \tau_0$

$$\theta = \bar{\Omega}(\tau - \tau_0). \tag{86}$$

Eqs. (79), (80), and (86) can be combined to give $A$ as a function of $\theta$

$$\frac{A}{A_0} = 1 + K_2\left(\frac{\theta}{2\pi N P_1}\right)^2 \tag{87}$$

where

$$N = \frac{\bar{\Omega}\tau_0}{2\pi}. \tag{88}$$

By definition, $D_\theta$ is the decibel gain reduction which results when the transit time of a transit-time device is changed from $\tau_0$ to $\tau$

$$D_\theta = -20\log\left[1 + K_2\left(\frac{\theta}{2\pi N P_1}\right)^2\right]. \tag{89}$$

For $D_\theta$ less than about 3 db, (89) can be solved approximately to give the desired relation between $K_2$ and $D_\theta/\theta^2$:

$$K_2 = -4.55 N^2 P_1^2 \frac{D_\theta}{\theta^2}. \tag{90}$$

The value of $\Delta\tau(\mathrm{opt.})$ can be taken with negligible error as that given by (53):

$$\Delta\tau(\mathrm{opt.}) = \bar{n}\pi/\bar{\Omega}. \tag{91}$$

Combination with (82) gives the optimum value of $\Delta V$

$$\frac{\Delta V(\mathrm{opt.})}{V_0} = \frac{\bar{n}}{2N P_1}. \tag{92}$$

Substitution of (90) and (92) into (84) leads to

$$q_2\bigg|_{\Delta V=\Delta V(\mathrm{opt.})} = -1.14\bar{n}^2\frac{D_\theta}{\theta^2}. \tag{93}$$

From (38) and (85) we have

$$M_n(q, r) = M_n(0, r) - q_2 M_n''(0, r). \tag{94}$$

And since we assume $F = 0$ in this appendix, (51) gives for $M_n(0, r)$

$$M_n(0, r) = \frac{\sin(r - n\pi)}{r - n\pi}. \tag{95}$$

Differentiation of (95) and substitution along with (93) into (94) gives the spectral amplitudes $M_n(q, r)$

$$M_n(q, r) = \frac{\sin x}{x} + 1.14\bar{n}^2\frac{D_\theta}{\theta^2}\left(-\frac{\sin x}{x} - 2\frac{\cos x}{x^2} + 2\frac{\sin x}{x^3}\right)$$

$$x = r - n\pi. \tag{96}$$

Substitution of (91) into (96) and simplification yields

$$M_{\bar{n}}[q, \bar{\Omega}\Delta\tau(\mathrm{opt.})] = 1 - 0.377\bar{n}^2\frac{D_\theta}{\theta^2} \tag{97}$$

$$M_{\bar{n}\pm 1}[q, \bar{\Omega}\Delta\tau(\text{opt.})] = 0.229\bar{n}^2 \frac{D_\theta}{\theta^2}. \qquad (98)$$

Eqs. (98) and (97) when substituted into (49) and (48) give (13) and (14).

## APPENDIX VI

### BANDWIDTH

Eq. (53) shows that $\Delta\tau(\text{opt.})$ is a single-valued function of the output frequency $\bar{\Omega}$. Thus, if the output frequency is to be varied over a band and $\Delta\tau$ is to be held constant, $\Delta\tau$ can be optimized for only one frequency in the band.

We designate by $\bar{\Omega}_0$ the band center of the output signal and by

$$\bar{\Omega}_0 \left(1 \pm \frac{B}{2}\right) \qquad (99)$$

the upper and lower band edges, where $B$ is the total fractional bandwidth. Let us set $\Delta\tau$ to the optimum value at the center of the band, by use of (53) ($F=0$ in this appendix)

$$\Delta\tau = \frac{\bar{n}\pi}{\bar{\Omega}_0}. \qquad (100)$$

From (51), we have

$$M_n(0, r) = \frac{\sin(r - n\pi)}{r - n\pi}. \qquad (101)$$

Substitution of (99) and (100) into (101) gives the following results for operation at the band edges

$$M_{\bar{n}}\left[0, \bar{\Omega}_0\left(1 \pm \frac{B}{2}\right)\Delta\tau\right] = \frac{\sin \bar{n}\pi B/2}{\bar{n}\pi B/2} \qquad (102)$$

$$M_{\bar{n}\pm 1}\left[0, \bar{\Omega}_0\left(1 + \frac{B}{2}\right)\Delta\tau\right] = \frac{\sin \bar{n}\pi B/2}{-\bar{n}\pi B/2 \pm \pi} \qquad (103)$$

$$M_{\bar{n}\pm 1}\left[0, \bar{\Omega}_0\left(1 - \frac{B}{2}\right)\Delta\tau\right] = \frac{\sin \bar{n}\pi B/2}{-\bar{n}\pi B/2 \mp \pi}. \qquad (104)$$

The maximum magnitude of (103) and (104) with respect to the plus-or-minus signs is

$$\left| \frac{\sin \bar{n}\pi B/2}{|\bar{n}|\pi B/2 - \pi} \right|. \qquad (105)$$

Substitution of (105) and (102) into (49) and (48) gives (17) and (18), respectively.

# CORRECTION

Benjamin Lax, author of "Frequency and Loss Characteristics of Microwave Ferrite Devices," which appeared on pages 1368–1386 of the October, 1956 issue of PROCEEDINGS has requested the editors to make the following reference correction to his paper. Footnote 20, in column one of page 1381, should be

[20] S. Weisbaum and H. Boyet, "Broad-band nonreciprocal phase shifts—analysis of two ferrite slabs in rectangular guide," *J. Appl. Phys.*, vol. 27, pp. 519–524; May, 1956.

# Direct-Coupled-Resonator Filters*

SEYMOUR B. COHN†, SENIOR MEMBER, IRE

*Summary*—A new analysis is given of direct-coupled-resonator filters that results in excellent response at much greater bandwidths than has previously been possible. The method relies on the fact that the coupling elements can be made into perfect impedance inverters, or "quarter-wave" transformers, by the addition of negative elements in lumped-constant circuits, or of short negative lengths of line in waveguide and transmission-line circuits. Specific design formulas are given for filters constructed of lumped-constant elements, waveguide, and strip or other TEM transmission line, and for pass band response functions of the maximally flat and Tchebycheff types. The formulas include a simple frequency transformation that corrects for the frequency sensitivity of the coupling reactances, and thereby greatly improves the design accuracy for both lumped-constant and microwave types when the bandwidth is more than a few per cent. Exact response curves computed from typical filter designs are compared to the prototype-function response curves, and it is shown that the design formulas give good results with bandwidths of at least 20 per cent in guide wavelength in the case of waveguide filters, or 20 per cent in frequency for TEM-mode transmission-line and lumped-constant filters.

## INTRODUCTION

IN ORDER to achieve a narrow bandwidth in bandpass filters at high frequencies, it is necessary to utilize high-Q resonant circuits or cavities coupled loosely to each other in cascade. Fig. 1 shows lumped-constant, waveguide, and strip-transmission-line examples of this coupled-resonator class of filters. The relative steepness of cutoff increases with the number of resonators, and therefore methods of synthesizing filters with any number of resonators to have any desired physically realizable response are of particular interest. Considerable work has been done prior to this program on such methods, and for sufficiently narrow bandwidths adequate design formulas are available [1–11]. However, as the bandwidth is increased above 1 per cent, the accuracy of most previous design formulas deteriorates, and therefore a reexamination of the problem was considered desirable. This has been done for the case of direct coupling between resonators by utilizing a different analytical approach than has been used before, and the resulting design equations are given in an easy-to-use form in this paper. In the case of waveguide, strip-line, and other transmission-line filters of this class, a substantial improvement in design accuracy is obtained over prior results, and precise designs are now possible for bandwidths of at least 20 per cent in guide wavelength in the case of waveguide filters, or 20 per cent in frequency for the TEM-mode transmission-line filters. The design equations for the lumped-constant circuits of Fig. 1 are included for complete-

(a) LUMPED CONSTANTS, SERIES CAPACITIVE COUPLING

(b) LUMPED CONSTANTS, MUTUAL INDUCTIVE COUPLING

(c) LUMPED CONSTANTS, SHUNT INDUCTIVE COUPLING

GEN.←— →LOAD

(d) WAVEGUIDE

GEN.←— →LOAD

(e) STRIP TRANSMISSION LINE

Fig. 1—Direct-coupled-resonator filters of various constructions.

ness, and because they are in a form particularly convenient for design application. In addition, the formulas include a simple frequency transformation that corrects for the frequency sensitivity of the coupling reactances (which have been previously assumed to be constant), and thereby greatly improves the design accuracy for both lumped-constant and microwave types when the bandwidth is more than a few per cent.

The work described in this paper on direct-coupled-resonator filters began originally as a survey of design methods for narrow-band waveguide filters. Such filters have been constructed in two ways. In one, waveguide cavities are formed by inductive irises spaced approximately a half-wavelength apart, the individual cavities being coupled to each other through quarter-wavelength waveguide transformers. In the other, adjacent cavities are coupled directly to each other by single irises. The quarter-wave-coupled type permits tuning of the individual cavities before assembly, and the iris dimensions are relatively noncritical. On the other hand, the direct-coupled type is much more compact, has fewer parts, and is capable of successful operation over a greater bandwidth. Recent advances in alignment procedures [5] and iris design have largely eliminated the disadvantages of direct-coupled-cavity filters, and therefore it is believed that their use will predominate in the future.

In the case of quarter-wave-coupled waveguide-cavity filters, Mumford's design method is a convenient one and sufficiently accurate for bandwidths up to a few per cent [2]. His formulas are based on a lumped-constant filter prototype having the maximally flat response, but they may be extended easily to apply to other responses, such as the Tchebycheff equal-ripple response. However, a study of the literature on direct-coupled waveguide filters has revealed three published design methods [3, 8, 9], each of which gives different results from the others. In order to evaluate these methods, and to determine the extent of the approximations involved, a fourth approach mentioned above has been used. The result is a still different set of design relationships. This new analysis is given in this paper and its accuracy is compared with the other three. For extremely narrow bandwidths, all of the methods are good, but for bandwidths exceeding approximately 1 per cent the new method is the most precise, and it may be applied successfully to far greater bandwidths.

The analysis in this paper relies on the fact that the coupling elements can be made into perfect impedance inverters, or "quarter-wave" transformers, by the addition of negative elements in lumped-constant circuits, or of short negative lengths of line in waveguide and transmission-line circuits. These elements or line lengths may then be absorbed into the resonators. In this manner an exact design is achieved in the neighborhood of resonance. Because the "quarter-wave" quality of the impedance inverters is a broad-band one, good accuracy is maintained over a wide range. However, the variation of the coupling reactances, or susceptances, with frequency causes the response of the filter to be unsymmetrical, although a close approximation of this effect may be taken into account by means of the frequency-transformation formulas given in this paper.

## Design Formulas

### Low-Pass Prototype

The design formulas for the various types of direct-coupled-resonator filters considered in this paper are based on the low-pass filter prototype shown in Fig. 2. General formulas are given in the figure for the element values that will yield either the maximally flat or the Tchebycheff (equal ripple) insertion-loss response [12, 13]. For convenience, the following conditions are assumed: 1) the pass band edge occurs at $\omega' = 1$ (*i.e.*, at $f' = 1/2\pi$), 2) the left-hand load resistance is one ohm, and 3) the first element, $g_1$ is a shunt capacitance. Thus the elements of odd order $g_1, g_3, g_5, \cdots$, are shunt capacitances, and their values given by the formulas are in farads, while the elements of even order $g_2, g_4, g_6, \cdots$, are series inductances, and are given in henries. The last element, $g_n$, is a shunt capacitance if $n$ is odd, and a series inductance if $n$ is even. Inspection of the formulas shows that the right-hand resistance, $r$, is one ohm for all cases considered, except that of Tchebycheff response with $n$ even. Also, with the exception of that
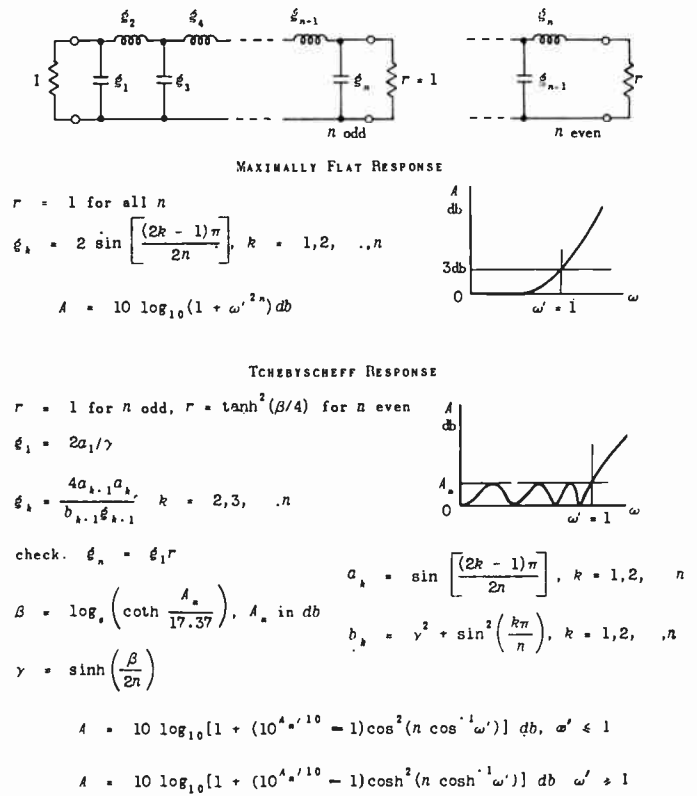


Fig. 2—Prototype low-pass filter and its design equations for maximally flat and Tchebycheff response.

case, the set of element values $g_1, g_2, g_3, \cdots, g_n$ is symmetrical. All needed quantities are defined in the formulas or response curves of the figure. The insertion-loss functions are also given in Fig. 2 for both maximally flat and Tchebycheff response.

### Lumped-Constant Filter Formulas

Fig. 3 gives the design relationships for lumped-constant coupled-resonant-circuit filters for both capacitive and inductive coupling. The derivation of the formulas is outlined later on. The load resistances may have any value, and may be equal or unequal to each other, as desired. The resonant-circuit elements $L_{rk}$, $C_{rk}$, may be selected to have any convenient sizes as long as they are large enough that the actual inductances and capacitances in the composite filter circuit are all positive. The formulas for the individual elements utilize the values $g_1, g_2, g_3, \cdots, g_n$ of the low-pass prototype filter, which are computed from the formulas in Fig. 2 for the particular insertion-loss response desired.

The insertion-loss curves sketched in Fig. 3 show the relationships between the band-pass and prototype responses. The band-pass response is obtained from the prototype response by transforming the frequency scale such that $f_0$ of the band-pass filter corresponds to $f' = 0$ of the low-pass filter, and the band edges $f_1$ and $f_2$ of the band-pass filter correspond to the band edge $f_1'$ of the prototype. In the design equations, $f_1'$ should be set equal to $1/2\pi$ if it is desired to have the insertion loss at
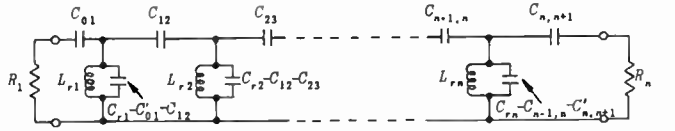
$$L_{rk}C_{rk} = \frac{1}{\omega_o^2}$$

$$C_{01} = \frac{1}{\omega_o}\left(\frac{w'C_{r1}/R_1\mathcal{E}_1}{1 - w'C_{r1}R_1/\mathcal{E}_1}\right)^{\frac{1}{2}}$$

$$C_{n,n+1} = \frac{1}{\omega_o}\left(\frac{w'C_{rn}r/R_n\mathcal{E}_n}{1 - w'C_{rn}R_n/\mathcal{E}_n}\right)^{\frac{1}{2}}$$
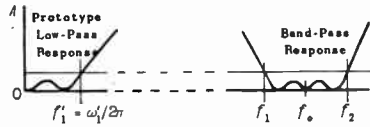
$$C_{jk} = w'\left(\frac{C_{rj}C_{rk}}{\mathcal{E}_j\mathcal{E}_k}\right)^{\frac{1}{2}} \text{ for } C_{12} \text{ to } C_{n-1,n}$$

$$C'_{01} = \frac{C_{01}}{1 + \omega_o^2 C_{01}^2 R_1^2}$$

$$C'_{n,n+1} = \frac{C_{n,n+1}}{1 + \omega_o^2 C_{n,n+1}^2 R_n^2}$$

$\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ = prototype elements in farads and henries from Fig. 2.

$r$ = right-hand load resistance in Fig. 2.

Note: The filter circuit is symmetrical if $R_1 = R_n$ and $C_{rk} = C_{r,n+1-k}$.

$$f_o = f_1 + f_2 - [(f_2 - f_1)^2 + f_1 f_2]^{\frac{1}{2}}$$

$$w' = \left(\frac{f_o}{f_1} - \frac{f_o}{f_2}\right)\left(\frac{f_o}{f_1}\right)$$

For $f_2/f_1 < 0.05$:

$$f_o \approx \sqrt{f_1 f_2}$$

$$w' \approx \frac{f_2 - f_1}{f_1}$$

**(a)**



$$\omega_o^2 L_{rk} C_{rk} = 1 \text{ for } k = 2 \text{ to } n-1$$

$$\omega_o^2 L_{r1} C_{r1} = 1 + \frac{w'L_0}{\mathcal{E}_1 R_1}$$

$$\omega_o^2 L_{rn} C_{rn} = 1 + \frac{w'L_{n+1}r}{\mathcal{E}_n R_n}$$

$$\omega_o M_{jk} = \pm w'\sqrt{\frac{L_{rj}L_{rk}}{\mathcal{E}_j\mathcal{E}_k}} \text{ for } M_{12} \text{ to } M_{n-1,n}$$

$$\omega_o M_{01} = \pm\sqrt{\frac{w'(R_1^2 + \omega_o^2 L_0^2)}{\omega_o^2 C_{r1}\mathcal{E}_1 R_1}}$$

$$\omega_o M_{n,n+1} = \pm\sqrt{\frac{w'r(R_n^2 + \omega_o^2 L_{n+1}^2)}{\omega_o^2 C_{rn}\mathcal{E}_n R_n}}$$

$\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ = prototype elements in farads and henries from Fig. 2.

$r$ = right-hand load resistance in Fig. 2.

$$f_o = f_1 + f_2 - [(f_2 - f_1)^2 + f_1 f_2]^{\frac{1}{2}}$$

$$w' = \left(\frac{f_o}{f_1} - \frac{f_o}{f_2}\right)\left(\frac{f_o}{f_1}\right)$$

For $f_2/f_1 < 0.05$:

$$f_o \approx \sqrt{f_1 f_2}$$

$$w' \approx \frac{f_2 - f_1}{f_1}$$

**(b)**

Fig. 3—(a) Design formulas for capacitive-coupled lumped-constant filter. (b) Design formulas for inductive-coupled lumped-constant filter.

the pass-band edge be 3 db in the maximally flat case or be equal to the ripple level in the Tchebycheff case. However, an $f_1'$ corresponding to any other pass band-edge insertion loss may be used if desired. The insertion-loss curves of the band-pass filter and the low-pass prototype vary in the same manner as a function of $|f - f_0|$ and $f'$. For narrow bandwidth, the band-pass insertion loss may be computed from the formula for the low-pass filter by replacing $f'/f_1'$ by $2|f - f_0|/(f_2 - f_1)$. For bandwidths of more than a few per cent, if one assumes that the coupling reactances do not vary with frequency, a better transformation requires that $f'/f_1'$ be replaced by $|f/f_0 - f_0/f|/(f_2/f_0 - f_0/f_2)$, where $f_0$ is related to $f_1$ and $f_2$ by $f_0 = \sqrt{f_1 f_2}$. However, the coupling reactances necessarily vary with frequency, and for bandwidths of more than a few per cent a superior transformation for the filter circuits of Fig. 3 is to replace $f'/f_1'$ by:

$$\frac{f'}{f_1'} = \left|2 - \frac{f_0}{f} - \frac{1}{2 - \frac{f_0}{f}}\right| \Bigg/ \left[2 - \frac{f_0}{f_2} - \frac{1}{2 - \frac{f_0}{f_2}}\right] \quad (1)$$

where $f_0$ is now related to $f_1$ and $f_2$ by

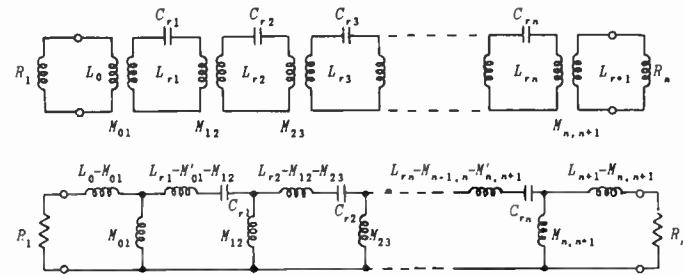$$f_0 = f_1 + f_2 - \sqrt{(f_2 - f_1)^2 + f_1 f_2}. \quad (2)$$

For constant coupling reactance, the proper formula for $w'$ in Fig. 3 would be $w' = (f_2 - f_1)/f_1'$. However, a close approximation of the effect of reactance variation with frequency is obtained when the following relationship for $w'$ is used:

$$w' = \left(\frac{f_0}{f_1} - \frac{f_0}{f_2}\right)\frac{f_0}{f_1'}. \quad (3)$$

In order to evaluate the accuracy of the design formulas for a moderately wide bandwidth, a six-resonant-circuit lumped-constant filter was designed and its exact response computed. The result is compared to the prototype function in Fig. 4. The response was designed to be maximally flat with 3-db points $\omega_2$ and $\omega_1$ in the ratio 1.20:1. The agreement between the curves is seen to be very good, despite the fact that the response is highly unsymmetrical on a linear frequency scale.

### Waveguide Filter Formulas

Fig. 5 gives the formulas for direct-coupled waveguide filters. The electrical lengths, $\phi_1$, of the cavities and the normalized inductive reactances, $X_{i,i+1}$, of the irises are computed in terms of the prototype elements $g_1$, $g_2$, $g_3$, $\cdots$, and the desired guide-wavelength pass band limits $\lambda_{g1}$ and $\lambda_{g2}$. As in the case of the lumped-constant filters, the formulas take into account the variation of the inductive coupling reactances with frequency. The iris reactance is assumed to vary in inverse proportion to guide wavelength (this is an excellent approximation for inductive windows or apertures in thin walls and for inductive posts [14]). If the coupling reactances were
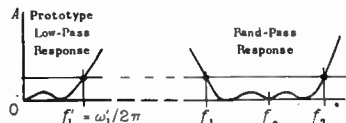
Fig. 4—Insertion loss of lumped-constant filter with six resonant circuits designed for maximally flat response.



Fig. 5—Formulas for inductive-iris-coupled waveguide filters.

As in the lumped-constant case, $\omega_1'$ should be set equal to unity if the pass band-edge insertion loss is desired to be 3 db for maximally flat response or to be the equal-ripple level for Tchebycheff response. Otherwise, $\omega_1'$ should be given the value that corresponds to the desired edge insertion loss in the low-pass prototype function.

The formulas of Fig. 5 differ from other formulas for direct-coupled filters in a number of important respects. First, all other published design formulas are based on frequency-independent coupling reactances, which are physically unrealizable. For bandwidths of more than a few per cent, this leads to errors in the pass band limits and in the steepness of cutoff. Aside from this, Southworth's formula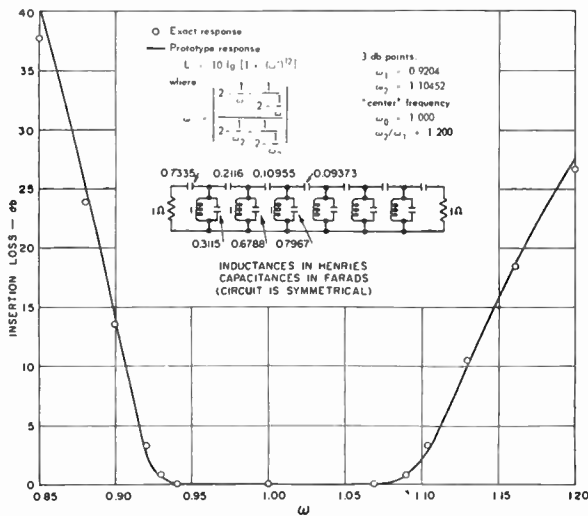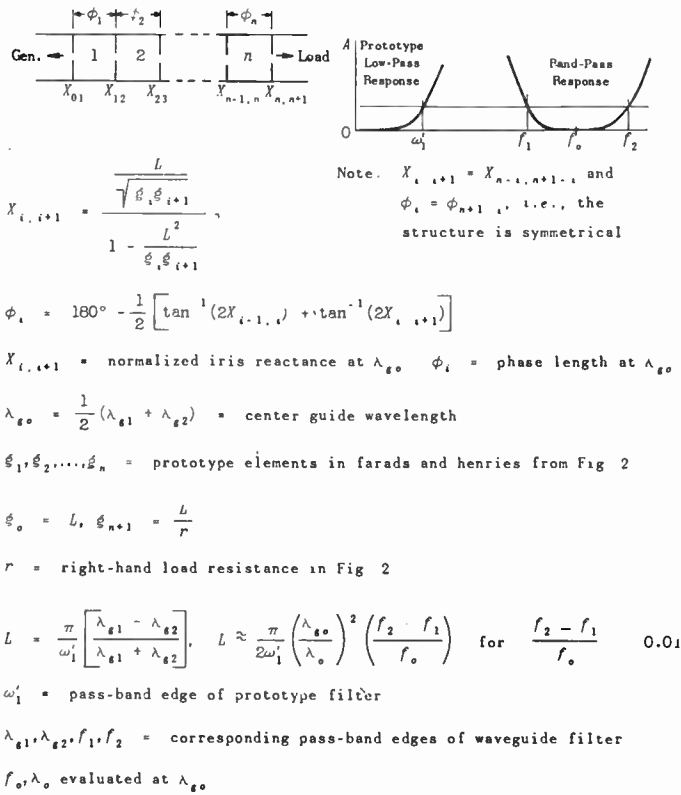 [8] for $X_{i,i+1}$ differs from that of Fig. 5. If the denominator $(1 - L^2/g_i g_{i+1})$ of the relation for $X_{i,i+1}$ in Fig. 5 is replaced by unity, the result is equivalent to Southworth's. However, the omission of the term $L^2/g_i g_{i+1}$ leads to appreciable error in a design for even a 1 per cent bandwidth, although for extremely narrow bandwidths the discrepancy becomes negligible. Fano and Lawson's analysis [3] has been found to contain typographical errors. When the errors are corrected, the result is the same as Southworth's, except for a minor difference in the formula for $\phi_i$. Riblet's formulas [9] are in a completely different form from those of Fig. 5, and hence a direct comparison cannot be made.

Fig. 6 shows a comparison of the exact vswr of filters designed by various methods to have the same bandwidth and the same response function. Perfect agreement occurs between the maximally flat prototype function and the design of this paper for the assumed bandwidth of 3 per cent. Equally good agreement was obtained on an insertion-loss scale up to at least 50 db [Fig. 6]. Although Southworth's design deviates considerably from the prototype function at low vswr values, on an insertion-loss scale it would appear to agree much better. Riblet's design method gives its best results at low vswr's in the example of Fig. 6. At the 3-db points the bandwidth error is about 8 per cent. For greater bandwidths the error in design would increase.

The vswr and insertion loss of a six-cavity waveguide filter designed by the new method for Tchebycheff response is shown in Fig. 7. The agreement between the exact computed response and the prototype function is excellent for the design bandwidth of 2 per cent.

The new method was also tested for moderately wide bandwidths. The results for a six-cavity filter designed for maximally flat response and bandwidth of about 20 per cent are shown in Fig. 8, and for Tchebycheff response and bandwidth of about 10 per cent in Fig. 9. In both cases the agreement is good, although minor deviations from the prototype functions are observed. The pass band vswr, which gives the most exacting test, shows the most discrepancy, but the maximum pass band vswr is very low and is certainly suitable for most practical applications. The dissymmetry of these responses caused by frequency variation of the coupling

independent of frequency, the response of the filter would be expected theoretically to be symmetrical on a $1/\lambda_g$ scale. However, the guide-wavelength dependence of the coupling reactances has the very interesting effect of making the filter response symmetrical on a $\lambda_g$ scale. The predicted insertion-loss response may be computed from the formulas given in Fig. 2 for the low-pass prototype by substituting

$$\frac{\omega'}{\omega_1'} = 2\left(\frac{\lambda_{g0} - \lambda_g}{\lambda_{g1} - \lambda_{g2}}\right). \qquad (4)$$

Fig. 6—VSWR and insertion loss of direct-coupled six-cavity wave-guide filter designed for maximally flat response—bandwidth 3 per cent.



Fig. 7—Insertion loss and vswr of direct-coupled six-cavity wave-guide filter designed for Tchebycheff response—bandwidth 2 per cent.



Fig. 8—Insertion loss and vswr of direct-coupled six-cavity wave-guide filter designed for maximally flat response—bandwidth 20 per cent.
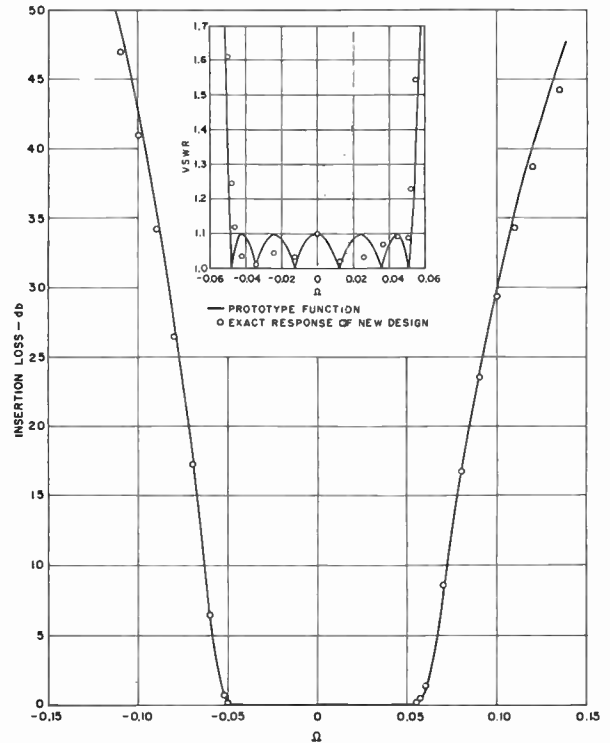


Fig. 9—Insertion loss and vswr of direct-coupled six-cavity wave-guide filter designed for Tchebycheff response—bandwidth 10 per cent.

reactances is quite large, but is compensated for very well by the frequency-scale transformation that was introduced into the design formulas.

*Strip-Line Filter Formulas*

Strip-line direct-coupled filters can be made conveniently with either series-capacitive or shunt-inductive coupling between resonant line lengths. The formulas of Fig. 5 for waveguide filters apply in the case of inductive coupling if free-space wavelength $\lambda$ is substituted everywhere for guide wavelength $\lambda_g$. The shunt inductances might consist of metal posts or strips joining the strip line to the two ground planes. However, a usually more convenient construction method uses capacitive coupling, where gaps in the strip line form the series capaci-

tances. The filter layout and design formulas are shown in Fig. 10. The filter circuit is the dual of the circuit containing inductive coupling elements, and the formulas were written directly from Fig. 5 by means of the usual duality relationships, and the substitution of free-space wavelength for guide wavelength.
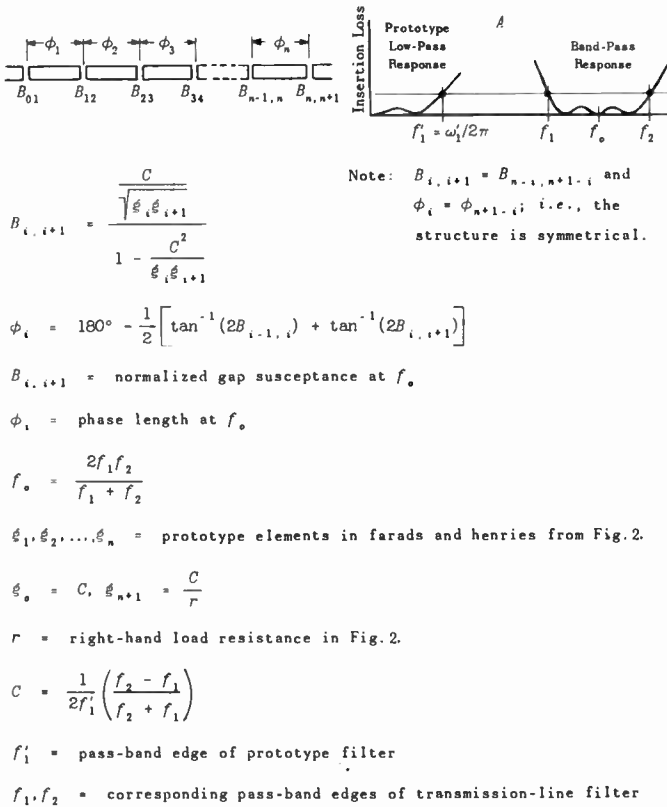


Fig. 10—Formulas for capacitive-gap-coupled transmission-line filter.

The calculation of the coupling values and line lengths in terms of the low-pass prototype elements is done the same way as for the waveguide filter. The gap dimensions to yield the required series capacitances may be obtained from experimental curves or a theoretical formula to be found in several references [10, 11, 15, 16]. If the gap is small compared to the ground-plane spacing the resonant line lengths should be measured from the center of the gap. The line-length error in this procedure is proportional to the square of the ratio of gap to ground-plane spacing, and is usually negligible. However a line length correction for the presence of the gap may be made from available formulas and data [15, 16].

The theoretical response curves of Figs. 6 to 9 apply exactly to strip-line direct-coupled filters if guide wavelength is replaced by free-space wavelength. Thus, as in the waveguide case, the response of either series-capacitive or shunt-inductive coupled strip-line filters is symmetrical on a wavelength scale rather than on a frequency scale. Therefore, to compute the insertion-loss response from that of the low-pass prototype given in Fig. 2, one should use (4) with $\lambda$ replacing $\lambda_g$.

## DERIVATION OF FORMULAS

### Basic Transformations Applicable to All Types

Fig. 11(a) shows the lumped-constant prototype and Fig. 11(b) the inverse-arm band-pass filter derived from it by the frequency transformation: $\omega' \propto (\omega/\omega_0 - \omega_0/\omega)$. In both cases the first arm is taken always to be a shunt arm. Impedance inverting transformers[1] are next inserted between the successive arms of the filter to obtain the circuit of Fig. 11(c), in which all resonant arms are series arms. The impedance-inversion property that makes this last transformation possible may be expressed simply by the well-known relationships for a quarter wavelength of transmission line: $Z_1 = K^2 Y_2$ and
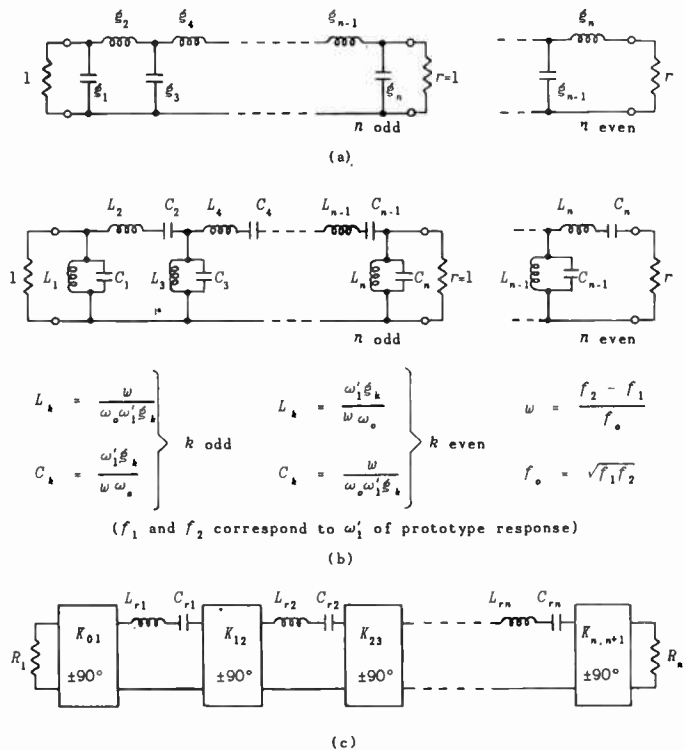


Fig. 11—Transformations from low-pass prototype to band-pass filter containing quarter-wave inverters.

$Y_1 = Z_2/K^2$, where $K$ is the characteristic impedance (or image impedance) of the quarter-wave line (inverter), the subscript "2" denotes load impedance or admittance connected to the inverter, and the subscript "1," input impedance or admittance. Thus, for example, a shunt capacitance $C$ viewed through the inverter appears to be a series inductance of value $L = K^2 C$, while a series inductance $L$ appears to be a shunt capacitance of value $C = L/K^2$. The detailed steps using this concept are carried out below and lead to the desired relationships between the elements $L_k$, $C_k$ in Fig. 11(b) and $L_{rk}$, $C_{rk}$ in Fig. 11(c).

[1] Impedance inverters have had frequent use in network transformation. For example, see [1] and [15].

Consider the portion of the inverse-arm band-pass circuit shown in Fig. 12(a). The elements $L_2$ and $C_2$ already constitute a series arm as desired, but at an improper impedance level. This may be corrected as indicated in Fig. 12(b) by multiplying all inductances in the filter (and the load resistances) by $L_{r2}/L_2$, and all capacitances by $L_2/L_{r2}$. Next, an inversion network is introduced in Fig. 12(c) to transform the shunt arm of Fig. 12(b) into a series arm. In order for the performance of the filter to be the same in both cases, $Y_3$ in Fig. 12(b) and (c) must be identical. Hence it is necessary that

$$j\left(\omega C_3 - \frac{1}{\omega L_3}\right)\frac{L_2}{L_{r2}} + Y_4$$

$$= \frac{1}{K_{23}^2}\left\{j\left(\omega L_{r3} - \frac{1}{\omega C_{r3}}\right) + Z_4'\right\}.$$
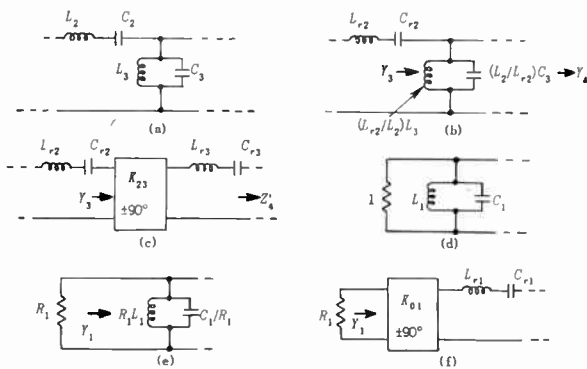
Fig. 12—Detailed transformations used in analysis.

After identifying the quantities that have the same dependence on $\omega$, one may write

$$K_{23} = \sqrt{\frac{L_{r2}L_{r3}}{L_2 C_3}}, \qquad L_{r2}C_{r2} = L_{r3}C_{r3} = \frac{1}{\omega_0^2}.$$

With the aid of the relations in Fig. 11(b), we may write, in general for arms $k$ and $k+1$,

$$K_{k,k+1} = \frac{\omega_0 w}{\omega_1'}\sqrt{\frac{L_{rk}L_{r,k+1}}{g_k g_{k+1}}}, \quad L_{rk}C_{rk} = \frac{1}{\omega_0^2}. \quad (5)$$

The first and last $K$ in the filter may be computed in a similar manner. Thus, Fig. 12(d) shows the input load resistance and first shunt arm of the inverse-arm band-pass filter. In Fig. 12(e), the impedance level has been changed from one ohm to $R_1$ ohms. In Fig. 12(f), an inversion network is introduced to transform the shunt arm to a series arm. In order for $Y_1$ to be the same in Fig. 12(e) and (f), $K_{01}$ must be chosen so that

$$K_{01} = \sqrt{\frac{L_{r1}R_1}{C_1}} = \sqrt{\frac{\omega_0 w}{\omega_1'}\cdot\frac{L_{r1}R_1}{g_1}}. \quad (6)$$

Similarly, the last inversion network is given by

$$K_{n,n+1} = \sqrt{\frac{\omega_0 w}{\omega_1'}\frac{L_{rn}R_n r}{g_n}} \quad (7)$$

for $n$ odd or even.

*Impedance-Inversion Networks*

A number of networks that have the inversion properties of a quarter-wave transformer over a broad bandwidth are shown in Fig. 13. The characteristic impedance, $K$, and phase shift, $\beta$, of each network may be derived in various ways. For example, in Fig. 13(a) the short- and open-circuit input impedances of half the network are $Z_{sc} = -j/\omega C$, $Z_{oc} = j/\omega C$, and through the use of well-known relationships

$$K = \sqrt{Z_{sc}Z_{oc}} = \frac{1}{\omega C},$$

$$\beta = 2\tan^{-1}\pm\left(\frac{-Z_{sc}}{Z_{oc}}\right)^{1/2} = \pm 90°.$$
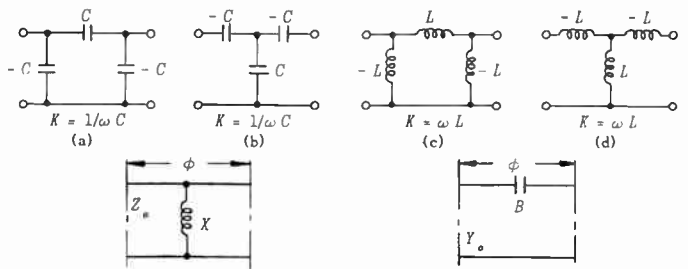
Fig. 13—Broad-band impedance-inversion networks.

For the network of Fig. 13(a) the negative sign for $\beta$ is correct, but in this application the particular sign is unimportant. The important point to notice is that $\beta$ is not frequency dependent, and hence the impedance-inversion property of Fig. 13(a), and also of Fig. 13(b), (c), and (d), holds at all frequencies, while the characteristic impedance has the same variation with frequency as does the kind of reactance in the network. The most suitable inversion network to choose for a given filter is one whose negative elements can be absorbed into the resonant circuits so that all element values of the composite filter are positive.

Networks in TEM transmission line and waveguide may also be devised that provide impedance inversion over very wide bandwidths. Cases utilizing a shunt inductive reactance and a series capacitive reactance appear in Fig. 13(e) and (f). For example, in the case of the

shunt inductive reactance, $X = \omega L$, negative electrical lengths of line are required on each side of the reactance. The formulas in the figure may be derived conveniently from the short- and open-circuit impedances of half the structure:

$$\beta = 2 \tan^{-1} \pm \sqrt{\frac{-X_{sc}}{X_{oc}}}$$

$$= 2 \tan^{-1} \pm \sqrt{\frac{-Z_0 \tan \dfrac{\phi}{2}}{Z_0 \tan \left\{ \dfrac{\phi}{2} + \tan^{-1}\left(\dfrac{2\omega L}{Z_0}\right)\right\}}} \quad (8)$$

$$K = Z_0 \sqrt{- \tan\left(\frac{\phi}{2}\right) \tan\left[\frac{\phi}{2} + \tan^{-1}\left(\frac{2\omega L}{Z_0}\right)\right]}. \quad (9)$$

If we set $\beta = -90°$, it is necessary at $\omega_0$ that

$$\phi = - \tan^{-1}\left(\frac{2\omega_0 L}{Z_0}\right) \quad \text{and} \quad K = Z_0 \tan \frac{\phi}{2}. \quad (10)$$

The numerator and denominator of (8) are both almost exactly proportional to $\phi$, and this causes the frequency variation of $\beta$ from $-90°$ to be slight enough to be neglected. Also, $K$ is very closely proportional to $\phi$.

For use in design work, a formula for $X = \omega_0 L$ in terms of $K$ is needed. This is obtained as follows with the aid of a trigonometric identity.

$$\frac{X}{Z_0} = \frac{1}{2} \tan(-\phi) = \frac{1}{2} \tan\left(2 \tan^{-1}\frac{K}{Z_0}\right)$$

$$= \frac{K/Z_0}{1 - (K/Z_0)^2}. \quad (11)$$

### Equivalent Circuit of a Half-Wavelength Resonator

In the case of TEM-transmission-line or waveguide direct-coupled filters, the equivalent circuit of a half-wavelength of line is required. This is given exactly in Fig. 14(b). The three reactive elements provide the frequency dependence of the circuit, while the ideal transformer provides the phase reversal of a half-wave line. Since this reversal plays no part in the filter performance, it is neglected in the approximate equivalent circuit of Fig. 14(c).

When the half-wave resonators are used in a waveguide filter, the shunt arms of Fig. 14(c) are connected almost directly in parallel with the large shunt susceptances of the inversion networks. In Fig. 14(b), it is seen that $B/Y_0$ is much less than unity even for moderately wide bandwidths, while the normalized susceptance of the coupling irises is much larger than unity. Hence the shunt arms may be neglected as in Fig. 14(b), where only the series-resonant arm remains. Reference to the

filter schematic of Fig. 11(c) shows that the last half-wave-line equivalent may be used as the resonators, and the shunt reactance and associated negative line length of Fig. 13(e) may serve as the inversion networks. First, however, an expression for $L_{r1}$ in terms of waveguide parameters must be obtained.
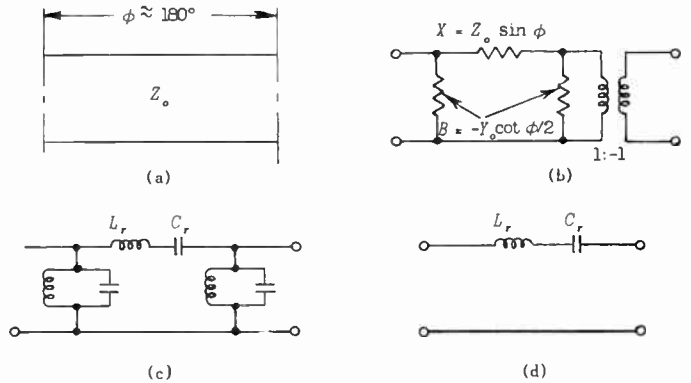


Fig. 14—Equivalent circuit of half-wavelength resonator.

Fig. 14(b) gives the reactance of the series arm as $X = Z_0 \sin \phi$. In the vicinity of $\phi = \pi$, this may be written

$$X = Z_0(\phi - \pi) = \pi Z_0\left(\frac{\lambda_{g0}}{\lambda_g} - 1\right). \quad (12)$$

In terms of the equivalent-circuit elements,

$$X = \omega L_r - \frac{1}{\omega C_r}. \quad (13)$$

A relationship between $L_r$ and $\lambda_g$ is most easily obtained by equating $dX/d\omega\big|_{\omega = \omega_0}$, computed from (12), to the same quantity from (13). Thus,

$$\frac{dX}{d\omega}\bigg|_{\omega = \omega_0} = 2L_r = Z_0 \frac{d\phi}{d\omega}\bigg|_{\omega = \omega_0} \approx \pi Z_0 \frac{\lambda_{g0}/\lambda_{g2} - \lambda_{g0}/\lambda_{g1}}{\omega_2 - \omega_1}$$

and therefore

$$L_r = \frac{\pi Z_0}{2\omega_0 w}\left(\frac{\lambda_{g0}}{\lambda_{g2}} - \frac{\lambda_{g0}}{\lambda_{g1}}\right) \quad (14)$$

where $w = (\omega_2 - \omega_1)/\omega_0$ and $\omega_0 = (\omega_1 + \omega_2)/2$. A combination of (5), (6), (7), and (14) then gives

$$\frac{K_{k,k+1}}{Z_0} = \frac{L}{\sqrt{g_k g_{k+1}}} \quad (15)$$

where, for convenience,

$$L = \frac{\pi}{2\omega_1'}\left(\frac{\lambda_{g0}}{\lambda_{g2}'} - \frac{\lambda_{g0}}{\lambda_{g1}'}\right) \quad (16)$$

$$g_0 = L, \quad g_{n+1} = \frac{L}{r}, \quad R_1 = R_n = Z_0 \qquad (17)$$

$$\frac{1}{\lambda_{g0}} = \frac{1}{2}\left(\frac{1}{\lambda_{g1}'} + \frac{1}{\lambda_{g2}'}\right). \qquad (18)$$

The prime marks placed on $\lambda_{g1}'$ and $\lambda_{g2}'$ indicate that these are the pass band limits when the inversion-network characteristic impedances $K_{k,k+1}$ are assumed to be constant with frequency. However, $K_{k,k+1}$ varies with frequency, causing the filter response to be an unsymmetrical function of $\lambda_{g0}/\lambda_g$, and therefore the unprimed quantities $\lambda_{g1}$ and $\lambda_{g2}$ will henceforth be reserved for the true pass band limits. A simple relationship between $\lambda_g'$ and $\lambda_g$ will now be obtained.

*Frequency-Scale Correction for Waveguide Filter*

The bandwidth of a multiple-resonator direct-coupled filter is dependent mainly upon the size of the internal coupling elements, and only slightly upon the input and output coupling elements. Eqs. (15) and (16) show that the relative bandwidth expressed in reciprocal guide wavelength is proportional to $K/Z_0$ of the internal inverting networks, while (11) shows that $K/Z_0$ is approximately proportional to $X/Z_0$. A shunt inductive iris or post in waveguide has a reactance variation very nearly proportional to $\lambda_{g0}/\lambda_g$. Hence, to a good approximation, $\lambda_{g0}/\lambda_g - 1 \propto K/Z_0 \propto \lambda_{g0}/\lambda_g$. For very narrow bandwidths, $\lambda_g' \to \lambda_g$, and this permits the proportionality relationship to be completed as follows:

$$\left(\frac{\lambda_{g0}}{\lambda_g}\right) - 1 = \left(\frac{\lambda_{g0}}{\lambda_g'} - 1\right)\frac{\lambda_{g0}}{\lambda_g}. \qquad (19)$$

Through the use of (18) and (19), $\lambda_{g0}$ is related to the pass band limits by $\lambda_{g0} = (\lambda_{g1} + \lambda_{g2})/2$. When this is substituted in (16), the formula for $L$ given in Fig. 5 results.

Upon joining the series of half-wave lines and inversion structures, the electrical length of each resonator is 180° plus the $\phi/2$ lengths associated with the adjoining coupling reactances. With the aid of (10), the electrical length of the $i$th resonator at $\omega_0$ is as given in Fig. 5.

Thus the various design equations of Fig. 5 for the waveguide direct-coupled filter have been verified. Also, Fig. 10 for the capacitive-gap-coupled transmission-line filter is seen to be verified when it is considered that $\lambda_g$ in the waveguide filter formulas should be replaced by $\lambda$, and that this circuit is the dual of the waveguide circuit.

*Lumped-Constant Filter Design*

The resonant elements $L_r$, $C_r$ in the first circuit of Fig. 3(b) are in series resonant arms. Therefore, the discussion pertaining to Fig. 11 applies, and the inversion-network characteristic impedances are given by (6). The most suitable inversion network is that of

Fig. 13(d), since its negative inductance elements may be subtracted from $L_r$. However, this cannot be done with the first and last inversion network of the filter, since a negative inductance would be required at the input and output of the filter. This has been avoided by making $L_0 - M_{01} > 0$ and by choosing the value of $M_{01}$ such that the proper load resistance $R'$ (Fig. 15) will be inserted in series with the first resonant circuit. The equivalent series-inductance component $M_{01}'$ then becomes a part of the resonant circuit, as indicated in Fig. 3(b).
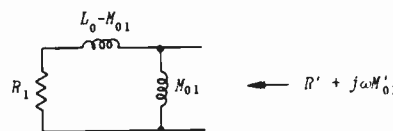


Fig. 15—Equivalent load impedance of inductive-coupled lumped-constant filter.

The capacitance-coupled lumped-constant filter of Fig. 3(a) is the dual of the circuit of Fig. 3(b) (if $L_0 = |M_{01}|$), and the formulas for one case may be obtained readily from those of the other. It will be noticed that the formulas for $C_{01}$ and $M_{01}$ have a somewhat different appearance, but this difference disappears when $L_0 - |M_{01}| = 0$, and the equation giving $M_{01}$ is solved explicitly for $M_{01}$. However, when mutual inductance is used in a filter design, a primary inductance $L_0 > |M_{01}|$ must be provided, and the spacing to the secondary should then be adjusted to yield the necessary value of $M_{01}$.

The frequency-scale correction for the lumped-constant filter to allow for the frequency dependence of the inversion-network characteristic impedance is derived in the same way as for the waveguide filter. The resulting frequency relationships are given in (1), (2), and (3).

*Generalized Case of Coupled Resonators—Coupling Coefficient*

In the analysis of lumped-constant coupled-resonant-circuit filters, it has been customary to define a coupling coefficient $k$ between adjacent resonant circuits. For narrow bandwidth, the coupling coefficient between resonators $i$ and $i+1$ is related to the pass band frequencies and prototype elements by

$$k_{i,i+1} = \frac{1}{\omega_1'\sqrt{g_i g_{i+1}}}\cdot\left(\frac{f_2 - f_1}{f_0}\right).$$

This parameter is also useful in the design of narrow-band filters containing resonant cavities of any shape. The coupling element (aperture, loop, probe, etc.) to yield the desired value of coupling coefficient may be determined by calculation or by an experimental method given by Dishal [5].

## BIBLIOGRAPHY

[1] Fano, R. M., and Lawson, A. W. "Microwave Filters Using Quarter-Wave Couplings," PROCEEDINGS OF THE IRE, Vol. 35 (November, 1947), pp. 1318–1323.
[2] Mumford, W. W. "Maximally Flat Filters in Waveguide," *Bell System Technical Journal*, Vol. 27 (October, 1948), pp. 648–714.
[3] Ragan, G. L. *Microwave Transmission Circuits*. New York: McGraw-Hill Book Company, Inc., 1948. Chapters 9 and 10 by R. M. Fano and A. W. Lawson.
[4] Dishal, M. "Design of Dissipative Band-Pass Filters Producing Desired Exact Amplitude-Frequency Characteristics," PROCEEDINGS OF THE IRE, Vol. 37 (September, 1949), pp. 1050–1069.
[5] Dishal, M. "Alignment and Adjustment of Synchronously Tuned Multiple-Resonant-Circuit Filters," PROCEEDINGS OF THE IRE, Vol. 39 (November, 1951), pp. 1448–1455.
[6] Dishal, M. "Two New Equations for the Design of Filters," *Electrical Communication*, Vol. 40 (December, 1953), pp. 324–337.
[7] Dishal, M. "Concerning the Minimum Number of Resonators and Minimum Unloaded-Resonator-$Q$ Needed in Filters," *Electrical Communication*, Vol. 31 (December, 1954), pp. 257–277. Also IRE TRANSACTIONS, PGVC-3 (June, 1953), pp. 85–117.
[8] Southworth, G. C. *Principles and Applications of Waveguide Transmission*. New York: D. Van Nostrand Company, Inc., 1950, pp. 285–319.
[9] Riblet, H. J. "Synthesis of Narrow-Band Direct Coupled Filters," PROCEEDINGS OF THE IRE, Vol. 40 (October, 1952), pp. 1219–1223.
[10] Bradley, E. H., and White, D. R. "Band-Pass Filters Using Stripline Techniques," IRE TRANSACTIONS, Vol. MTT-3 (March, 1955), pp. 163–169.
[11] Bradley, E. H. "Design and Development of Strip-Line Filters," IRE TRANSACTIONS, Vol. MTT-4 (April, 1956), pp. 86–93.
[12] Belevitch, V. "Tchebyscheff Filters and Amplifier Networks," *Wireless Engineer*, Vol. 29 (April, 1952), pp. 106–110.
[13] Orchard, H. J. "Formulas for Ladder Filters," *Wireless Engineer*, Vol. 30 (January, 1953), pp. 3–5.
[14] Marcuvitz, N. *Waveguide Handbook*. (Massachusetts Institute of Technology, Radiation Laboratory Series, Vol. 10.) New York: McGraw-Hill Book Company, Inc., 1951, pp. 221, 227, 229, 238, 257, etc.
[15] Keen, H., and Sion, E. "Progress During November 1954." Contract AF 19 (604)-780, Airborne Instruments Laboratory, Mineola, N. Y. (December 1, 1954).
[16] Oliner, A. A. "Equivalent Circuits for Discontinuities in Balanced Strip Transmission Line," IRE TRANSACTIONS, Vol. MTT-3 (March, 1955), pp. 134–143.
[17] Wheeler, H. A. Wheeler Monographs, Vol. I (Wheeler Laboratories, Great Neck, L. I., N. Y.), 1953, pp. 3–54.

# Distributed-Parameter Variable Delay Lines Using Skewed Turns for Delay Equalization*

F. D. LEWIS†, SENIOR MEMBER, IRE, AND R. M. FRAZIER‡, MEMBER, IRE

*Summary*—Delay equalization of distributed-parameter delay lines is accomplished by a new method making use of skewed turns in the winding. A simple analysis of the basis for the use of this method of equalization is given, and the performance of experimental variable delay lines is discussed.

A brief survey of the artificial-line delay-equalization problem is given with a discussion of some of the alternative methods of solving the problem.

## INTRODUCTION

THE REALIZATION of continuously variable electromagnetic delay lines providing adequate bandwidth with constant time delay requires the adoption of certain design features which combine to make the task difficult. The requirement for continuous variation instead of step-wise adjustment means that the variable delay line is preferably made as a continuously wound distributed-parameter line with a sliding tap. In the continuous type of winding, it becomes difficult to introduce compensating network sections unless they can be put in as distributed elements. Also, single-layer constant-pitch coils wound with commonly used sizes of wire and form dimensions possess distributed constants which provide approximate self-compensation

of time delay vs frequency *only* in the relatively high impedance range, or when the time delay per unit length of the line is very low. It is the purpose of this paper to describe a new method of delay equalization making use of a skewed winding to obtain a constant time delay vs frequency. This compensation method offers several important advantages, including the ability to provide delay equalization for characteristic impedances down to 150 ohms, or lower, without introducing wavelength-sensitive discontinuities. In Fig. 1, the pulse and step responses of a 1-$\mu$sec delay line using this new method of compensation are shown compared to the responses of an equivalent line without compensation. A simple analysis of the basis for using this method of compensation shows that the method is theoretically valid, and that the measured characteristics agree qualitatively with the results predicted by the analysis.

## DELAY VARIATION WITH RESPECT TO FREQUENCY IN ARTIFICIAL LINES

### Review of Delay Equalization Problem

Much engineering effort has been expended on the design of delay networks for various purposes. Lumped parameter pulse-forming networks are representative of specialized types of delay networks wherein the delay time is arranged to be constant at the required frequen-
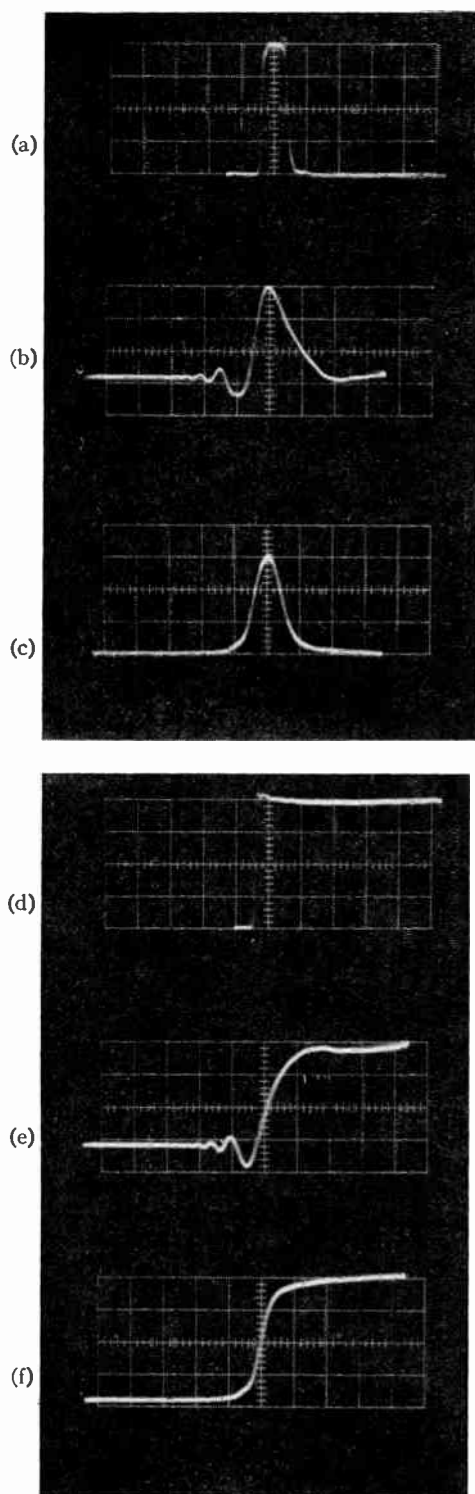
Fig. 1—Pulse and step responses of 1-μsec delay 500-ohm variable delay lines: (a) pulse input, (b) pulse out of uncompensated line, (c) pulse out of skewed-winding line, (d) step input, (e) step out of uncompensated line, (f) step out of skewed-winding line. Scope photos taken on Tektronix 541, 0.1-μsec/cm sweep.

cies in the pass band, but not necessarily so at all frequencies in the band covered.[1] Television transmission systems, in general, require delay networks with con-

stant time delay at all frequencies in the pass band.[2] Computers using pulse techniques frequently require wide-band distortion-free delay lines for equalization of time delay. In the two latter classes of applications, and in all pulse systems where the pulse shape must be transmitted relatively unaltered, constancy of time delay at all frequencies in the operating range of the delay line is of greatest importance. A brief review of the problem of delay equalization, and the solutions devised heretofore, is given below.

Various methods of filter design have been applied to lumped-parameter delay networks. However, these methods are not readily applicable to distributed-parameter continuously wound delay lines. The dissimilarity between these two types of delay networks is basic, and lies in the definition of lumped and distributed-parameter circuits. The current in a turn of a lumped inductance coil is assumed, by definition, to be in phase with the current in all the other turns of the same lumped inductance, or at most only a negligible phase difference is allowed. In a distributed-parameter delay line, the very time delay which is sought by the use of the delay line produces progressive phase shift between the currents in successive turns along the line at any given frequency. This phase shift then produces a reduction in the effective inductance (as defined in the Appendix) of the distributed winding as the frequency is increased.[3,4] This decrease in effective inductance results from the reduction of the in-phase component of the current in a given turn with respect to the current in a reference turn. The reduction in effective inductance produces a readily observable effect in a distributed-parameter delay line, namely, a roll-off of the time-delay-vs-frequency characteristic of the line. This effect is easily predictable since $T_d = \sqrt{LC}$ where $T_d =$ time delay per unit length and $L$ and $C$ are the inductance and capacitance, respectively, of the line.

What is desirable, then, in order to provide constant time delay vs frequency with a distributed-parameter delay line, is a method of compensation which will maintain the effective inductance at a constant value throughout the operating frequency range. Such a method should provide a constant time delay at any frequency within the pass band of the line.

### Capacitive Compensation Methods

Delay compensation by the use of floating patches has been described by Kallman.[4,5] This method is well suited to distributed-winding-type construction since the patches may be applied to the winding form or added to the winding after it is completed.

[1] G. N. Glasoe and J. V. Lebacqz, "Pulse Generators," vol. 5, M.I.T. Radiation Lab. Series, McGraw-Hill Book Co., New York, N. Y. (Part II, "The Line-Type Pulser"); 1948.

[2] A. H. Turner, "Artificial lines for video distribution and delay," *RCA Review*, vol. X, pp. 477–489; December, 1949.

[3] J. P. Blewett and J. H. Rubel, "Video delay lines," Proc. IRE, vol. 35, pp. 1580–1584; December, 1947.

[4] J. F. Blackburn, "Components Handbook," vol. 17, M.I.T. Radiation Lab. Series, McGraw-Hill Book Co., New York, N. Y. (ch. 6, "Electromagnetic Delay Lines," by H. E. Kallman); 1949.

[5] H. E. Kallman, "Equalized delay lines," Proc. IRE, vol. 34, pp. 646–657; September, 1946.

Two disadvantages arise from the use of patch compensation: The patches produce a wavelength-sensitive discontinuity in the delay-vs-frequency curve, and they also cause the effective impedance of the line to vary at some critical frequency associated with this wavelength sensitivity. This impedance difficulty could be anticipated since the decrease in effective inductance is compensated by an increase in effective distributed capacitance. Thus, a variation in capacitance which keeps the time delay constant results in a change in effective line impedance.

The wavelength-sensitive property of the patches was avoided by Blewett and Rubel[3] who applied a paint of high dielectric constant, composed of aluminum powder in polystyrene, to a distributed winding. This paint had the effect of increasing the natural longitudinal distributed capacitance of the winding, giving considerable improvement in the bandwidth over which the delay remained relatively constant and avoiding the difficulties associated with the size of the floating patches.

A further variation was described by Carley[6] making use of a bank winding to obtain a coil with adequate distributed capacitance to be approximately self-compensating over the range desired.

In view of the progressive phase change between the turns of the distributed winding, it is worth noting that all of these systems of compensation obtain constancy of delay time by effective longitudinal capacitance change as the frequency varies. As was mentioned above, an inductance-compensation scheme would be preferable.

### Lumped-Parameter Delay-Line Inductance Compensation Methods

Methods have been devised for keeping the effective inductance of lumped-element delay lines constant over a wide range by adjusting the mutual inductance between sections of the inductance on an individual coil basis. Pierce[7] obtained relatively constant time delay by this method. Gillen[8] obtained the required mutual inductances by winding the coils on individual forms bearing ground-capacitance strips in distributed form, and adjusting the positions of the coils. This method of compensation is effective out to the point at which the phase shift across the coil section becomes 180°, whereupon the phase and amplitude response of the delay line show a marked change.

Schwartz[9] obtains inductive compensation by adjustment of mutual inductance between coils wound on a single form, with alternating directions of winding, introducing the use of interposed short-circuited turns to control mutual coupling.

Control of the magnitude of the mutual inductance between sections of a lumped-constant delay line has been accomplished by the use of a suitably proportioned shielding enclosure.[10]

Golay[11] has made a theoretical study of a sectionalized delay line with mutual inductance between the various coils of the line, and has derived expressions for optimum compensation of the time-delay-vs-frequency characteristic below the cutoff frequency by the adjustment of sign and magnitude of the mutual inductances between sections. The mutual inductances are introduced as additional coils connected in series with the principal coil of the section, but placed on the winding form of the coil with which the coupling is desired.

These methods are not directly applicable to a continuously wound distributed-constant line as each of them requires sectionalizing the winding in some particular way in order to achieve control of the compensation of time delay vs frequency.

### Effect of Solid Shield on Cylindrical Distributed-Winding Delay Lines

Some control of the mutual inductance coupling between turns of a circular cross section uniform-pitch coil is exercised by applying a solid conducting tubular shield surrounding the winding.[12] The quantities thus controlled are limited to those associated with the magnitude and physical extent of the mutual inductances and capacitances, no control being possible over the sign of the mutual coupling solely by the addition of the shield. Nevertheless, comparatively large improvements can be effected in the constancy of time delay of a uniformly wound delay line by the use of such a shield if the losses introduced thereby are not important. Prache[12] gives data indicating that the time delay variation of an unshielded delay line with a distributed-winding coil can be changed from its original value of −5 per cent at 25 mc to a value more than +4 per cent from the low-frequency delay merely by installing a surrounding shield of the proper diameter. Obviously, a shield of proper size could be chosen to provide almost perfect compensation.

## A New Method of Inductance Compensation for Distributed-Parameter Delay Lines

### Description and Analysis of Skewed-Turn Inductance Compensation Method

To provide inductance compensation of a distributed-parameter delay line, some means should be devised to provide a corrective mutual inductance on a continuous, or turn-to-turn, basis. In the delay lines which are the subject of this paper, this has been accomplished by

[6] W. S. Carley, "Distributed-constant delay lines with characteristic impedances higher than 5000 ohms," 1953 IRE CONVENTION RECORD, part 5, pp. 71–80.

[7] G. W. Pierce, U. S. Patent 1,576,459; filed December 24, 1921.

[8] D. A. Gillen, "A simple method of phase compensation of video delay lines," IRE TRANS., vol. BTR-1, pp. 1–9; January, 1955.

[9] A. A. Schwartz, Jr., "Time-Delay Network," U. S. Patent 2,703,389; filed November 17, 1953.

[10] Blackburn, *op. cit.*, p. 215.

[11] M. J. E. Golay, "Corrected Delay Line," U. S. Patent 2,598,683; filed February 5, 1946.

[12] P. M. Prache, "La ligne à retard électromagnétique de structure uniforme," *Cables & Transmission*, 9ᵉ Année, pp. 112–143; April, 1955. (Published by Sotelec, 16 rue de la Baume, Paris VIII, France.)

skewing the turns of the winding with respect to the winding axis (see Fig. 2) at such an angle that the effective inductance remains nearly constant over a wide range of phase-shift per turn as shown in Fig. 3. The



(a)

(b)

(c)
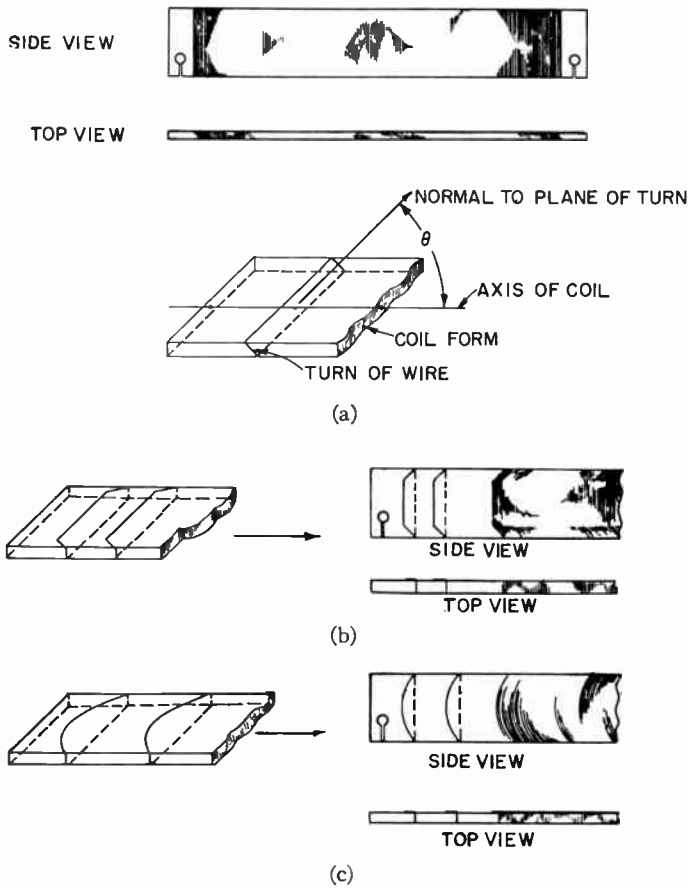
Fig. 2—Diagrams showing arrangement of skewed windings on forms of rectangular cross section: (a) rectangular turns, (b) equivalent trapezoidal turns, (c) equivalent D-shaped turns.
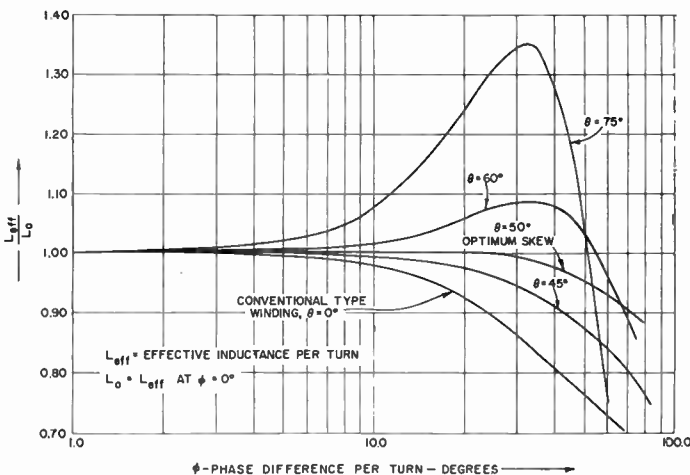


Fig. 3—Calculated curves showing effect of various skew angles on the effective inductance per turn vs phase difference per turn.

optimum skew angle is such that a simple, manufacturable, variable delay line may be constructed using a design similar to that of a standard wire-wound potentiometer. For ease of construction, variations in turn

shapes as shown in Fig. 2(b) and 2(c) were tried and found to produce the same results as the original rectangular turn.

The coil of a typical line is wound on a flat card similar to that used for the winding of the usual commercial potentiometer, with the addition of strips of copper foil to supply the distributed shunt capacitance over the length of the winding (see Fig. 4). These strips are applied to the winding form, and to the external insulating guard strip when additional capacitance is required, by etched-circuit techniques.
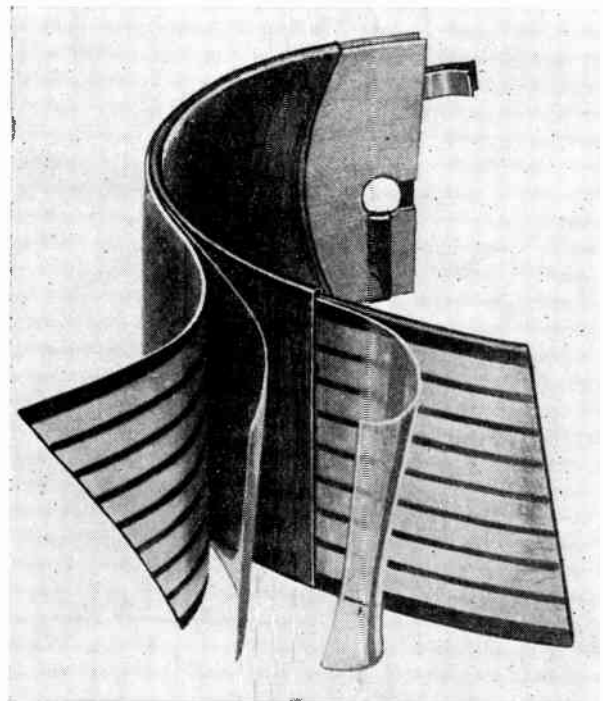


Fig. 4—Photograph of skewed-turn delay line winding opened out to shown ground capacitance strips and D-shaped skewed turns.

Neglecting end effects, the total effective inductance of a delay-line winding is equal to the product of the effective inductance per turn, as defined in the Appendix, and the total number of turns, and is, therefore, proportional to the effective inductance per turn.

The effective inductance per turn of a distributed-parameter delay line can be given by

$$L_{eff} = L + 2 \sum_{k=1}^{N} M_{n(n+k)} \cos k\phi. \tag{1}$$

(See Appendix.)

Where

$L$ = the self-inductance of one turn.

$M_{n(n+k)}$ = the mutual inductance between turn $n$ and turn $(n+k)$.

$\phi = \omega T$ = the phase change between two adjacent turns.

($T$ = the time delay per turn).

$N$ = number of turns with significant mutual inductance values, as discussed below.

To determine the effective inductance per turn at any given $\phi$, we must evaluate $L$ and

$$\sum_{k=1}^{N} M_{n(n+k)}.$$

With reference to Fig. 5, the self-inductance can be
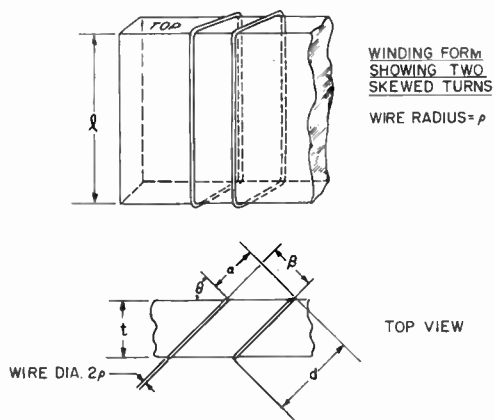
Fig. 5—Diagram of two turns on form of rectangular cross section.

approximated[13] by

$$L = 4l\left[\log\frac{d}{\rho} + \frac{1}{4}\right] \times 10^{-9} \text{ henrys} \qquad (2)$$

where $l$ is the length of the turn expressed in centimeters, and $l \gg d$.

It can be shown also that the mutual, $M$, between the two turns is very nearly

$$M = l \log\left[\frac{[(\alpha + d)^2 + \beta^2][(\alpha - d)^2 + \beta^2]}{(\alpha^2 + \beta^2)^2}\right]$$

$$\times 10^{-9} \text{ henrys}. \qquad (3)$$

(See Appendix.)

Fig. 6 shows typical curves of $M/\text{cm}$ vs $\beta$ and $M/\text{cm}$ vs $\alpha$. The assumptions made in calculating $M$ seem fully justified since an experiment performed using two large rectangular turns of wire produced results less than 2 per cent different from those obtained by use of the formula. Measured values are shown plotted with the computed values in Fig. 6(b). Note that negative mutual coupling is possible when the axes of the two turns are separated beyond a critical distance.

To show how skewing the turns produces a reasonably constant effective inductance over a wide frequency range, the effective inductance per turn vs phase shift

[13] F. W. Grover, "Inductance Calculations," D. Van Nostrand Co., Inc., New York, N. Y., p. 39; 1946.

(a)

(b)
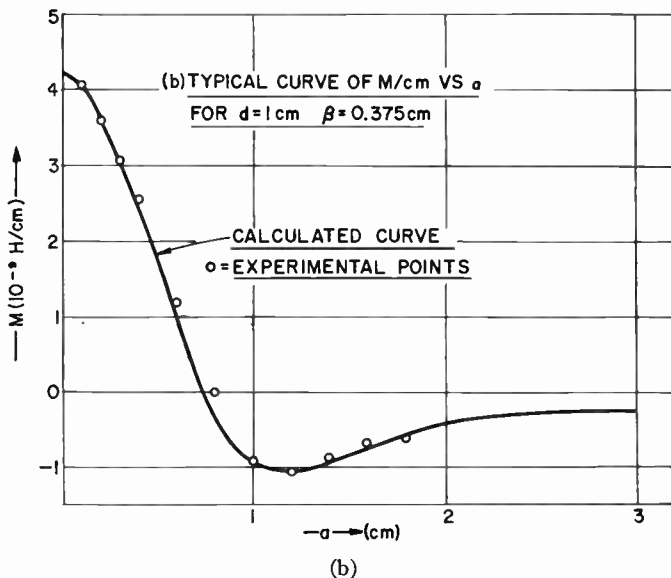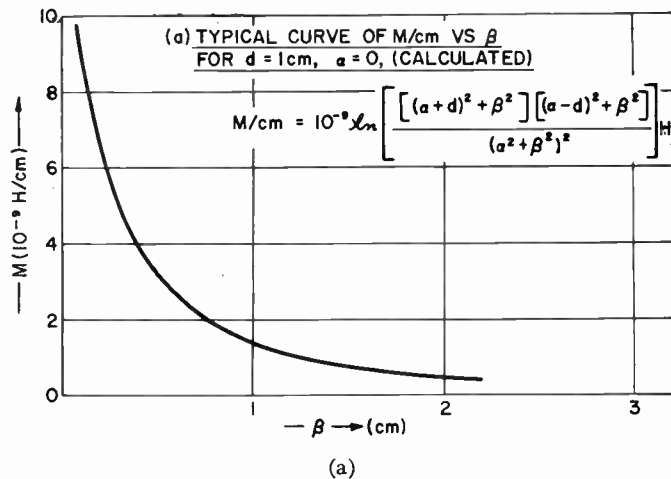
Fig. 6—Curves showing variation of mutual inductance, $M$, between two rectangular turns as (a) their axial separation, $\beta$, is varied, and (b) their displacement, $\alpha$, is varied. (See Fig. 5.)

per turn of a distributed-parameter delay line coil has been calculated for five different values of skew angle using (1), (2), and (3). In the illustrative example (Fig. 7), a value of $N = 6$ was chosen since, with a coil of the dimensions chosen, there is very little magnetic coupling between turns separated by more than six turns.

The low-frequency time delay per turn, $T = \phi/\omega$, is assumed the same in each case so that the five cases will have the same delay per unit length, turns per unit length, and form size. Referring to Fig. 7, the dimensions chosen for the computed examples were: $t = 1$ cm, $s = 1$ cm, and $\rho = 0.1$ cm.

Referring to Fig. 7, we can reduce (2) and (3) to

$$L = 4l\left(\log\frac{d}{\rho} + \frac{1}{4}\right) = 4l\left(\log\frac{t}{\rho\cos\theta} + \frac{1}{4}\right)$$

$$\times 10^{-9} \text{ henrys} \qquad (4)$$

$\theta$ = SKEW ANGLE
$\beta_k$ = $ks$ COS $\theta$
$a_k$ = $ks$ SIN $\theta$
$d$ = $\dfrac{t}{\cos\theta}$

(b)

Fig. 7—Diagram of skewed-turn coil showing quantites used in computation of effective inductance values given in Fig. 3.

and

$$M_{n(n+k)} = l \log\left(1 + \frac{d^4}{k^4 s^4} + \frac{2d^2}{k^2 s^2}\cos 2\theta\right) \times 10^{-9} \text{ henrys}$$

$$= l \log\left(1 + \frac{t^4}{k^4 s^4 \cos^4\theta} + \frac{2t^2 \cos 2\theta}{k^2 s^2 \cos^2\theta}\right)$$

$$\times 10^{-9} \text{ henrys.} \tag{5}$$

Fig. 3 shows the result of these calculations using values of $\theta$ from 0° to 75°. The desired result is evident, since the skewed coil has a more nearly constant inductance than the unskewed one. The curves suggest that any amount of skewing up to the optimum value is an improvement over the regular type of winding. Experimental lines have shown this to be the case. The curves also suggest that over-compensation is possible by increasing $\theta$. This has also been found to be the case, as some lines were constructed for which the time delay increased as a function of frequency up to a certain point.

Notice from (5) that the value of $M_{n(n+k)}$ changes sign at the point $\cos 2\theta = -d^2/2k^2 s^2$ so that even for large $k$ it is impossible to obtain any negative mutual coupling unless $\cos 2\theta < 0$ or $\theta > 45°$. The optimum skew angle for constant effective inductance is therefore always greater than 45°.

Eqs. (4) and (5) show that the optimum angle, $\theta_0$, is primarily a function of the ratio between $s$, $t$, and $\rho$. If the wire spacing, $s/\rho$ is held constant, $\theta_0$ is primarily a function of the winding form thickness, $t$, only.

If capacitance between turns could be neglected, a given coil with optimum skew would provide constant time delay up to the same phase change per turn for different values of ground (shunt) capacitance (or characteristic impedance). However, higher impedance lines require a smaller $\theta_0$ because of the larger phase-correcting effect of the capacitance between turns. When the turns are not rectangular, for example, as in Fig. 2(c), corrections must also be made. Once $\the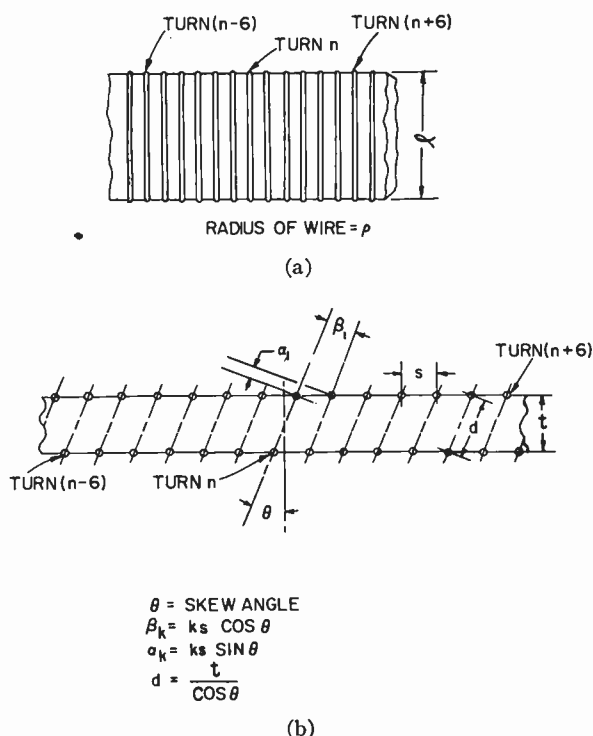ta_0$ is calculated from the equations, construction and measurement of the characteristics of a line shows whether $\theta$ should be increased or decreased slightly.

For the example given, $\mu/\mu_0$ was assumed equal to unity. However, the results should apply for other values of $\mu$ if losses in the permeable material are neglected. Addition of a permeable material in and around the winding may also affect the various capacitances in the circuit, depending on the dielectric constant of the permeable material.

A useful by-product of the application of skewed turns in the manner described is the increase in the inductance of the skewed coil wound on a card of a given thickness with a given number of turns per unit length compared with the inductance of a similar constant-pitch coil on the same card. This result can be used to provide an increase in time delay for a given winding length, and thus an over-all decrease in size for any given performance. Alternatively, larger wire and fewer turns can be used for the same inductance, thus obtaining a higher Q and lower attenuation.

### Experimental Results with Skewed-Turn Coils

Experimental work on variable delay lines constructed on wire-wound rheostat forms was suggested to the authors by W. A. Higinbotham who had constructed similar lines for laboratory applications.[14] Although some delay lines of this form without special delay compensation have been manufactured, the time delay vs frequency of those lines has been constant in most cases over only a relatively small frequency band. In general, the performance of such uncompensated lines is represented by the pulse and step responses of Fig. 1(b) and Fig. 1(f), and the time-delay-vs-frequency curve of Fig. 9(b) for the line with no compensation. Although the first lines of this type were made by Higinbotham with thin copper sheet ground capacitance, many of the commercial lines have made use of interconnected copper strips or comb-shaped capacitance strips for the reduction of eddy-current losses.

Two representative variable delay lines constructed with skewed windings on thin rectangular cards are shown in Fig. 8. The 0.5-$\mu$sec 200-ohm line [Fig. 8(a)] is of the type on which most of the early experimental data were obtained. It became apparent after a few experi-

[14] W. C. Elmore and M. Sands, "Electronics-Experimental Techniques," McGraw-Hill Book Co., Inc., New York, N. Y., p. 352; 1949.
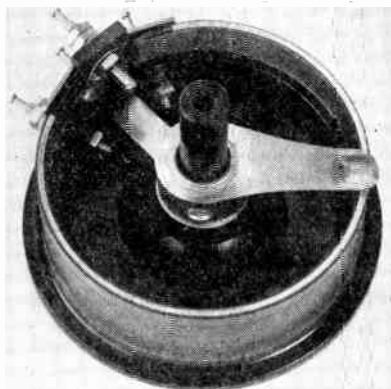
(a)
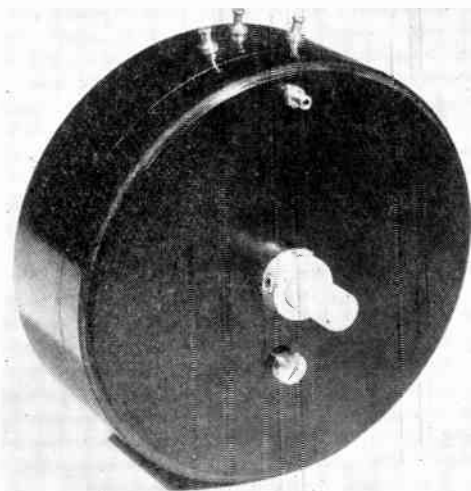


(b)

Fig. 8—(a) Variable delay line of 0.5-μsec maximum delay, 200 ohms impedance, (b) variable delay line of 1.0-μsec maximum delay, 500 ohms impedance.
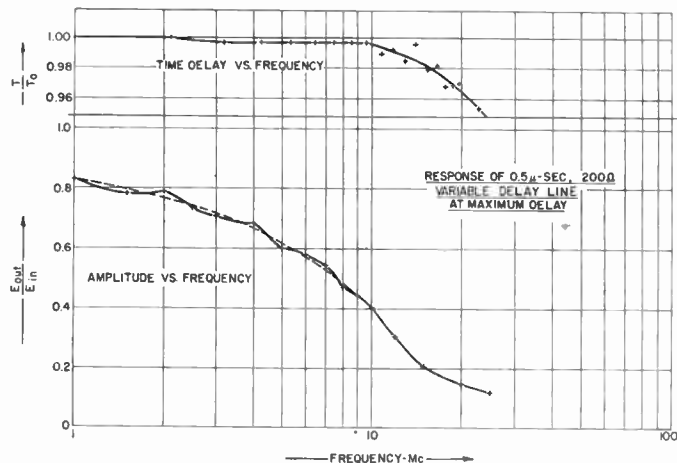


Fig. 9—Time delay and amplitude vs frequency with resistive termination as measured at full delay on 0.5-μsec, 200-ohm variable delay line with skewed winding [see Fig. 8(a)]. Dashed line is smooth curve among measured points (crosses).



Fig. 10—Pulse and step response of 0.5-μsec, 200-ohm experimental variable delay line with skewed winding; (a) pulse input, (b) pulse output at 0.5-μsec delay, (c) step input, (d) step output at 0.5-μsec delay. Scope photos taken on Tektronix 541, 0.1-μsec/cm sweep.

mental lines had been constructed that constancy of time delay as a function of frequency was a more important property than low attenuation over a wide band. It was later found that most of the attenuation observed could be accounted for by copper and dielectric losses, and that the cutoff characteristic of the equivalent low-pass filter cannot be observed since it is overshadowed by the losses.

The time-delay-vs-frequency and the amplitude-vs-frequency curves for this 200-ohm line are given in Fig. 9. Note that the time delay is within 2 per cent of its low-frequency value up to approximately 16.5 mc while the amplitude drops 3 db below the low-frequency value at approximately 6 mc. The pulse and step responses of this line are shown in Fig. 10. The small ripple preceding the main output signal at full delay is probably caused by the fall-off of time delay at high frequencies. It may be noted that the same effect appears in the oscillogram of the step response.

The 1-μsec 500-ohm line, shown in Fig. 8(b), gives the pulse response of Fig. 1(a) and 1(c), and the step response of Fig. 1(d) and 1(f). The pulse and step re-

sponses of a comparable unskewed, uniform-pitch, 500-ohm, 1-$\mu$sec line wound on the same size winding form with the same delay per unit length are shown in Fig. 1(b) and 1(e). The time delay vs frequency of the skewed-winding line, and of the unskewed-winding line for comparison, are shown in Fig. 11(a). Note the re-
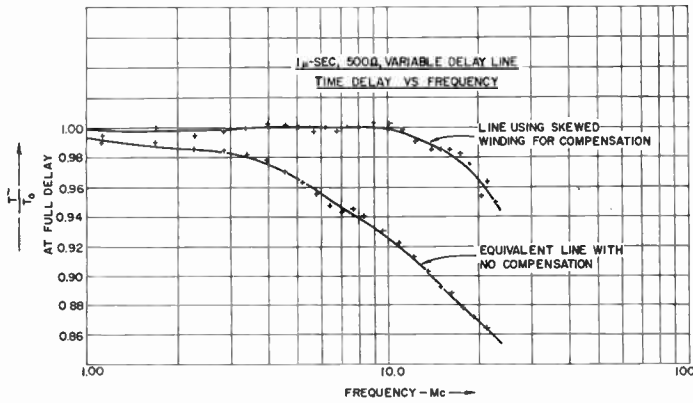


(a)
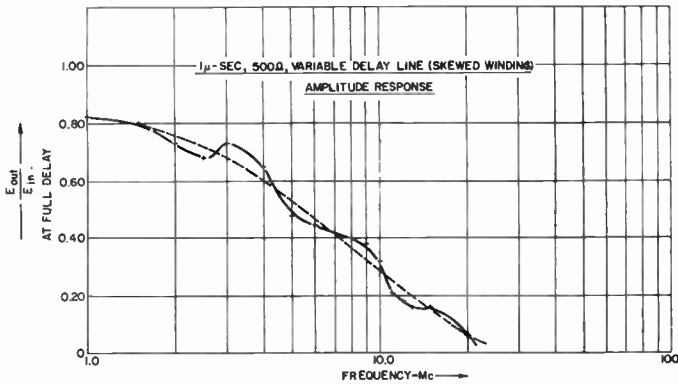


(b)

Fig. 11—Time delay and amplitude vs frequency for 1.0-$\mu$sec, 500-ohm experimental variable delay lines; (a) time delay vs frequency measured at full delay on equivalent lines wound on flat cards with skewed and unskewed windings, (b) amplitude vs frequency for skewed-winding delay line measured at full delay, resistive termination.

semblance to the appropriate curves of Fig. 3. It is apparent from Fig. 3 that it may be possible to achieve a reasonably constant value of time delay out to a limiting frequency, *i.e.*, to the frequency for which the phase change per turn is less than approximately 25° in the example given. For frequencies beyond this value, it is apparently not possible to obtain constant time delay by skewing the winding alone, and hence the optimum design for a skewed-winding line must take into account the highest frequencies it will be necessary to pass. The value of allowable phase shift per turn indicated by the computed illustrative example is approximately four times that actually realized in the experimental delay lines described here. However, it should be noted that the geometry of these examples is considerably different, especially with respect to the turn spacing, $s/\rho$.

The amplitude response vs frequency of the 500-ohm impedance, 1-microsecond, skewed-winding line is given in Fig. 11(b). The attenuation of the unskewed line, which is not shown, was slightly higher because the $Q$ of an equivalent unskewed coil is lower.

The two experimental lines shown in Fig. 8 were optimized to pass pulses of 0.1-$\mu$sec duration with approximately 3-db attenuation at maximum delay within the p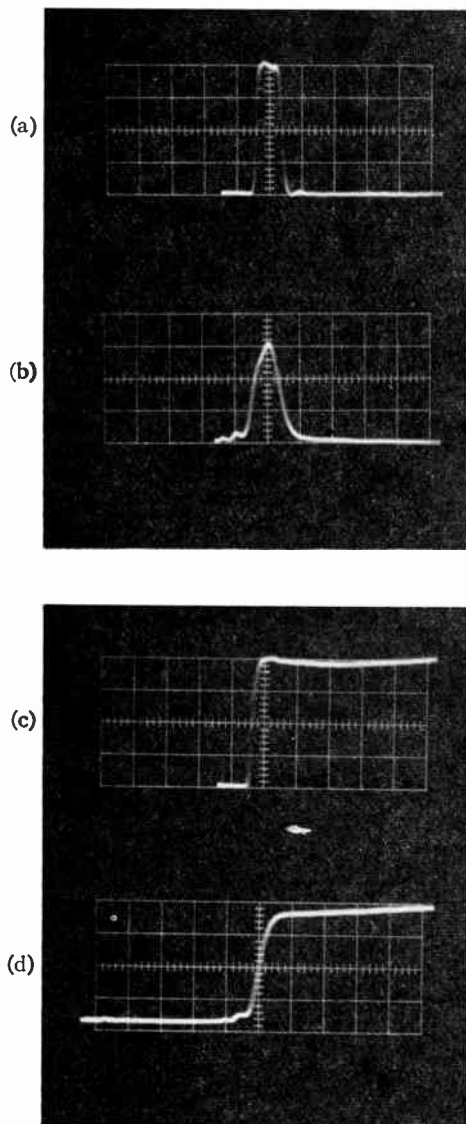hysical size limitations of the standard potentiometer assembly. As can be seen from the oscillograms of Figs. 1 and 10, pulse stretching of the order of 40 per cent at the 10 per cent-amplitude level occurs, while the increase in duration at the half-amplitude point is well below this value. The transient response of the two delay lines to step-function input signals indicates that they are free from many of the overshoot troubles commonly experienced.

### Conclusion

It is apparent that the skewed-turn inductance compensation method offers a new technique for delay equalization of continuously wound distributed-parameter delay lines wound on magnetically isotropic coil forms. The use of this equalization method extends to distributed-parameter delay lines the control over the mutual inductances between individual turns which was formerly available only between the coils of a lumped-element artificial line. It thus becomes possible to produce not only fixed-delay distributed-parameter lines having excellent transient response, but also to construct equivalent variable delay lines incorporating many of the desirable design features of wire-wound rheostats. It is of interest, also, that this method enables easy construction of distributed-parameter delay-equalized variable delay lines of characteristic impedance of 150 ohms or less.

### Appendix

A simplified equivalent circuit of a distributed-parameter delay line showing each turn as a filter section is given in Fig. 12, where $C$ is the capacitance to ground
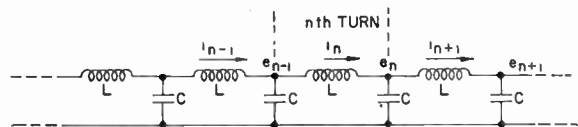


Fig. 12—Simplified equivalent circuit of a distributed-parameter delay line showing each turn as a filter section. Mutual inductance between the turns is treated in the Appendix.

per turn, $L$ is the self-inductance per turn, and there is mutual inductance, $M_{pq}$, between any two turns $p$ and $q$. The capacitance between turns and any possible losses are neglected.

If the current in turn $n$ is given by $i_n = I_n e^{j\omega t}$, then the currents in the preceding and following turns are given by $i_{n-1} = i_n e^{j\omega T}$ and $i_{n+1} = i_n e^{-j\omega T}$, respectively, where $T$ is the time delay per turn.

In the figure $M_{n(n+k)} = M_{n(n-k)}$ and

$$(e_{n-1} - e_n) = L \frac{di_n}{dt} + M_{n(n+1)}\left(\frac{di_{n+1}}{dt} + \frac{di_{n-1}}{dt}\right)$$

$$+ M_{n(n+2)}\left(\frac{di_{n+2}}{dt} + \frac{di_{n-2}}{dt}\right) + \cdots$$

$$= j\omega i_n(L + 2M_{n(n+1)} \cos \omega T$$

$$+ 2M_{n(n+2)} \cos 2\omega T + \cdots).$$

The term

$$(L + 2M_{n(n+1)} \cos \omega T + 2M_{n(n+2)} \cos 2\omega T + \cdots)$$

will be defined as the effective inductance per turn and can be more conveniently expressed as

$$L_{\text{eff}} + L + 2 \sum_{k=1}^{N} M_{n(n+k)} \cos k\omega T$$

where $N$ can be chosen large enough to include the significant mutual inductance terms.

Fig. 13 shows two turns of a card-wound (*i.e.*, rectangular turns) distributed-parameter line which are separated by a distance $\beta$ and displaced by a distance $\alpha$.

It is assumed that the radius of the wire $\rho \ll d$ and that the magnetic field $H$, produced by the ends of turn $A$ can be neglected ($l \gg d$). It is further assumed that the phase shift of the current within turn $A$ is equal to zero, whereas in reality there is a progressive phase shift across the turn.

Having made the above assumptions, the expression for $M_{AB}$, the mutual between turn $A$ and turn $B$, can be derived as follows:

$$M_{AB} = \phi/i, \text{ where } \phi = \mu_0 \int H_n dA; \quad dA = l dx$$

so that

$$M_{AB} = \mu_0 l \int \frac{H_n}{i} dx.$$

Where

$H$ = the magnetic field produced by $i$ in turn $A$.

$H_n$ = the normal component of $H$ at point $P$ in Fig. 13.
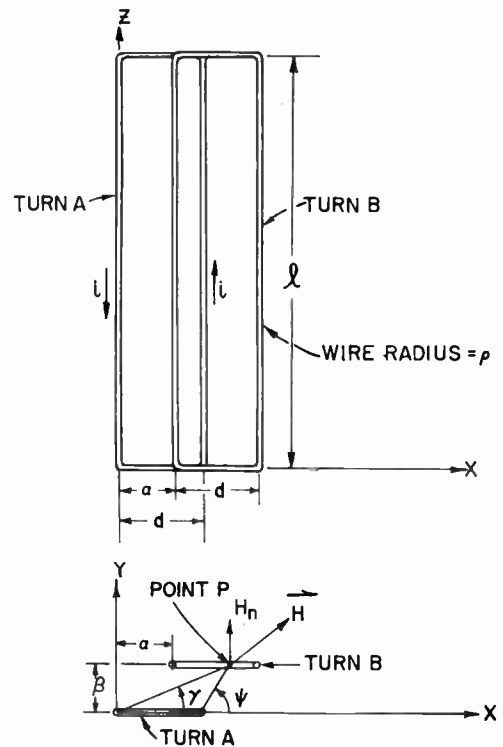


Fig. 13—Diagram of two rectangular wire loops representing two turns of a card-wound coil which are caparated by a distance $\beta$ and displaced by a distance $\alpha$. (See Appendix.)

$$H = \frac{i}{2\pi\sqrt{x^2 + y^2}} - \frac{i}{2\pi\sqrt{(x - d)^2 + y^2}} \cdot$$

$$H_n = \frac{i \cos \gamma}{2\pi\sqrt{x^2 + y^2}} - \frac{i \cos \psi}{2\pi\sqrt{(x - d)^2 + y^2}}$$

$$= \frac{ix}{2\pi(x^2 + y^2)} - \frac{i(x - d)}{2\pi[(x - d)^2 + y^2]}$$

$$M_{AB} = \frac{\mu_0 l}{4\pi i} \int_{\alpha}^{\alpha+d} \left[\frac{ix}{(x^2 + y^2)} - \frac{i(x - d)}{[(x - d)^2 + y^2]}\right] dx$$

$$\therefore M_{AB} = \frac{\mu_0 l}{4\pi} \log\left[\frac{[(\alpha + d)^2 + \beta^2][(\alpha - d)^2 + \beta^2]}{(\alpha^2 + \beta^2)^2}\right],$$

or

$$M_{AB} = l \log\left[\frac{[(\alpha + d)^2 + \beta^2][(\alpha - d^2) + \beta^2]}{(\alpha^2 + \beta^2)^2}\right]$$

$\times 10^{-9}$ henrys

when $l$ is in centimeters.

# Microwave Power Measurements Employing Electron Beam Techniques*

HAROLD A. THOMAS†, SENIOR MEMBER, IRE

*Summary*—A new electron beam technique for measuring microwave power flow, either cw or pulse, in waveguides is described. This technique consists of accelerating an electron beam transversely through an evacuated section of waveguide carrying power in the $TE_{10}$ mode. The transit time of the electrons is adjusted to a value which gives maximum interaction of the field in the guide with the electrons, *i.e.*, electrons gain maximum energy. The energy gained by the electrons is measured by means of a dc stopping potential which can be related to the field. Power is then calculated from the Poynting vector. The instrument takes the form of a sealed-off vacuum tube having a short section of waveguide as part of the tube. The ends of the waveguide may be sealed using known window techniques. The theory for the ideal case is presented and then means of correcting for the various perturbations present in an actual tube are discussed. The theory and the corrections are verified by experiment. The technique appears to have definite value for monitoring or measuring high- and medium-level cw or pulse-power flow. Theoretically the device is self-calibrating and therefore might make a good primary standard. Its suitability as a primary standard is under further investigation.

## INTRODUCTION

THE IMPORTANCE of power measurements at microwave frequencies in today's technology need not be amplified. Low- and medium-level cw power measurements at microwave frequencies may readily be made with good accuracy by various means; however, high-level power measurements introduce greater problems. High-level cw power measurements may be made by means of a calorimeter with the disadvantage that all of the power being measured must be dissipated in the calorimeter. Other methods of cw power measurement include the use of a torque vane or directional coupler and bolometer or thermistor techniques. High-level pulse-power measurements in particular are difficult to make with acceptable accuracy. Accuracies available with present techniques have been reviewed by R. Henning in a recent paper[1] in which he points out that in pulse-power measurements, with present techniques, errors in excess of 30 per cent are to be expected.

The technique described in this paper shows considerable promise as a method of measuring pulse-power with considerable improvement in accuracy. In addition, though it is fundamentally a peak-reading device, it can be used for medium- and high-level cw power measurements. Since measurement is made without disturbing

or affecting the power being measured it should prove particularly valuable for monitoring high-level cw or pulse-power flow.

Theoretically the device is self-calibrating and gives power flow in terms of easily measured physical dimensions, frequency, and dc potentials. For this reason its suitability as a primary standard for microwave pulse power is presently being investigated.

Briefly, the technique consists of observing the change in energy of electrons in a beam passing through a waveguide carrying the power to be measured. The transit time of the electrons across the guide is adjusted so that the interaction of the electric field with the electrons is on a maximum (there are many maxima) and the change in energy of the electrons due to the field is measured after they emerge from the guide by means of a dc stopping or cutoff potential.[2]

The basic arrangement is shown in Fig. 1 which in-
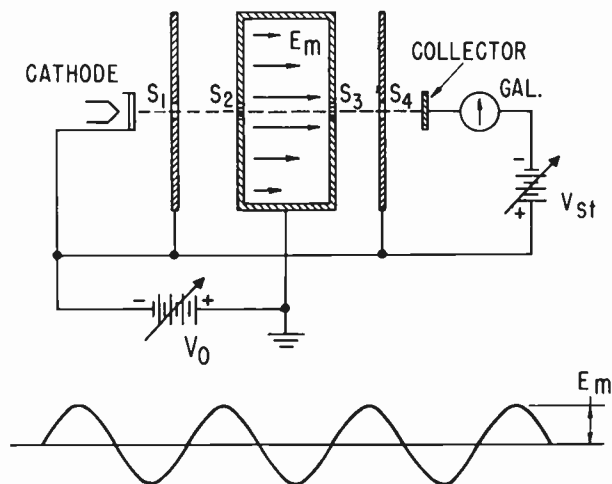


Fig. 1—Basic arrangement.

cludes a means of accelerating electrons through a waveguide parallel to the electric field and a means of measuring the increase in energy of the electrons due to the field by means of a stopping potential applied to the collector. The accelerating voltage $V_0$ is adjusted so that the transit time is $n$ half-cycles of rf where $n$ is an odd integer (for example $n = 7$ in Fig. 1). Naively then, elec-

[1] R. Henning, "Microwave Peak-Power Measurement Techniques," presented at the Conference on High Frequency Measurements at Washington, D. C.; January, 1955.

[2] Shortly after a paper on the work herein described was given at the URSI meeting in Washington, D. C., in May, 1956, it came to the author's attention that work on a similar technique, using a coaxial configuration, was being carried out at the Univ. of Illinois. This is reported by Dong un Sub, "The Concentric Line Power Meter," Tech. Rep. No. 1 to ONR dated June 1, 1956.

trons gain energy during one-half cycle, lose it on the next, etc. It can be seen that for $n-1$ half cycles the net energy gain is zero but on the $n$th half cycle the electrons gain a definite amount of energy which may be related to the field intensity. Upon emerging from $S_4$ electrons have remaining only the energy gained in the guide because $S_4$ is at the same potential as the cathode. The power flow may be obtained from the field intensity by means of the Poynting vector and is thus measured in terms of the dc stopping potential. It is not necessary to know the exact value of $V_0$ to adjust the transit time to that value where electrons gain maximum energy. In practice, $V_0$ is adjusted to that value that gives a maximum stopping potential for current cutoff and then an approximate value of $V_0$ will allow a determination of which nominal value of $n$ is being used.

The effect of various perturbing influences such as the effect of the holes for introducing the electron beam, vswr, etc., will be discussed following the basic theory.

## THEORY

Let monoenergetic electrons enter the guide at $S_2$ with velocity $v_0 = \sqrt{2 V_0 e/m}$. At this point $y=0$, $t=0$, and the field has an arbitrary phase angle $\theta$. The equation of motion in the waveguide space is:

$$\ddot{y} = -\frac{e}{m} E_m \sin (\omega t + \theta) \tag{1}$$

which for the above boundary conditions gives

$$\dot{y} = \frac{eE_m}{m\omega} [\cos (\omega t + \theta) - \cos \theta] + v_0 \tag{2}$$

and

$$y = \frac{eE_m}{m\omega^2} [\sin (\omega t + \theta) - \sin \theta] - \frac{eE_m}{m\omega} t \cos \theta + v_0 t. \tag{3}$$

The energy of electrons at slit $S_4$ is

$$W_{s4} = \int_{y=0}^{y=b} f_y dy = \int_0^\tau f_y \dot{y} dt \tag{4}$$

where $b$ is the distance across the guide, $\tau$ is the transit time, and $f_y$ is the force on the electron due to the electric field and is equal to $m\ddot{y}$. The values of $m\ddot{y}$ and $\dot{y}$ from (1) and (2) may then be substituted into (4) and integrated through the limits indicated. The resulting expression for the energy however must be subjected to the restriction that $t=\tau$ at $y=b$. This may be done by substituting $t=\tau$ and $y=b$ in (3) and combining this result with that already obtained from integrating (4) by eliminating $v_0$ between the two equations. This will give for the final expression for the electron energy at $S_4$ then

$$W_{s4} = \frac{eE_m}{\omega} \left\{ \frac{b}{\tau} [\cos (\omega\tau + \theta) - \cos \theta] \right.$$

$$+ \frac{eE_m}{4m\omega} [\cos 2(\omega\tau + \theta) - \cos 2\theta]$$

$$+ \frac{eE_m}{m\omega^2\tau} \left[ \sin (\omega\tau + 2\theta) - \frac{1}{2} \sin 2(\omega\tau + \theta) \right.$$

$$\left. \left. - \frac{1}{2} \sin 2\theta \right] \right\}. \tag{5}$$

Eq. (5) may be written in the form

$$W_{s4} = \frac{ebE_m}{\omega\tau} [\cos (\omega\tau + \theta) - \cos \theta]$$

$$+ A[f_1(\omega\tau, \theta)] + B[f_2(\omega\tau, \theta)]. \tag{6}$$

It may be shown that even at the highest values of field likely to be encountered in microwave practice, the values of $A$ and $B$ are only a fraction of a per cent of the quantity $ebE_m/\omega\tau$. Furthermore, for the particular condition of interest, i.e., at the values of $\theta$ and $\tau$ that make $W_{s4}$ a maximum the quantity $[\cos (\omega\tau+\theta)-\cos \theta]$ approaches the value 2 while $f_1(\omega\tau, \theta)$ and $f_2(\omega\tau, \theta)$ both vanish. It is interesting to note that the 1st term of (6) is the result that would have been obtained if in the analysis the electric field the electron sees as it moves across the guide with average velocity $b/\tau$ were expressed as a function of $y$; i.e., $E = E_m \sin [(\omega\tau/b)y+\theta]$. Use of this fact will be made later.

Now let $\tau = n\pi/\omega$ where $n$ may vary only slightly from an odd integer value. Practically, this may be accomplished by adjusting $V_0$. From (5) or (6) and dropping the 2nd and 3rd terms, there is obtained:

$$W_{s4} = \frac{ebE_m}{n\pi} [\cos (n\pi + \theta) - \cos \theta]. \tag{7}$$

A plot of this expression is shown in Fig. 2 for those
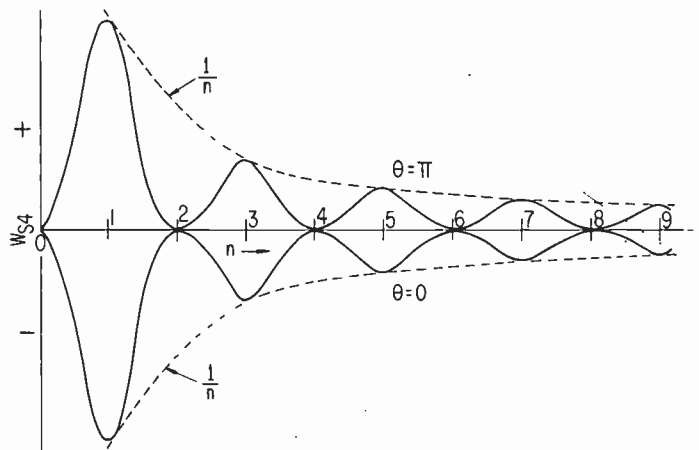


Fig. 2—Electron energy as a function of $n$.

electrons that enter when $\theta = \pi$ (gain maximum energy) and $\theta = 0$ (lose maximum energy). Because of the $1/n$ dependence it can be seen that the energy is maximum at values of $n$ slightly below the odd integer values.

The exact values of $n$ and $\theta$ that make $W_{s4}$ a maximum, may be obtained by taking the partial derivatives of $W_{s4}$ in (7) with respect to $\theta$ and $n$, setting these equal to zero and solving the resulting equations simultaneously. These do not permit of an exact analytic solution, but can be reduced to the expressions: $\tan\theta = \sin n\pi/(1-\cos n\pi)$ and $n\pi(\sin n\pi) = 2(1-\cos n\pi)$, which may readily be solved by approximate methods by letting $n = (n'-\epsilon)$ and $\theta = (\pi + \delta\pi)$, where $n'$ is an odd integer and $\epsilon$ and $\delta$ are small variations from the nominal values of $n$ and $\theta$. Table I shows exact values of $n$ and $\theta$ for the nominal values of $n$ as calculated from the above expressions. Also included is the error in field intensity that would result if the nominal values of $n$ and $\theta = \pi$ were used rather than the exact values. It is obvious from (7) that for the nominal values of $n$ and for $\theta = \pi$ that $W_{s4} = 2ebE_m/n\pi$ and that therefore the relation between field intensity and cutoff stopping potential is

$$V_{st_c} \cong \frac{2bE_m}{n\pi}. \qquad (8)$$

TABLE I

| $n'$ Nominal | $n$ Exact | $\theta$ Radians | Per Cent Error |
|---|---|---|---|
| 11 | 10.963 | 3.199 | 0.17 |
| 9 | 8.955 | 3.213 | 0.26 |
| 7 | 6.942 | 3.233 | 0.45 |
| 5 | 4.919 | 3.270 | 0.80 |

Table I can be used to compute the exact value $V_{st}$ in terms of exact values of $n$ and $\theta$.

The acceleration voltages required for different nominal values of $n$ may be determined by solving (3) for $v_0$ and using $t = \tau = n\pi/\omega$, $\theta = \pi$, and $y = b$. Combining this result with $v_0 = \sqrt{2V_0e/m}$ gives

$$V_0 = \frac{m}{2e}\left(\frac{b\omega}{n\pi}\right)^2 - \frac{bE_m}{n\pi} + \frac{e}{2m\omega^2}E_m^2. \qquad (9)$$

As might be expected (9) indicates that the field in the guide affects the value of $V_0$ required. The 2nd and 3rd terms of (9), however, are small compared to the first term even for very high values of field intensity produced by high power levels. Fig. 3 shows the accelerating voltage required for different values of $n$ at a particular frequency in the $X$ band. It can be seen that values of voltage required, for example, for $n = 5$ or 7 are entirely practical and affected only slightly by power level.

Using the Poynting vector to obtain the relation between field intensity and power flow and using (8) there is obtained the relation for power flow in terms of stopping potential

$$P = V_{st_c}^2\sqrt{\epsilon/\mu}\sqrt{1-(fc/f)^2}\ (n\pi/4)^2 \cdot a/b \qquad (10)$$

where $a$ and $b$ are the dimensions of the guide, $f$ and $fc$
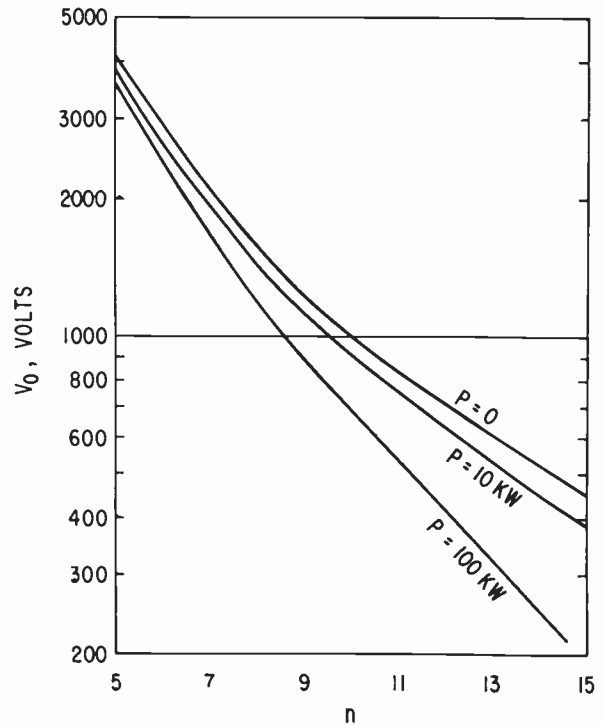


Fig. 3—Accelerating voltage as a function of $n$.

the frequency and cutoff frequency of the guide respectively, and $\sqrt{\epsilon/\mu}$ has the approximate value of $1/120\pi$ for free space. Eq. (10) gives the power in watts when $V_{st_c}$ is in volts.

It should be remembered that electrons are arriving continuously at $S_2$ and therefore all values of $\theta$ are possible. However, it should be clear from the above discussion that only those electrons that enter at $\theta$ equal or near $\pi$ will gain maximum energy for a given value of $n$ or $V_0$. It is interesting to look at the distribution of electron energies arriving at slit $S_4$. The effect of varying $\theta$ about the optimum value $\pi$ may be investigated by substituting $\theta = \pi + \alpha$ into (7) and adjusting $V_0$ to make $n$ equal to an odd integer. Then there results

$$V_\alpha = V_{st_c}\cos\alpha \qquad (11)$$

where $V_{st_c}$ is the cutoff stopping potential for $\alpha = 0$. Let $I_c$ be the average current reaching the collector for a given stopping potential and $I_4$ be the current entering the guide. Assuming no current loss in the guide or on $S_4$, $I_c/I_0 = \alpha/\pi$ where $2\alpha$ is the angle of acceptance of electrons by the collector for a retarding potential $V_\alpha$. Combining this with (11) gives

$$V_\alpha = V_{st_c}\cos\frac{I_0}{I_c}\pi. \qquad (12)$$

This expression is important because it shows that current cutoff is approached abruptly on the voltage axis of the current-voltage curve which makes cutoff easy to determine. It also indicates that theoretically at least the field intensity in the guide could be measured by looking for the stopping potential for which the current

to the collector starts to decrease. This would measure the energy of electrons which had lost, rather than gained, maximum energy. Practically, this method presents some difficulties so there is no obvious advantage of using this method over that of current cutoff.

The above discussion and (12) of course assumed monoenergetic electrons entering the waveguide and the thermal energy distribution present will affect this result. Redhead and Crowell[3] have calculated the effect of thermal energies on the cutoff characteristic of an ion beam in a retarding field and their result could be applied to the above analysis. Also Burkhard[4] has analyzed the effects of thermal distribution for electron beams. However, the details will not be included in this discussion because it would serve no useful purpose, since by an appropriate experimental technique it is possible to essentially eliminate the effect of the thermal energy distribution.

## PERTURBING EFFECTS

There are a number of effects present in a practical operating tube that were not included in the above analysis for the ideal case. Some of these effects, the errors introduced, and the corrections for them, will now be considered in the order of their importance.

### Effect of VSWR

Since the device measures the field at only one point in the waveguide, a large vswr could introduce large errors in the power calculations. The maximum error possible due to a vswr of $\rho$ can be determined as follows. The true net power flow $P_T = k(E_+^2 - E_-^2) = k(E_+ + E_-) \cdot (E_+ - E_-)$ where $E_+$ and $E_-$ represent the crest values of the voltage waves traveling toward the load and away from the load respectively. The measured power would have maximum error if the electron beam passed through the field at a point where it was maximum, *i.e.*, where $E = (E_+ + E_-)$. The measured power would then be $P_m = k(E_+ + E_-)^2$. From the above and remembering that $\rho = (E_+ + E_-)/(E_+ - E_-)$ it follows that the maximum error is

$$\text{max error} = \frac{P_m - P_T}{P_T} = \rho - 1. \tag{13}$$

A standing wave ratio of 1.05 then might introduce as much as 5 per cent error in the power determination. Even though the load might be perfectly matched, the mica, glass, or ceramic windows that allow the microwave power to flow through the evacuated section of the waveguide will introduce some vswr. This could be corrected for by knowing the distance from the load side window to the electron beam and the reflection characteristics of the window for all frequencies in the band to be covered.

Other methods would be to either move the beam with respect to the field or to use a tube with 3 fixed beams. Either of these would allow a complete determination of the power flow in both directions and the value of $\rho$ regardless of the source of the vswr.

### Hole Effect

Holes are required on either side of the evacuated section of waveguide to pass the electron beam through. Since the field is perturbed in the vicinity of the holes, the maximum energy gained by the electrons is affected. The error introduced, if not corrected for, can be considerable, depending on the ratio of the hole diameter to the distance across the guide and the value of $n$. In order to solve this problem it is assumed that a static field solution may be obtained and then the time variation imposed on the static solution. An analytic solution for the field about the hole in an infinitely thin diaphragm has been obtained.[5] However, in this case where the hole is long compared to its diameter, an analytic solution is not known. Hence the field distribution was obtained by means of the wedge shaped electrolytic tank technique[6] for axial symmetry. All distance measurements were expressed in units of hole diameters and field intensities were expressed relative to the uniform value $E_0$ in the center of the guide. Analytic expressions were then fitted to different parts of the curve and it was found that the distribution could be closely represented by means of four exponential functions. Fig. 4 (next page) shows a plot of two of the functions chosen (solid line), and the experimental distribution (points). The correspondence between the two is quite adequate since this curve is to be used only to obtain a correction for an error that would not ordinarily exceed 10 per cent.

The functions assumed to describe the field the electrons see during transit across the guide are as follows:

$$E_1 = 0.5E_0e^{\alpha y'} \qquad\qquad -\infty < y' < 0 \text{ (inside hole)}$$

$$E_2 = E_0(1 - 0.5e^{-\alpha y'}) \qquad 0 < y' < b'/2$$

$$E_3 = E_0(1 - 0.5e^{-\alpha(b'-y')}) \quad b'/2 < y' < b'$$

$$E_4 = 0.5E_0e^{\alpha(b'-y')} \qquad\quad b' < y' < \infty \text{ (inside hole)}$$

where $E_0 = E_m \sin(\omega\tau + \theta)$.

It has already been shown in connection with the derivation of (7) that the field can be expressed as a function of $y$ rather than $t$, *i.e.*, $E_0 = E_m \sin[(\omega\tau/b)y + \theta]$. The energy of the electrons at electrode $S_4$, including the hole effect is then

$$W_{s4H} = -e\left[\int_{-\infty}^{0} E_1 dy + \int_{0}^{b/2} E_2 dy + \int_{b/2}^{b} E_3 dy + \int_{b}^{\infty} E_4 dy\right].$$

[3] P. A. Redhead and C. R. Crowell, "Analysis of the linear rf mass spectrometer," *J. Appl. Phys.*, vol. 24, pp. 331–337; March. 1953.
[4] D. G. Burkhard, private communication, April, 1956

[5] Y. K. Zworykin, G. A. Morton, E. K. Ramberg, J. Hillier, and A. W. Vance, "Electron Optics and the Electron Microscope," John Wiley and Sons, Inc., New York, N. Y., p. 383; 1945.
[6] *Ibid.*, pp. 389–393.
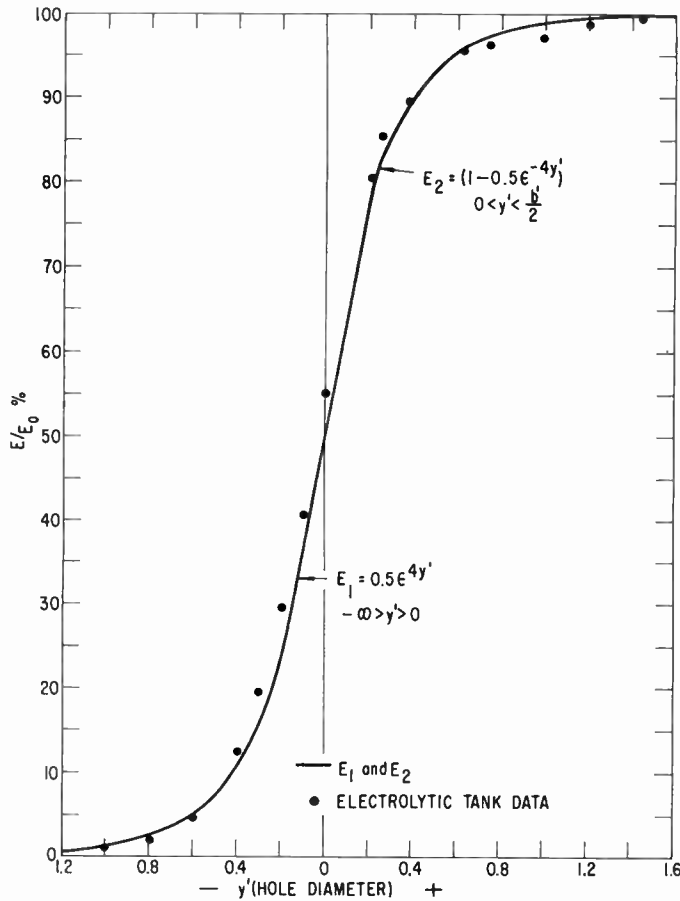
Fig. 4—Experimental field intensity in the vicinity of the hole.

Carrying out the above integrations, with $\omega = n\pi$, and using the very good approximation that for small hole diameters $e^{-\alpha b/2d} \cong 0$ (since $b/d$ is very large) there is obtained

$$W_{s4H} = \frac{ebE_m}{n\pi}\left(1 - \frac{(n\pi d)^2}{(b\alpha)^2 + (n\pi d)^2}\right)$$
$$\cdot [\cos(n\pi + \theta) - \cos\theta]. \quad (14)$$

The similarity of this expression to (7) should be noted.
If

$$W_{s4} = \frac{ebE_m}{n\pi}\left[\cos(n\pi + \theta) - \cos\theta\right]$$

is the energy gained for no hole effect, the error due to holes [using (14)] is

$$\text{error} = (W_{s4} - W_{s4H})/(W_{s4H} = (n\pi d/b\alpha)^2. \quad (15)$$

The last column of Table II shows the errors for different values of $n$ and $b/d$ that would result if no correction were applied using $\alpha = 4$ (see Fig. 4).

Taking the partial derivatives of (14) with respect to $n$ and $\theta$, and proceeding as before to find the values of $n$ and $\theta$ that will give maximum energy, yield the following expressions

$$\tan\frac{\epsilon\pi}{2} = \frac{2}{n\pi}\left[1 + 2\left(\frac{n\pi d}{\alpha b}\right)^2\right] \quad (16)$$

and

$$\theta = \pi + \frac{\epsilon\pi}{2} \quad (17)$$

where $n$ is the exact value, $n'$ the odd integer or nominal value and $\epsilon = n' - n$. Solving (16) and (17) by approximation methods gives the exact values of $n$ and $\theta$ shown in Table II for $\alpha = 4$. The values are only slightly different than those shown in Table I. It is interesting to note that for the larger hole diameters and values of $n$ the error due to the holes becomes quite large.

TABLE II

| $n'$ Nominal | $n$ Exact | $\theta$ Radians | Error Per Cent |
|---|---|---|---|
| | $b/d = 40$ | $\alpha = 4$ | |
| 5 | 4.916 | 3.273 | 0.93 |
| 7 | 6.939 | 3.236 | 1.86 |
| 9 | 8.952 | 3.217 | 3.80 |
| 11 | 10.959 | 3.205 | 4.63 |
| | $b/d = 20$ | $\alpha = 4$ | |
| 5 | 4.911 | 3.280 | 3.72 |
| 7 | 6.933 | 3.247 | 7.41 |
| 9 | 8.943 | 3.230 | 12.35 |
| 11 | 10.949 | 3.221 | 18.50 |
| | $b/d = 10$ | $a = 4$ | |
| 5 | 4.993 | 3.208 | 15.4 |
| 7 | 6.907 | 3.287 | 29.4 |
| 9 | 8.909 | 3.283 | 49.0 |
| 11 | 10.908 | 3.286 | 73.4 |

The final expression for the stopping potential by using (14) is:

$$V_{stc} = \frac{bE_m}{n\pi}\left[\frac{(\alpha b)^2}{(\alpha b)^2 + (n\pi d)^2}\right][\cos(n\pi + \theta) - \cos\theta]. \quad (18)$$

The expression $[\cos(n\pi + \theta) - \cos\theta]$ for maximum energy electrons can be simplified by using (17) and reduces to $2\cos\epsilon\pi/2$. The corrected expression for power flow is then

$$P = \sqrt{\epsilon/\mu}\sqrt{1 - (fc/f)^2}\left(\frac{n\pi}{4}\right)^2\frac{a}{b}V_{stc}^2$$
$$\cdot\left[1 + \left(\frac{n\pi d}{\alpha b}\right)^2\right]^2\left(\cos\frac{\epsilon\pi}{2}\right)^{-2}. \quad (19)$$

For the experimental tubes that have been constructed and for $n = 7$ or less, (19) may be simplified to

$$P = \sqrt{\epsilon/\mu}\sqrt{1 - (fc/f)^2}\left(\frac{n\pi}{4}\right)^2\frac{a}{b}V_{stc}^2$$
$$\cdot\left[1 + 2\left(\frac{n\pi d}{\alpha b}\right)^2\right] \quad (20)$$

without introducing an error in excess of 0.8 per cent.

*Thermal Energy Distribution*

The thermal distribution creates a tail on the current cutoff curve and makes true cutoff difficult to determine at low power levels. However, by subtracting the apparent cutoff voltage obtained without power flow in the guide from that value obtained with the power flow a true cutoff potential is determined. This is illustrated in the experimental curve of Fig. 5 in which the cutoff potential thus obtained is plotted as a function of the collector current sensitivity. It is seen that when the current sensitivity is such that it will measure less than 1 per cent of the current $I_0$ into the guide, the cutoff potential doesn't change. This technique will also eliminate the error due to contact potential.

For very high-level power measurements, the values of $V_{st_c}$ will be of the order of a few hundred volts and both the contact potential and the small cutoff error due to thermal distribution will be negligible.

*Space Charge*

Space charge could affect the potential distribution in the guide and thus the velocity, and to a very small extent, the energy gained by the electrons. This effect cannot even be observed at the low input beam current of a few microamperes which have been used in this experiment.

### EXPERIMENTAL RESULTS

Experimental tubes have been constructed and the theory and corrections previously described have been checked. To provide flexibility, the tubes were of the demountable type using teflon gaskets and continuous evacuation by mercury pumps. It was found that oil pumps and rubber gaskets were not suitable because the hydrocarbon molecules present in the system would settle out on the electrodes and become polymerized by electron bombardment.[7,8] In the case of the collector, the resulting tar-like film, being highly insulating, builds up a negative charge and thus causes a stopping potential different from that actually applied.

In an early tube, magnetic focusing was used; however, electrons with insufficient energy to reach the collector, being trapped by the magnetic field, would make many oscillations through the tube. This resulted in excessive positive ion current to the collector and had other undesirable results. Electrostatic focusing has also been used; but while this increased the total space current through the tube, it did not increase the current in the vicinity of cutoff sufficiently to warrant the additional complexity. This type of focusing will usually introduce additional spread in the energies of the electrons. Since only very small currents through the tube are required, the present experimental tubes have no focusing.

[7] R. L. Stewart, "Insulating films formed under electron and ion bombardment," *Phys. Rev.*, vol. 45, pp. 488–490; April, 1934.
[8] A. E. Ennos, "The sources of electron-induced contamination in kinetic vacuum system," *British J. Appl. Phys.*, vol. 5, pp. 27–31; January, 1954.
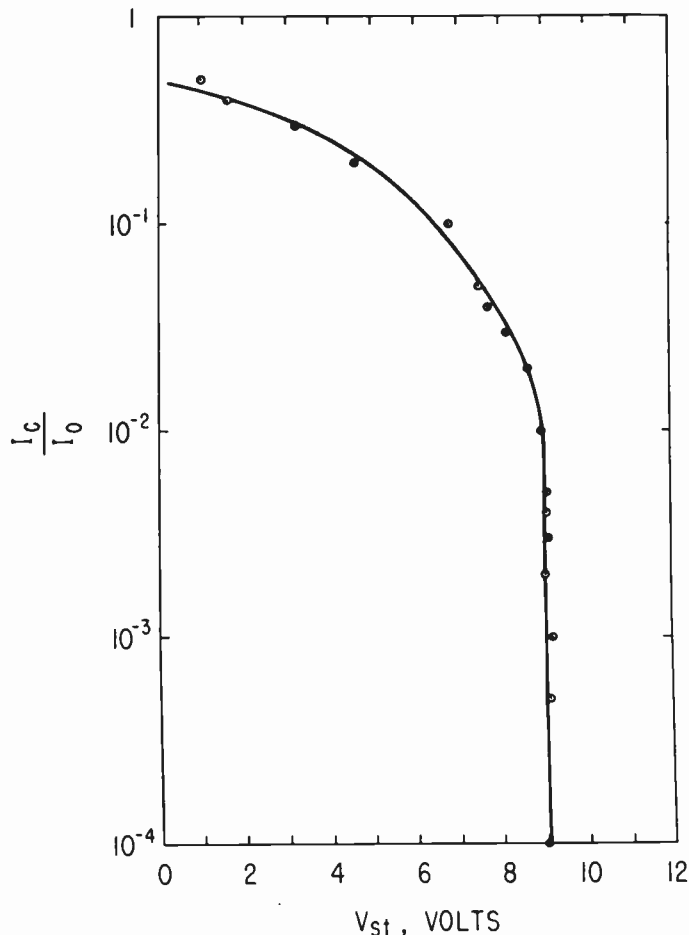


Fig. 5—Relative collector current as a function of stopping potential.

They do use collimating diaphragms having holes of 0.02 inch in diameter in front of the filament and also in front of the collector. The collector is not a flat plate, but is actually a Faraday cage to prevent the loss of secondary electrons.

In a sealed-off version of the tube, an oxide-coated, indirectly-heated cathode would be desirable. Since the experimental versions were demountable, tungsten filaments were used to prevent cathode poisoning.

The tube was constructed by using a short section of $X$-band waveguide (RG-52/u) and sealing the two ends at the flanges with 0.003-inch mica windows seated on teflon gaskets. Klystron grid stock was used in each side of the waveguide to pass the electron beam through and at the same time keep the field distortion in the guide to a minimum. The klystron grid stock, obtained through the courtesy of Varian Associates, produced a grid the same thickness as the waveguide wall (0.05 inch) and having many holes 0.01 inch in diameter. The metal separating these holes is only 0.001 inch thick and hence the transparency is excellent. The diameter of the transparent section was 0.170 inch; however, the beam was defined by the collimating diaphragms mentioned above. The 0.01-inch holes give a ratio of $b/d = 40$, which is quite satisfactory.

The section of waveguide was surrounded by an adequate metal vacuum envelope with provisions for bringing in the necessary leads. Fig. 6 shows a photograph of an experimental tube and the associated measuring equipment. All measurements to date have been made with an X-band, 20-watt klystron. No pulse power measurements have been made as yet because of the lack of a pulse-power source. However, the cw source used has been quite adequate for checking the basic theory of the technique and the corrections involved.
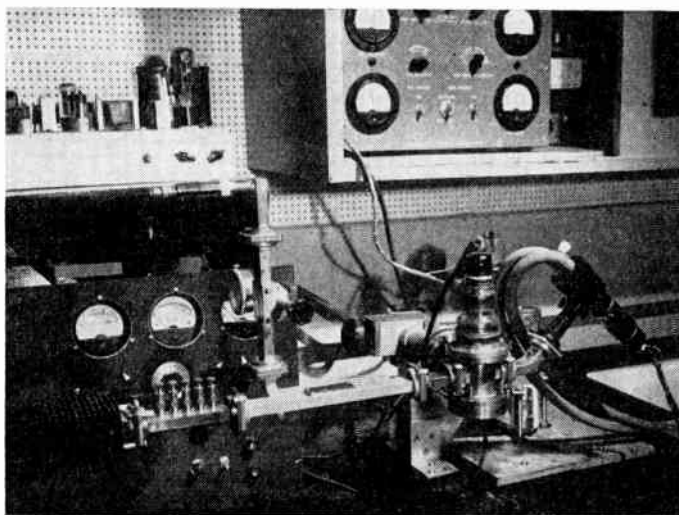


Fig. 6—Laboratory setup showing experimental tube. Pumping equipment is seen in the background, water cooled X-band klystron on the right, and load on the left.

Fig. 7 shows the experimental verification of the theory. The power in the guide was held constant and the various values of accelerating voltage which gave maxima in the cutoff potential were found. These experimental data (x points) are compared with the theory [(9), solid curve] in Fig. 7. The experimental points should be slightly higher than the theoretical values because the theoretical expression for $V_0$ did not include the effect of the fringing field. The other curves of Fig. 7 show the measurement of the same value of power for different values of $n$. In the one case the power is calculated from (10) using no corrections and this is compared with the power values obtained when (20) is used. A value of $\alpha = 4$ was used for the hole correction. Another series of measurements on a later tube of slightly different design gave values of power that were consistent to within 0.5 per cent for various values of $n$; in this case the value of $\alpha$ required to give constant power for the different values of $n$ was approximately 3 instead of 4. This indicates that the form of the equation derived from the electrolytic tank experiment is correct but that the value of $\alpha$ should be determined for each tube design. In practice, values of $n$ of about 7 would probably be used and the difference in the power correction, using $\alpha = 4$ or 3 for $n = 7$, is about 2 per cent. Since this is greater than the observed consistency of the data,
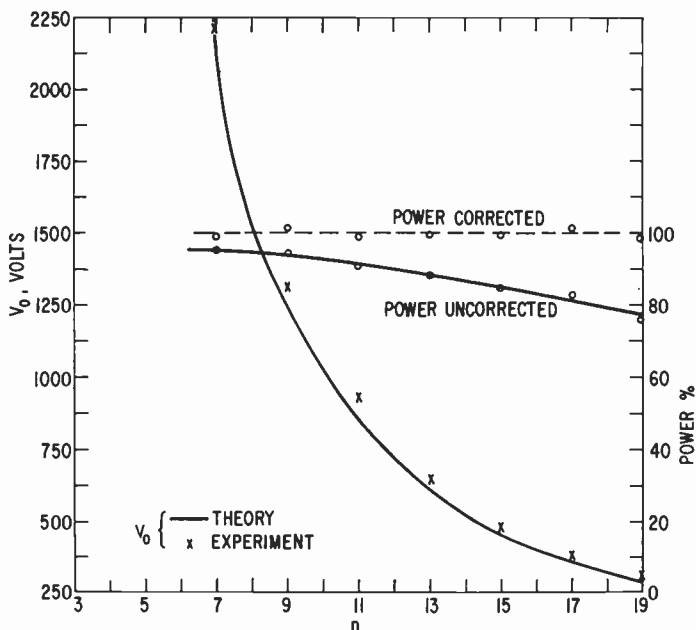


Fig. 7—Experimental verification of theory. Power was held constant for the different values of $n$. Corrected power curve uses the derived correction.

it would be possible to determine the exact value of $\alpha$ for a given tube by taking measurements of the same power level for different values of $n$. This would then be a constant for that particular tube.

In the above measurements, the saturation current $I_0$ was approximately $3 \times 10^{-7}$ amperes. If the tube is used for pulse power measurement, the ratio of the saturation current to the sensitivity will have to be increased by the duty cycle factor. This does not impose any difficulty, however, for sensitivities of $10^{-13}$ can readily be obtained and the space current in the guide could be materially increased with no ill effects. A practical tube of this type would ordinarily be sealed off and then an oxide-coated heater-type cathode could be used. No doubt this would improve the energy distribution in the vicinity of cutoff because of the lower cathode temperature.

## Conclusion

The experimental data obtained serves as a good check on the theory of the device described. It appears to have considerable value where it is desired to measure or monitor high-power levels for either cw or pulse without disturbing the power being measured. Theoretically it is self-calibrating; however, more experimental work will have to be done and comparisons made with other standards before its value as a primary standard can be thoroughly evaluated.

## Acknowledgment

# VHF Pulse Techniques and Logical Circuitry*

D. E. ROSENHEIM†, SENIOR MEMBER, IRE, AND A. G. ANDERSON†, ASSOCIATE MEMBER, IRE

*Summary*—Techniques and components for use in systems handling pulses of 10-millimicrosecond width have been investigated. Bandwidth requirements have led to the use of secondary emission pentodes in amplifier service. The limitations of time delay in feedback type circuitry have made necessary the use of special logical reshaping circuits. A test program on commercially available semiconductor diodes resulted in the selection of high-conductance, gold-bonded junction diodes for use in switching circuits. Multivibrator circuits have been designed for gating and delay functions. Electromagnetic delay lines of both the coaxial and helical-wound types have been used for delay and for pulse generation.

These components have been applied to the design of an arithmetic unit which performs binary addition, multiplication, and dynamic storage at a pulse repetition rate of 50 megacycles per second.

## INTRODUCTION

PULSE-CODED information systems have found application in communications, digital computers, and nuclear instrumentation. The desire to handle a maximum amount of information in a minimum amount of time has led to the development of systems using increasingly narrow pulses at increasingly high repetition rates. Resolution of millimicrosecond pulses at low repetition rates has been demonstrated in nuclear instrumentation research.[1,2] A 50-megacycle per second pulse repetition rate using pulses 10 millimicroseconds wide was considered by the authors to be a practical upper limit on the basis of the available components and the state of measuring techniques. Having established this repetition rate, components and circuits were developed to perform amplification, pulse shaping, and switching. An interesting application is the demonstration that digital computer operations can be performed at a 50-megacycle pulse repetition rate using these components and circuits.

## PULSE AMPLIFIERS

The rise time[3] in seconds of an uncompensated pentode video amplifier is given by

$$t_r = 2.2RC,$$

where $R$ is the plate load resistance in ohms, and $C$ is the total capacity in farads at the plate. For a 10-millimicrosecond pulse, assuming an allowable rise time of 4.4 millimicroseconds, and a value of $C$ of 10 micromicrofarads, $R$ is found to be 200 ohms. When using an amplifier to re-establish pulse amplitude and to decrease pulse rise and fall times, the amplifier gain should be of the order of 3 to 5. Postulating a minimum gain, $A_0$, of 3 from such a stage the average $g_m$ requirement is given by:

$$g_m = A_0/R = 3/200 = 15,000 \ \mu\text{mhos}.$$

Using vacuum tubes either clamped on or biased off, signal levels of from 5 to 10 volts are required. Assuming an 8-volt signal level and 200-ohm load resistance the required peak tube current is 40 ma.

Although much progress is being made in the field of high-speed semiconductor devices, there are no available transistors at the time of this writing which will operate at a 50-megacycle per second pulse repetition rate. An examination of available vacuum tubes led to the choice of the Western Electric 417A/5842 triode and the Philips EFP-60 secondary emission pentode. The Western Electric 417A/5842 triode has a $g_m$ of 25,000 $\mu$mhos, a plate dissipation of 4 watts, a maximum average plate current of 35 ma, a cathode to grid and heater capacity of 9 micromicrofarads, and a plate to grid and heater capacity of 1.8 micromicrofarads. The Philips EFP-60 secondary emission pentode has a grid to anode transconductance of 25,000 $\mu$mhos, a plate dissipation of 2 watts, a dynode dissipation of 1 watt, a maximum average cathode current of 8 ma, and a total input plus output capacity of about 15 micromicrofarads. The EFP-60 has the advantage of having an additional active electrode, the dynode, which allows the tube to be used as a noninverting amplifier with a grid to dynode transconductance of 20,000 $\mu$mhos. One may make use of the active electrodes of the EFP-60 to obtain circuits which would require 2 tubes of conventional design.

A noninverting limiter amplifier is shown in Fig. 1. Power supply decoupling and parasitic suppression resistors have been employed, and shunt peaking has been used to decrease the rise time of the signal on the dynode to about 3 millimicroseconds. This amplifier is normally biased below cutoff and is driven positive by pulses fed to its control grid through a dc restorer circuit, which maintains the grid at a constant dc bias voltage. The use of positive pulses, feeding the cutoff grid, has the advantages of elimination of baseline noise and the reduction of average power dissipation. The diode-resistor combination in the cathode circuit of the amplifier allows the tube to draw 10-ma cathode current without degeneration. However, if the tube draws more than 10-ma current, the diode opens and the gain is reduced by a factor of twenty-five. The 10-ma cathode current produces approximately 40 ma of dynode current so that the peak output voltage from the

[1] I. A. Lewis and F. H. Wells, "Millimicrosecond Pulse Techniques," McGraw-Hill Book Company, Inc., New York, N. Y.; 1954.
[2] N. F. Moody, G. J. R. Maclusky and M. O. Deighton, "Millimicrosecond pulse techniques," *Electronic Eng.*, vol. 24, pp. 214–219; May, 1952.
[3] Committee Personnel, "IRE standards on pulses: methods of measurement of pulse quantities, 1955," PROC. IRE, vol. 43, pp. 1610–1616; November, 1955.
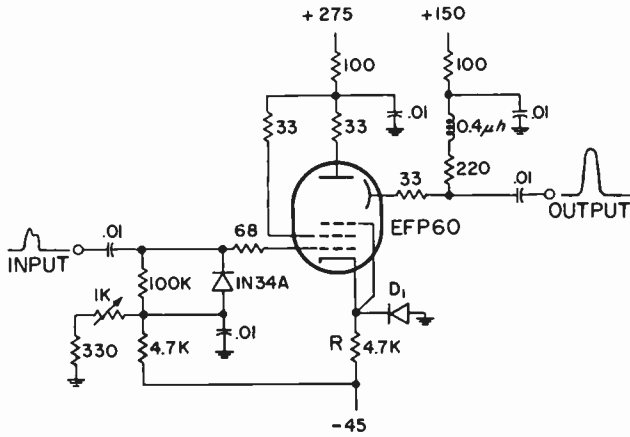
Fig. 1—The limiter amplifier circuit shown above includes decoupling, dc restoration, parasitic suppressors, and shunt peaking. *R* and *D*¹ provide limiting of the output pulses and the grid bias is set so that the tube is normally off.

dynode is limited at about 10 volts positive. The output of this amplifier typically drives a 5842 cathode follower which has an output impedance of 40 ohms and which may be used therefore to drive low impedance loads. Cathode follower power requirements are reduced through the use of positive pulses.

## LOGICAL CIRCUITRY

An important property of semiconductor diodes for use in logical circuitry at high pulse repetition rates is their switching speed. Previous experience with germanium diodes at a 1-megacycle repetition rate has shown that they are limited by the magnitude of their "reverse recovery time" effect.[4-7] This effect is the tendency for a diode to maintain high reverse conductance upon the application of a reverse bias, following forward diode conduction, and is due to the finite time required to remove the carriers from the conducting diode. The "forward recovery time effect," which is the tendency for the diode to maintain a high forward resistance after application of a forward voltage, is also of significance. At 200-ohm impedance levels it is necessary that the resistance of a diode be large relative to 200 ohms for it to be considered "open"; for it to be considered "closed," it is necessary that the diode resistance be small compared to 200 ohms.

With these criteria in mind, the circuits of Fig. 2 were used to measure recovery times. The pulse generator used is the Spencer-Kennedy Laboratory's Model 503 delay-line pulse generator[8] which supplies adjustable

[4] J. H. Wright, "Transient Response Limitations of Various Semi-Conductor Diodes," Natl. Bur. Standards Rep., No. 3638; July 15, 1954.

[5] T. E. Firle, M. E. McMahon, and J. F. Roach, "Recovery Time Measurements of Point-Contact Germanium Diodes," Hughes Aircraft Co. Rep., Res. and Dev. Labs.; June, 1954.

[6] R. H. Kingston, "Switching time in junction diodes and junction transistors," PROC. IRE, vol. 42, pp. 829–834; May, 1954.

[7] M. C. Waltz, "On some transients in the pulse response of point-contact germanium diodes," PROC. IRE, vol. 40, pp. 1483–1487; November, 1952.

[8] R. L. Garwin, "A pulse generator for the millimicrosecond range," Rev. of Sci. Instr., vol. 21, pp. 903–904; November, 1950.
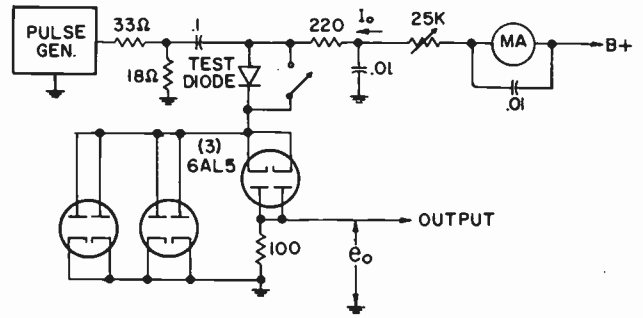
(a)—Circuit used to measure "reverse recovery time" effect.
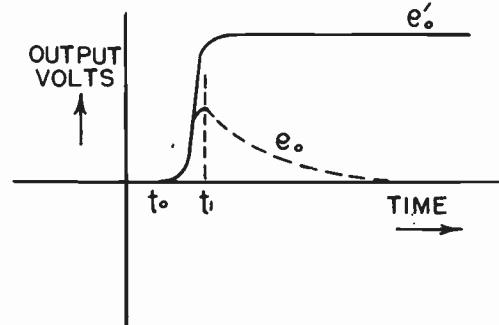


(b)—Typical set of output waveforms, using the circuit of Fig. 2(a), under the conditions of the test diode shorted ($e_0'$) and unshorted ($e_0$).
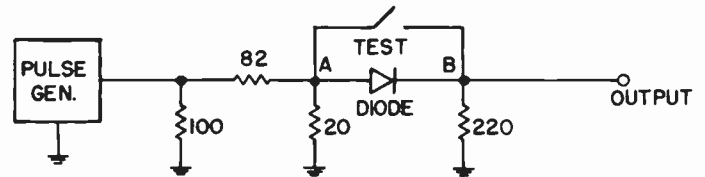


(c)—Circuit used to measure "forward recovery time" effect.

Fig. 2.

width pulses having a rise time of less than 1 millimicrosecond at a pulse repetition rate of 120 cycles per second. Referring to the circuit of Fig. 2(a), which is used to measure the "reverse recovery time" effect, the 25 K potentiometer is adjusted to provide the variable forward current $I_0$, which passes through the diode under test to ground through the two 6AL5's in parallel. The application of a negative pulse from the pulse generator reverse-biases the diode and a transient appears at the output, the amplitude of which is determined by the back resistance of the test diode as a function of time. By also observing the output voltage transient with the test diode shorted, the dynamic resistance of the test diode may be determined. A typical set of observed output transients is shown in Fig. 2(b).

The diode "forward recovery time effect" or the dynamic forward resistance was observed by using the circuit of Fig. 2(c). By observing the output waveforms with a positive pulse applied to the circuit under the conditions of the test diode shorted and unshorted, the forward resistance of the diode as a function of time can be determined.

Many commercial point-contact germanium diodes were found to have satisfactory reverse recovery times when working into a 200-ohm load. The need for a diode with a dynamic forward resistance which is small compared to 200 ohms led to an investigation of junction types, gold-bonded types, and various "high conductance" types of recent origin. In general, it was found that large junction area types were satisfactory in regard to forward resistance, but that their "reverse recovery" characteristics were inferior to the average point-contact diode. The gold-bonded junction types were found to have high forward conductance and, in some cases, satisfactory "reverse recovery" characteristics. Several diode types [*e.g.* 1N308 (CK741),[9] T6G, T-IN283] have been found which give satisfactory recovery characteristics with minimum variation of parameters during service.

The requirements for a logical AND circuit using these diodes can best be understood by referring to a standard 2-input AND circuit, as shown in Fig. 3(a). The initial rate of rise when driving a capacitive load is given by

$$de/dt = i/c.$$

For a rate of rise of 2 volts per millimicrosecond with a capacitive load of 10 micromicrofarads (*e.g.*, tube input capacity) a current of 20 ma is necessary. With currents as high as these and 5–10-volt signal levels, the impedances of the sources driving the AND circuit inputs must be kept low in order to eliminate unwanted responses at the output of the AND circuit when fewer than all of the inputs to an AND circuit rise. Driving source impedances can be reduced by using 5842 cathode followers to drive the AND circuit inputs or by the use of the AND circuit design shown in Fig. 3(b). In this circuit the catcher diodes and resistors compensate for the source impedance, and the rise of the output when one leg rises is only a fraction of a volt.

In the case of the OR circuit which is shown in Fig. 3(c), the output rises when either or both inputs rise. The output rise time is a function of the output impedance of the pulse source at the OR circuit input, the forward impedance of the diode, the OR circuit resistor $R$, and the load impedance. The fall time is determined by the OR circuit resistor $R$, and the load impedance. Rise and fall times of approximately 5 millimicroseconds are typical for the AND and OR circuits.

### PULSE GENERATORS, MULTIVIBRATORS, AND GATES

Test work at vhf pulse repetition rates requires pulse generators, multivibrators, and gating circuits as standard test equipment. As outlined above, few components are available with which to construct these circuits. Because of its possibilities for noninverted power-gain with either grid or cathode input terminals (dynode or
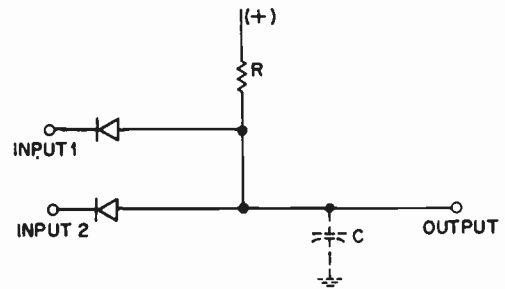


(a)—Standard 2-input AND circuit using semiconductor diodes.
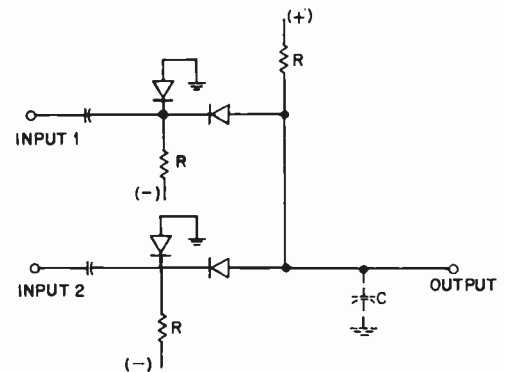


(b)—Modified 2-input AND circuit which effectively reduces the driving source impedances.
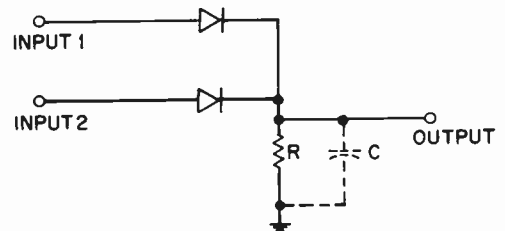


(c)—Standard 2-input OR circuit.

Fig. 3.

plate output terminals), the EFP-60 has been the preferred tube for use in high-speed regenerative circuits. These circuits are of general interest and are outlined below.

The cathode to plate current gain of a secondary emission pentode is utilized in the plate-cathode coupled multivibrator[10] of Fig. 4.

In the astable configuration shown the repetition rate is determined by $C$, $R_0$, $R_1$, and the $B$ voltage, while the pulse width is determined by $C$, $R_0$, and the tube characteristics. The grid may be used as a high impedance sync or trigger terminal, while the dynode may be used as an isolated output terminal supplying positive pulses. The condition for regeneration, neglecting stray capacitance, may be derived from the high-frequency equivalent circuit shown in Fig. 4(b) with the requirement that current-gain from cathode-to-plate-to-cathode be greater than one.

---

[9] No longer in production.

[10] F. H. Wells, "Fast pulse circuit techniques for scintillation counters," *Nucleonics*, vol. 10, pp. 28–33; April, 1952.
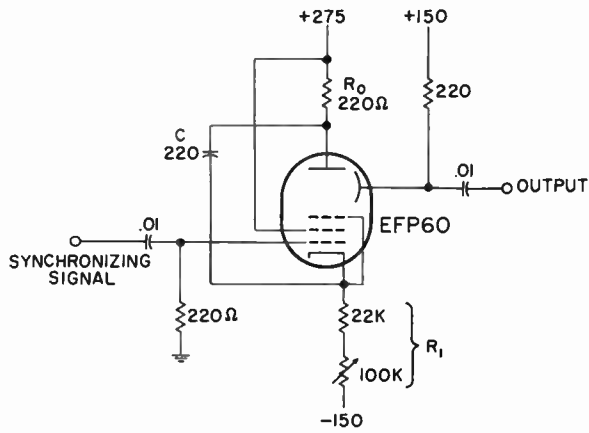
(a)—Plate-cathode coupled multivibrator. In this circuit the secondary-emission current gain of the EFP-60 pentode makes plate-cathode regeneration possible. Repetition rate is varied by means of the 100 K potentiometer.
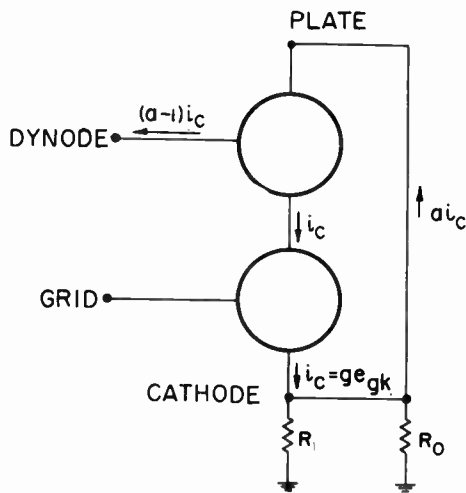


a = SECONDARY EMISSION CURRENT GAIN

g = CATHODE-GRID TRANSCONDUCTANCE

$e_{gk}$ = a.c. GRID-CATHODE VOLTAGE

(b)—High-frequency equivalent circuit of plate-cathode coupled multivibrator circuit.
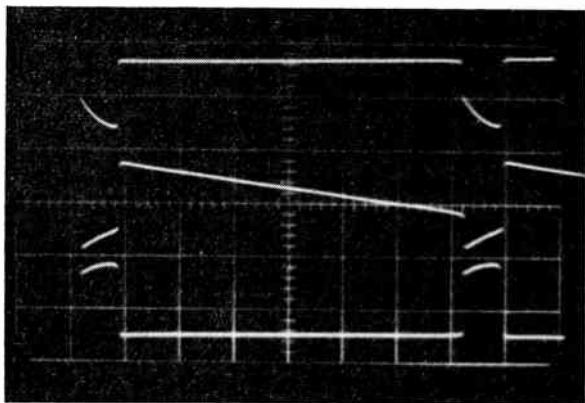


(c)—The plate, cathode, and dynode waveforms (reading from top to bottom) for the circuit of Fig. 4(a). Sweep speed is 2 microseconds per major division.

Fig. 4.

This gives:

$$a\left(\frac{R}{R + 1/g}\right) > 1 \quad \text{or} \quad (a - 1)Rg > 1,$$

where

$R$ = total parallel resistance of $R_0$ and $R_1$,

$a$ = secondary emission current gain,

$g$ = cathode-grid transconductance.

For the EFP-60, $a \cong 5$, $g \cong 5000$ $\mu$mhos, and therefore regeneration occurs only if $R > 50$ ohms.

The dynode grid multivibrator is shown in Fig. 5.[11]
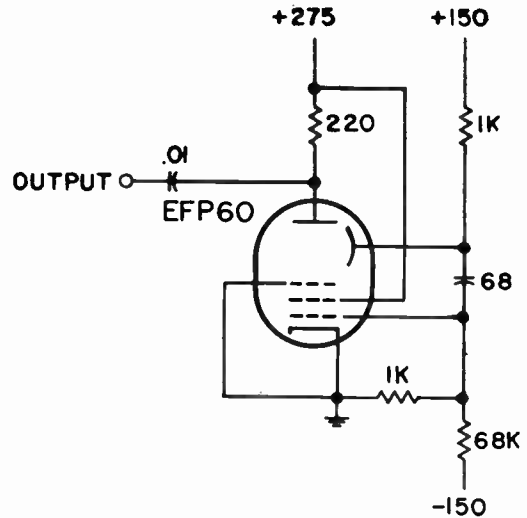


(a)—Dynode-grid coupled multivibrator. In this unstable circuit the noninverting voltage amplification characteristic of the secondary emission pentode is used to obtain multivibrator operation.
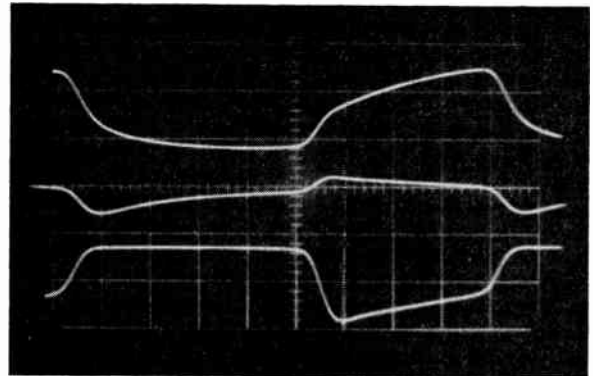


(b)—The dynode, grid, and plate waveforms (top to bottom) for the circuit of Fig. 5(a). Sweep speed is 100 millimicroseconds per major division.

Fig. 5.

A modification of this circuit has been used for synchronized frequency division (see Fig. 9) and time delay. The monostable time delay circuit is shown in Fig. 6. This latter circuit has an "off" period determined by $C$, $E_0$, $R$, and the grid base of the EFP-60. Supply volt-

[11] N. F. Moody, G. J. R. Maclusky, and M. O. Deighton, "Millimicrosecond pulse techniques," *Electronic Eng.*, vol. 24, pp. 214–219; May, 1952.
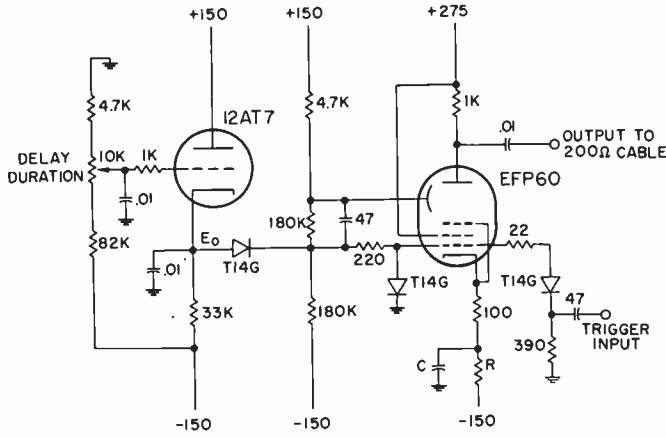
(a)—Dynode-grid monostable multivibrator. The stable operating condition is with the EFP-60 "on" and with the EFP-60 grid at ground potential. Upon application of a negative trigger pulse the tube cuts off regeneratively and the grid falls to $E_0$. EFP-60 remains cutoff until the voltage across $C$ has fallen to a value sufficiently low to allow conduction. At this time the stable operating point is regeneratively re-established.
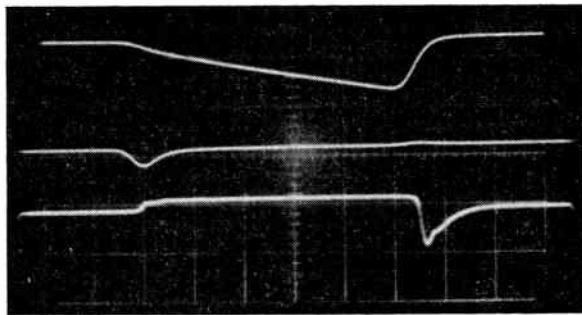


(b)—The cathode, trigger, and plate waveforms (top to bottom) for the circuit of Fig. 6(a). The cathode and trigger waveforms were obtained by using the Tektronix 531 amplifiers, which have a rise time of 35 millimicroseconds, while the plate waveform was obtained by direct connection to the deflection plates through 200-ohm cable. The sweep speed is 100 millimicroseconds per major division.

Fig. 6.

age variations have only a second order effect on this period, since the EFP-60 is used only to charge $C$ between limits set by $E_0$ and ground, and since supply voltage changes which produce a change in current through $R_1$ produce a compensating change in $E_0$.

Pulse rise times in the EFP-60 multivibrator circuits are as short as 10 to 15 millimicroseconds and pulse widths may be as short as 25 millimicroseconds. The 4-millimicrosecond transit time of the EFP-60 accounts for part of the rise time when used in regenerative circuitry. Output amplitudes are from 10 to 14 volts across 200 ohms.

Fig. 7 shows a "regenerative pulse generator"[12] which has been used as a source of short pulses. The operation of the circuit is such that pulses are produced having a width determined by circuit bandwidth, and a repetition rate determined by the total delay time. Repetition rates as high as 80 megacycles, pulse widths

[12] C. C. Cutler, "The regenerative pulse generator," PROC. IRE. vol. 43, pp. 140–148; February, 1955.

as short as 4 millimicroseconds, and amplitudes of the order of 10 volts across 200 ohms have been obtained with this circuit. By injecting a synchronizing signal this circuit may also be used for synchronized frequency division. (See Fig. 9.)
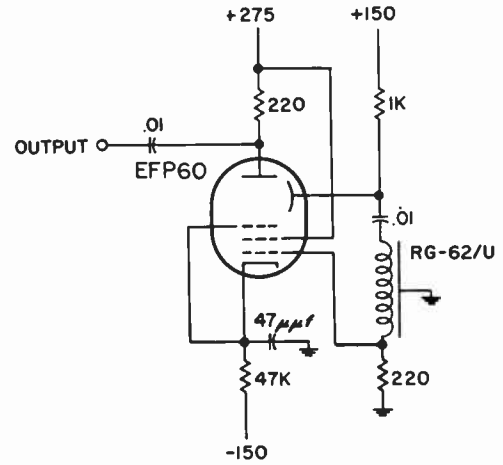


Fig. 7—Regenerative pulse generator. Pulses were produced having a width determined by circuit bandwidth and a repetition rate determined by the total delay time through the tube and the RG-62/U coaxial cable.

A method for generating a group of coded pulses 10 millimicroseconds wide with a 20-millimicrosecond pulse spacing from a single 10-millimicrosecond pulse is shown in Fig. 8. Cathode followers 1 to 5 are successively driven negative by the pulse which travels down the tapped line. External switches set up the current through tubes 1 to 5. The output pulses are mixed in the diode OR circuit and amplified in the following stages.

When working at a 50-megacycle pulse repetition rate synchronized lower frequency pulses are needed for triggering oscilloscopes as well as synchronizing auxiliary circuitry. These pulses are provided by the circuit of Fig. 9, making use of the methods described above.

### RESHAPING

As a pulse proceeds through successive stages of circuitry, both its timing and wave-shape deteriorate. In order to compensate for these effects retiming and reshaping circuitry must be provided. The function of the reshaper is to produce a retimed fixed amplitude output pulse if, and only if, an input pulse is present at a specified "clock" time. The reshaper normally includes input, output and clock terminals, as well as logical circuitry and an amplifying device.

Clock and input timing is set so that the input pulse is at full amplitude at the time of the appearance of the clock pulse. During the coincidence time of the clock and input pulses, circuits are set up which insure the production of a standard output pulse. Since the input, output, and clock pulses are normally of approximately
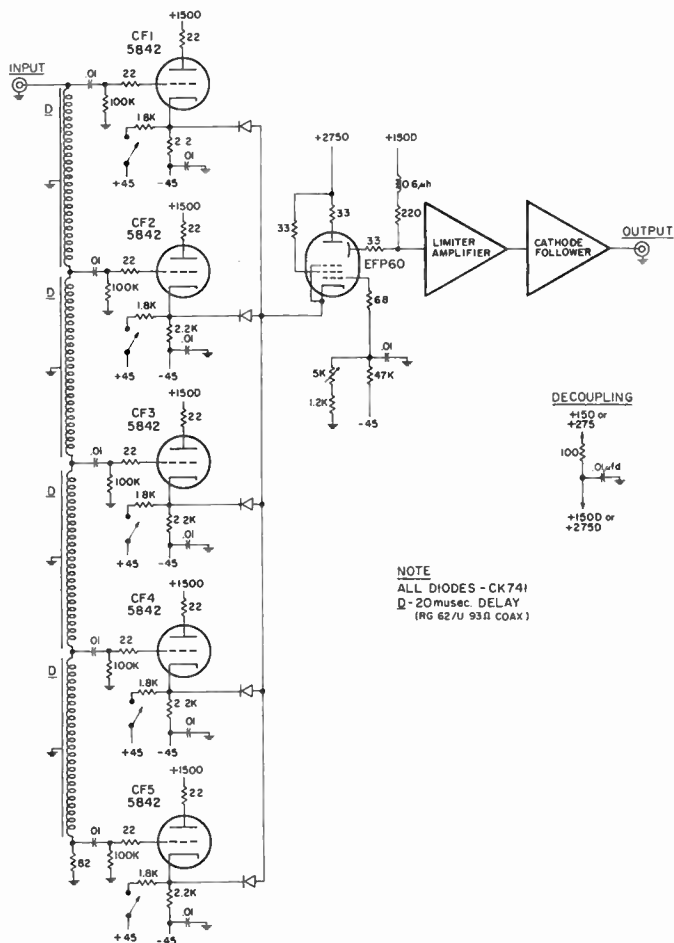
Fig. 8—Circuitry used to generate a group of pulses 10 millimicro-
seconds wide from a single 10-millimicrosecond pulse.
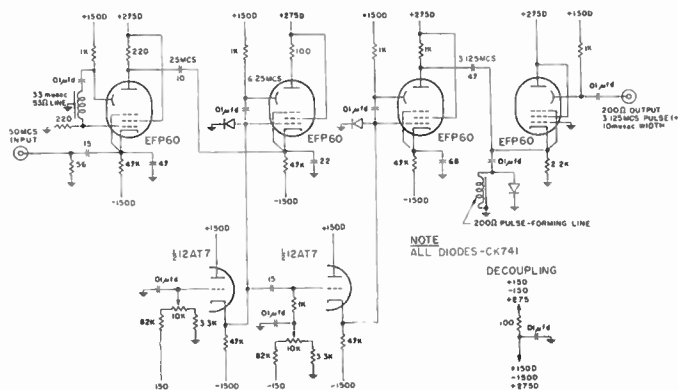
Fig. 9—Multivibrator divider chain used to produce synchronized
pulses at a submultiple of the input frequency.

the same width, and since the input pulse must arrive
early enough with respect to the clock to insure that the
rise of the output is determined by the leading edge of
the clock pulse, a hold circuit must be provided to
maintain the output until the fall of the clock pulse. A
limiting device is generally needed in order to maintain
constant amplitude output pulses.

In the microsecond range these requirements can be

met in many ways.[13-15] For pulses of 10-millimicro-
second width the use of "regenerative broadening"[14] is
made impractical by the magnitude of transit time and
rise-time delays in available amplifiers. Fig. 10 shows a
method of reshaping, adopted here, which avoids the
use of "regenerative broadening." The condenser $C$ is
used for the "hold" operation, while the logical circuitry
provides that the rise and fall of the output pulse are
timed by the clock.
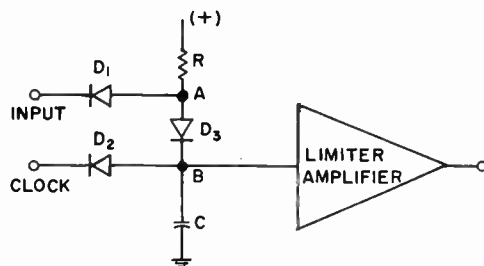
(a)—Reshaper circuit which eliminates the need for
feedback.
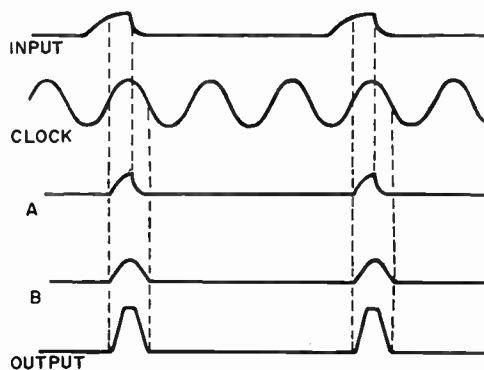
(b)—Idealized reshaper waveforms.

Fig. 10.

## Delay Lines

Electromagnetic delay lines have been used to pro-
vide a storage medium for use with the pulse circuitry
described. Both standard coaxial cables and specially
constructed helically-wound lines have been used for
this purpose. For delays of from 10 to 100 millimicro-
seconds, coax cables such as RG62U, RG58U, etc., or
special helical lines[16] of the type shown in Fig. 11(a) are
satisfactory. For delays greater than 100 millimicro-
seconds the attenuation limitation becomes more seri-

[13] H. Ross, "Arithmetic element of the IBM type 701 computer,"
Proc. IRE, vol. 41, pp. 1287–1294; October, 1953.
[14] R. D. Elbourn and R. P. Witt, "Dynamic circuit techniques
used in SEAC and DYSEAC," Proc. IRE, vol. 41, pp. 1380–1387;
October, 1953.
[15] Q. W. Simkins and J. H. Vogelson, "Transistor amplifiers for
use in a digital computer," Proc. IRE, vol. 44, pp. 43–55; January,
1956.
[16] I. A. Lewis and F. H. Wells, "Millimicrosecond Pulse Tech-
niques," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 42–47;
1954.

ous. Since skin effect attenuation in the frequency range of interest is inversely proportional to wire size, and since characteristic impedance is determined by geometry, the problem is solved for a given impedance by the use of larger diameter coax (*e.g.*, RG-15U), or larger diameter spiral lines. An example of the latter is shown in Fig. 11(b).
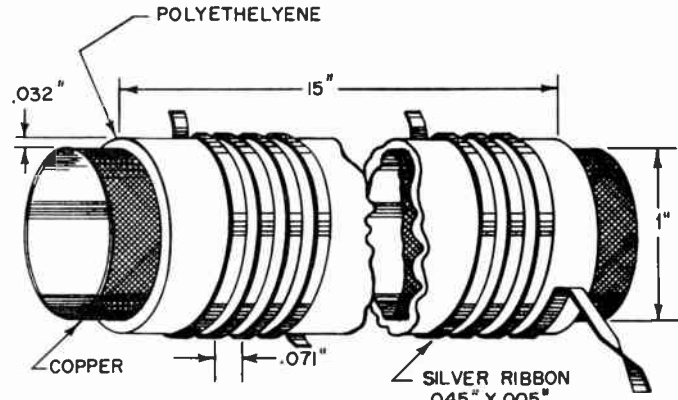


(a)—Helical delay line. The measured impedance is 93 ohms and the delay is 84 millimicroseconds. The measured attenuation is approximately 3 db at 70 megacycles.
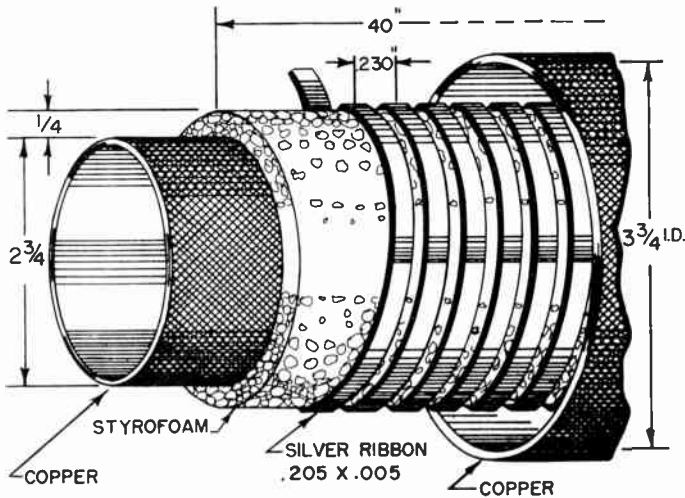


(b)—Helical delay line. The measured impedance and delay are 200 ohms and 160 millimicroseconds respectively. The large structure is necessary to minimize attenuation due to skin effect.[17] The attenuation measured was 2 db at 100 megacycles per second.

Fig. 11.

## Construction and Testing

When working with pulses in the millimicrosecond range extreme care is necessary in pulse handling. As a typical example of the problems involved, a standard half-watt composition resistor has a shunt capacity of $\frac{1}{2}$ micromicrofarad, and therefore has a time constant of $R/2000$ millimicroseconds. Therefore, a resistance voltage divider consisting of a 100 K and a 10 K resistor (of the same type) would have an output waveform as shown in Fig. 12 for a step function input. This example

[17] L. H. Thomas, "Propagation in Helical Lines" (unpublished).
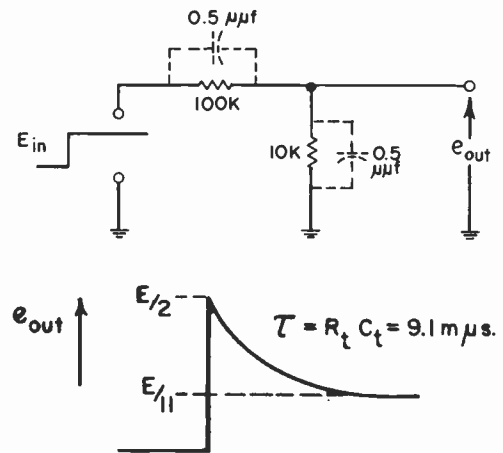


Fig. 12—Resistive voltage divider made up of standard half-watt composition resistors. Output waveform for step function input.

ignores all additional stray capacities but illustrates the fact that one cannot accept the low-frequency values of components when working with pulses in the millimicrosecond range.

A further example is the problem encountered with small values of resistance where a series inductance of 1 millimicrohenry produces a time constant of $1/R$ millimicroseconds. It is desirable that all connections of over a few inches in length be made through terminated coaxial cables.

The Tektronix 517 Oscilloscope, having an inherent rise time of 7 millimicroseconds, is used for test work at low repetition rates. The Tektronix 531, with provision for direct connection to the deflection plates (see Fig. 13)
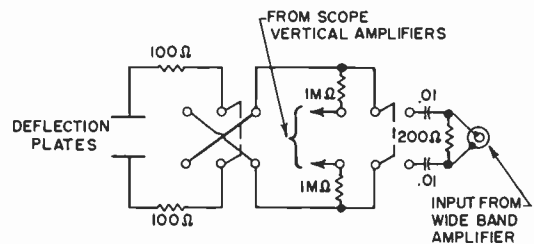


Fig. 13—Switching for oscilloscope defection plates. The 100-ohm resistors prevent high-frequency ringing by damping the resonant circuit consisting of lead inductance and deflection plate capacity. When the output of the wide band amplifier is to be observed, connection through 1-megohm resistors to the scope vertical amplifiers is maintained in order to utilize the oscilloscope vertical centering control.

is used for observation of high repetition rate phenomena. A cathode follower probe (*e.g.*, Tektronix 517 Probe) driving a wide band distributed amplifier (*e.g.*, Hewlett-Packard 460B or Instruments for Industry 500) is used to drive the scope deflection plates. The measured rise time, when using a cathode follower probe feeding the wide-band amplifier, is 3 millimicroseconds and the deflection sensitivity is approximately 5 volts/cm.

## Experimental Arithmetic Unit

The components described have been applied to the design of an arithmetic unit which performs binary addition, multiplication, and dynamic storage at a pulse repetition rate of 50 megacycles per second. A block diagram of the system is shown in Fig. 14.
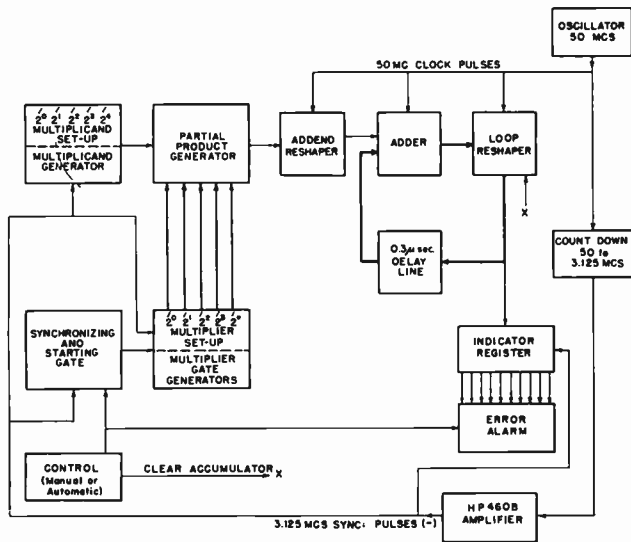


Fig. 14—Experimental arithmetic unit. The 50-megacycle oscillator and the count-down circuitry provide "clock" and "sync" signals respectively. When no input is present from the addend reshaper, the adder, loop reshaper, and delay line form a dynamic storage element which retains any information present until receipt of a clear signal at X. When the starting gate is energized the multiplier gate generator scans a set of five switches on five successive "sync" cycles and gates the output of the multiplicand generator to the addend reshaper. On each of these five gates the multiplicand output is shifted by 20 millimicroseconds so that the final result in the circulating loop is the binary product of the multiplicand and the multiplier.

The oscillator provides 10-volt sine wave clock signals at a 50-ohm impedance level at appropriate points throughout the system. A series of EFP-60 multivibrator circuits are used as frequency dividers (Fig. 9) to produce pulses 10 millimicroseconds wide at a 3.125 megacycle per second repetition rate. These pulses are amplified by the Hewlett-Packard 460B wide-band amplifier, and will hereinafter be referred to as the "sync" signal. The total delay time around the dynamic storage loop (shown in heavy lines in the block diagram) is made equal to the period of the "sync."

The multiplicand generator, Fig. 8, produces a 5-bit (binary digit) multiplicand with a bit spacing of 20 millimicroseconds and a group repetition rate equal to that of the "sync." The 5-bit multiplicand is fed to a delay

line tapped at 20-millimicrosecond intervals in the partial product generator. The outputs from the successive taps of this delay line are selected during successive "sync" periods by the multiplier gates, which in turn are conditioned by the multiplier set-up keys. The binary adder consists of logical AND and OR circuits, limiter amplifiers, and cathode followers of the types previously described and forms, with the delay line and loop reshaper, an accumulator which stores the result of the five bit by five bit multiplication.

## Discussion

The use of higher pulse repetition rates will undoubtedly be motivated by the desire for faster digital computers. The development of higher speed random access memories and the desire to perform certain calculations involving real time have previously been motivating factors for the development of higher speed arithmetic units. The components and techniques described herein utilizing a 50-megacycle pulse repetition rate will enable one to perform a complete multiplication in a time comparable to the access time of the fastest available random access memories.

When using a 50-megacycle pulse repetition rate the finite velocity of signal propagation (*i.e.*, approximately one foot per millimicrosecond) makes careful construction necessary in order to maintain accurate timing between various signals throughout a system, but at the same time it allows one to easily perform delay functions, and therefore arithmetic shifting operations.

Although it appears that the use of vhf pulse repetition rates for the arithmetic sections of digital computers is worth real consideration, the use of these pulse rates for an entire system is somewhat impractical now because of the need for special components such as the EFP-60 and 5842 vacuum tubes. These tubes are costly, bulky, and inefficient. However, it is hoped that the present efforts toward development of high-frequency transistors will result in desirable amplifying devices for vhf pulse repetition rates, and therefore make the prospect of the development of entire systems using such rates more attractive.

## Acknowledgment
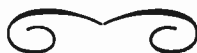
# Back-Scattering Characteristics of the Sea in the Region from 10 to 50 KMC*

J. C. WILTSE†, ASSOCIATE MEMBER, IRE, S. P. SCHLESINGER‡, MEMBER, IRE, AND C. M. JOHNSON§

*Summary*—Measurements of radiation back scattering from the ocean have been made over the frequency range from 10 to 50 kmc by means of several cw, Doppler, microwave systems operating simultaneously. The systems were mounted on the bow of a ship and various antenna depression angles from 0° (horizontal) to 90° were used. Data were obtained for vertical, horizontal, and circular polarization, and for the cross-polarized component of vertically polarized radiation. Horn antennas of 20-db gain were used at all of the operating frequencies and, in addition, higher and lower gain antennas were used at certain selected frequencies. Signals recorded from a variety of sea conditions have been used to calculate the back-scattering cross section per unit area, $\sigma^\circ$, and these calculations indicate that $\sigma^\circ$ is nearly constant with microwave frequency and increases as the antenna depression angle is increased. The Doppler frequency characteristics of the returned signals have also been obtained from these measurements. A comparison is made between the experimental results and the characteristics predicted from simple scattering mechanisms.

## INTRODUCTION

THE OPERATION of microwave systems close to the ocean surface is seriously handicapped by the back-scattered radiation from the surface, and the desire to overcome this interference continually stimulates sea scattering studies. Over the past several years quantitative measurements of the magnitude of sea return in the microwave region below 10 kmc have been made by various groups,[1-3] and considerable progress has been made toward developing a theory for this type of scattering. Except in a few instances, these lower frequency data were obtained at small grazing angles, and pulsed systems were used. In the higher frequency microwave region, previous quantitative measurements appear to be almost nonexistent.

To obtain quantitative information on sea-scattering in the frequency range from 10 to 50 kmc, two sea tests have been conducted by the Radiation Laboratory. In these tests the back-scattering measurements were made with several (usually three) cw, Doppler, microwave systems operating simultaneously from the bow of a moving ship. In the first test, conducted in December, 1954, microwave frequencies from 9.6 to 38 kmc were used, and data were obtained at antenna depression angles

from zero (horizontal) to 80°, but only for vertical, polarization.[4] In the second and more conclusive test, conducted in November, 1955, the frequency range was extended to 48.7 kmc and data were recorded for vertical, horizontal, circular, and cross polarization over the full range of depression angles from 0 to 90°. Also during this second test, several other parameters such as antenna beam width, height above the ocean, and azimuth with respect to the wind direction were varied individually in order to determine or verify the effect of each on the back-scattering results.[5] A wide range of sea conditions was encountered during these two tests since the first voyage lasted two days and covered a 1000-mile distance, and the second voyage extended over a period of a week and over a 3000-mile distance.

From the large amount of data obtained, the dependence of both the back-scattering cross section of the ocean surface and the Doppler frequency of the signals on the various parameters has been determined, and these relations form the basis for drawing several conclusions about the nature of the scattering mechanism.

## EXPERIMENTAL

### Microwave Systems

In these tests microwave systems of the type shown in Fig. 1 were operated at 9.6, 15, 24, 26, 35, 38, and 48.7 kmc. Each system uses a single reflex klystron operating cw to supply both a reference signal and a transmitted signal. The reference signal is applied to the crystal detector through two directional couplers, and for optimum detection sensitivity this reference power on the crystal should be between −20 and −30 dbm. When the transmitted signal illuminates a moving target, the reflected signal of constantly changing phase is mixed in the crystal detector with the constant-phase reference signal producing an audio output at the phase-modulation frequency. This phase-modulation frequency, $f$, produced by a target moving with a velocity, $V$, is given by

$$f = \frac{2V}{\lambda} \cos \theta, \qquad (1)$$

where $\lambda$ is the wavelength of the microwave radiation and $\theta$ is the acute angle between the propagation direction of the radiation and the target's direction of mo-

[1] D. E. Kerr, "Propagation of Short Radio Waves," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 481–527; 1951.
[2] M. Katzin, "Back scattering from the sea surface," 1955 IRE CONVENTION RECORD, Part 1, pp. 72–77.
[3] F. C. MacDonald, "Correlation of radar sea-clutter on vertical and horizontal polarizations with wave height and slope," 1956 IRE CONVENTION RECORD, Part I, pp. 29–32.

[4] C. M. Johnson, S. P. Schlesinger, J. C. Wiltse, and C. W. Smith, "Sea Scattering Measurements in the Region from 9.6 to 38 kmc," Tech. Rep. No. 27, Johns Hopkins Univ. Rad. Lab.; 1955.
[5] J. C. Wiltse, S. P. Schlesinger, and C. M. Johnson, "Back-Scattering Characteristics of the Sea in the Region from 10 to 50 kmc," Tech. Rep. No. 35, Johns Hopkins Univ. Rad. Lab.; 1956.
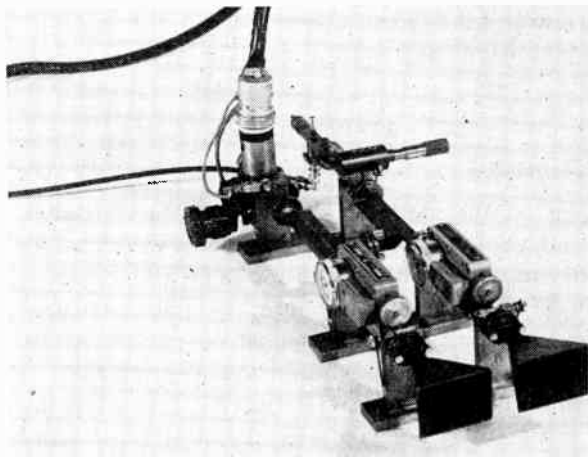
Fig. 1—A typical synchronous detection system as used for the sea-scattering measurements.

tion. This phase-modulation frequency is, of course, the same as the Doppler frequency produced by a moving scatterer.

Detection of the scattered signal by this synchronous method obeys a linear law (similar to superheterodyne detection), and thus the audio output amplitude is proportional to the square root of the received rf power. As used for these tests, with an audio bandwidth from 100 cps to 100 kc, this detection system has an average sensitivity of −90 dbm. A more complete description of this type of detection is given in an earlier paper.[6]

The various klystrons used for the different frequencies supplied between 20 and 200 milliwatts of power, except at 48.7 kmc where only about 0.2 milliwatt was available. At the lower frequencies (below K band) IN23 crystal detectors were used and at the higher frequencies IN26's were used.

Many different arrangements of microwave components were considered, such as a hybrid-tee system requiring only one horn, but the directional coupler arrangement using separate, identical, transmitting and receiving horns proved to be the least critical and most reliable system for field-test use. For the majority of the measurements horn antennas of optimum proportions and 20-db gain were used at each microwave frequency. In addition, 28-db optimum horns were used for some of the runs at 24 and 35 kmc, and 14.8-db optimum horns were sometimes used at 24 kmc. The 3-db beamwidth is about 7° for a 28-db horn, about 17° for a 20-db horn, and about 28° for a 14.8-db horn. Hence, the horns used covered a 4 to 1 ratio of beamwidths.

Although groups of runs were taken with horizontal, cross, and circular polarization of the radiation, systems were normally set up to operate vertically polarized. To obtain horizontal polarization, short 90° waveguide twists were installed behind both the transmitting and the receiving horns. For measurements of cross polariza-

[6] M. E. Brodwin, C. M. Johnson, and W. M. Waters, "Low-level synchronous mixing," 1953 IRE CONVENTION RECORD, Part 10, pp. 52–57.

tion, a twist was installed only behind the receiving horn, leaving the transmitted radiation vertically polarized. Circular polarization was produced by using circular waveguide containing a quarter wavelength dielectric wafer oriented at 45° to the E vector in the guide. With the dielectric wafers parallel in both the transmitting and the receiving arms of a two-horn system, a back-scattered signal which has undergone a single reflection (single bounce) from a flat plate normal to the direction of propagation is rejected; but a back-scattered signal which has undergone a double specular reflection (double bounce) from two flat plates oriented at the proper angles is accepted. With the wafers normal to each other in the two arms, the inverse is true. Both wafer orientations were used in the circular-polarization systems operated at 9.6 and 24 kmc. More exactly speaking, the polarization was not circular but elliptical, with ellipticities of about 1.3 at 9.6 kmc and 1.2 at 24 kmc.

### Shipboard Test Procedure

The tests were conducted aboard C2 type merchant ships which were making regular scheduled ocean trips. The first voyage was from Baltimore to Miami, Fla., and the second was from New York City to Puerto Rico and back to Baltimore. Data were taken on six different days for a total of 86 separate runs under a variety of sea conditions. In all cases the ships were in deep water, generally several hundred miles from the nearest land.

The microwave systems were mounted on a movable board which could be accurately set to give the desired antenna depression angle. Fig. 2 shows a photograph of
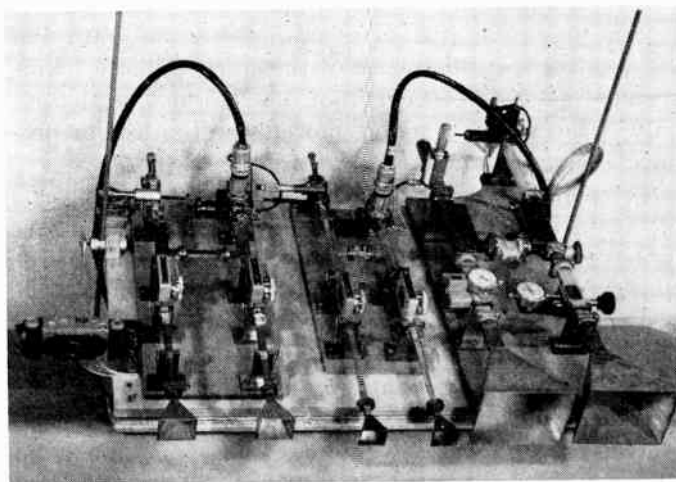


Fig. 2—Photograph of sea test microwave systems, mounting board, and stereographic camera.

the board with three microwave systems and a stereoscopic camera mounted on it. This assembly was leveled carefully and rigidly clamped to the bow of the ship. The bow rode at a constant average height above the water for long periods of time, and over all of the runs this height varied from 37 to 47.5 feet.

For the usual run three microwave systems were operated simultaneously, and the audio-frequency outputs of the crystal detectors were amplified by identical preamplifiers and recorded on a four-channel tape recorder (Ampex Model 307). The fourth channel of the recorder was used for timing markers and verbal information describing the runs. For each run the sea return signal was recorded for approximately three minutes which allowed enough time for the ship to pass through many wave crests and troughs.

Each microwave system was calibrated before and usually after each group of runs by facing the system inboard toward a vibrating corner reflector of known cross section and recording the back-scattered signal on the appropriate channel for comparison with the sea return signals. Calibrated attenuators were used during this procedure to reduce the signal level to approximately that expected from the sea. Once calibrated, the systems and the mounting board were turned to face forward and locked into position without interrupting the operation of the klystrons. Each klystron power supply was provided with a square-wave modulation voltage having its positive side clamped to the reflector voltage level. When this modulation was applied to the reflector, the output level due to the modulated reference signal could be observed. Thus, frequent checks could be made on the klystron power output and the detector sensitivity.

In order to determine the dependence of back-scattered signal on height above the surface a group of successive runs was made at the bow height (47.5 feet) and at a height of 26.5 feet. The lower height was achieved by lowering the microwave system from the bow on a 3-inch-diameter aluminum tube. The 24-kmc system operating at depression angles of 30°, 60°, and 90° was used in this phase of the work. The 48.7 kmc system was also operated from the 26.5-foot height but only because the available power was too low to produce a signal at the bow height.

To compare the back scattering from various aspects of the wave pattern, one ship executed a large square (about three miles on a side), traveling downwind, crosswind, and upwind. During the maneuver microwave systems were operated simultaneously at 9.6, 24, and 35 kmc with an antenna depression angle of 60°. Only vertically-polarized radiation was used.

Visual indications of sea conditions were recorded regularly, and, during the second test, wind speeds and directions were obtained from a windometer mounted above the bow. Photographs of the sea surface were taken with a 35-mm camera, a 16-mm movie camera, and a stereoscopic camera. Many of the stereophotographs were taken with the camera fastened to the mounting board (see Fig. 2). The depression angle for the camera was then the same as for the microwave horns, and the actual area of the sea producing the back-scattering was photographed. To obtain an estimate of wave height and length, stereophotographs of graduated scales were taken from the roof of a building under the same conditions of height and angle as used at sea.

## THEORETICAL

### Doppler Frequency

Eq. (1) gives the Doppler frequency observed when there is relative motion between the microwave system and a rigid target that intercepts only a small increment of solid angle when seen by the antenna. The ocean surface differs in two important respects from these criteria. First, the surface intercepts radiation over as large a solid angle as the antenna pattern permits. Second, the sea is, of course, by no means a rigid target. The many random scatterers that make up the surface have a wide distribution of instantaneous velocities. Both effects give rise to a more complicated Doppler signal.

In order to estimate the effect of target size on the Doppler frequency, the variation of the intensity of the returned signal with angle off the horn axis must be considered. The intensity depends upon both the antenna pattern function and the change in distance to the ocean surface as a function of the angle off axis. If the angle in the vertical plane subtended by the antenna beam is sufficiently large and the depression angle, $\theta$, of the horn axis sufficiently small, then incremental scatterers located on the beam center-line return lower intensity signals than those which are off-axis but closer to the antenna. For example, in the case of 20-db gain horns an analysis of the effects of range and antenna pattern on the back-scattered power shows that for an antenna depression angle of 0° the region from which appreciable signal is returned is more extended than the normal antenna beam width and the center of this region lies at an "effective depression angle" of approximately 15° below the horizontal. Similarly, the effective depression angle for $\theta = 15°$ is approximately 18°. For larger values of $\theta$ this off-axis effect is negligible. That these are the correct values for the effective depression angles is verified by the measured values of Doppler frequency.

The Doppler frequency produced by relative motion between a cw microwave system and a scattering element is determined by the relative velocity component along the direction from the microwave antenna to the scatterer. The frequency observed depends upon components of velocity due to the oscillatory motions of the scattering facets as well as the steady component due to the (horizontal) motion of the ship with respect to the main water mass. (The vertical motion of the ship had only a negligible effect on the Doppler frequency since the amplitude of this motion was generally small and the period was always greater than five seconds.) As long as the steady velocity component (given by $V \cos \theta$, where now $V$ is the ship's velocity) is at least as large as the peak velocity component due to the oscillatory facet motions, the average Doppler frequency is given by (1), since the oscillatory components average out. Thus (1)

is valid if $V \cos \theta$ is sufficiently large, which implies that the antenna depression angle must not be in the vicinity of 90°. Furthermore, the antenna beam width must also be small or $\theta$ must be replaced by a larger "effective depression angle" as discussed above.

Near normal incidence, where $V \cos \theta$ becomes smaller than the radial velocity components contributed by the motions of the individual scatterers, oscillatory components no longer average out since the frequency generated by a given radial velocity depends only on the magnitude and not on the direction. For example, at $\theta = 90°$, where $V \cos \theta = 0$, random scatterers on the surface will produce the same value of frequency at the detector output whether rising or falling at a given velocity.

If it is assumed that the vertical velocities change according to some simple relation, such as a sinusoidal function of time, then an expression can be written which describes the Doppler frequency for all depression angles. If the mean peak vertical velocity is given by $V_v$, then under the above assumption, the instantaneous Doppler frequency is given by

$$f = \left| V \cos \theta + (V_v \sin kt) \sin \theta \right| \cdot \frac{2}{\lambda} \qquad (2)$$

where $k/2\pi$ is the frequency of the vertical motion and $t$ is the time. The absolute value is used since the frequency is always positive. If $V_v$ is assigned a value of 4 knots, this equation gives a frequency for $\theta = 90°$ that is approximately equal to the measured values at this angle.

## Cross Section

For radar targets which intercept only a small portion of the antenna beam, the back-scattering cross section, $\sigma$, is ordinarily obtained from a solution of the well-known radar equation,

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \qquad (3)$$

where $P_r$ is the back-scattered power received, $P_t$ the transmitted power, $G$ the antenna gain, $\lambda$ the wavelength of the radiation, and $R$ the range to the target. However, for targets of practically infinite extent, such as the ocean surface, $\sigma$ is no longer independent of the measuring system, but becomes a function of range, antenna pattern, and depression angle. Therefore, it is necessary to consider the ocean target in detail and arrive at a more universal quantity that measures the scattering intensity.

For the beam widths and frequencies used in these measurements the ocean target can be resolved into a number of individual scatterers that are independent of each other.[1] The back-scattered power from the $j$th scatterer when the horn antennas are set at a depression angle $\theta$ is given by

$$(P_r)_j = \frac{P_t G^2 \lambda^2 f^2(\beta_j, \phi_j) \sigma_j}{(4\pi)^3 R_j^4}, \qquad (4)$$

where $\sigma_j$ is the cross section of the $j$th scatterer, $R_j$ the distance to the $j$th scatterer, and $f(\beta_j, \phi_j)$ the antenna power function evaluated at the $j$th scatterer. The angles $\beta$ and $\phi$ are measured from the horn antenna axis in the electric and magnetic planes, respectively.

The sum of these individual contributions gives the total received power; and, if the number of scatterers is large, the summation can be replaced by an integral. If these scatterers are assumed to be distributed uniformly and homogeneously over the surface of the ocean then for the case of narrow beam widths and reasonably large depression angles ($\theta$) the back-scattered power received from the ocean target becomes

$$P_r = \frac{P_t G^2 \lambda^2 \Theta \Phi}{(4\pi)^3 R^2 \sin \theta} \cdot \sigma°(\theta), \qquad (5)$$

where $\Theta$ and $\Phi$ are approximately equal to the $E$-plane and $H$-plane beam widths, respectively, and $\sigma°(\theta)$ is the back-scattering cross section of the ocean surface per unit area.

Comparison of (5) with (4) shows that

$$\sigma°(\theta) = \frac{\sigma \sin \theta}{R^2 \Theta \Phi} \cdot \qquad (6)$$

The quantity $\sigma°$, although independent of range and beam width, remains a function of depression angle, microwave frequency, sea condition, direction of polarization, and azimuth relative to the wave pattern on the ocean surface. The effect of each of these quantities on $\sigma°$ was measured during the sea tests.

## Data Analysis and Results

As mentioned above, the sea-return signals were recorded for about three minutes at each antenna depression angle. Calibration signals returned from a corner reflector of known cross section were recorded with each related group of runs; and likewise, system noise levels, with the transmitted microwave power attenuated but with full reference power on the crystal were also recorded. Together these recordings provided the basis for obtaining back-scattering cross section values. Timing markers (at about 0.3-second intervals) which were simultaneously applied to the fourth channel of the recorder while recording the sea-return signals on the other three channels made possible a comparison of Doppler frequencies and of signal amplitudes at simultaneous times for the different microwave frequencies used during a given run.

In the laboratory the tape-recorded signals were displayed on an oscilloscope, and a ten-second section of each run was photographed using a film transport speed sufficient to resolve the Doppler cycles. Figs. 3 and 4 show typical film samples (the dashed lines shown on the film strips are 120 cps markers which should not be con-

fused with the timing markers mentioned above). By means of a microfilm reader signal amplitudes (peak-to-peak) and Doppler frequencies were read from the films at thirty or, in a few instances, one hundred points equally spaced throughout each ten-second film sample. These readings were used to obtain the average amplitude and average frequency of each run. In addition to the average amplitude value, the largest peak-to-peak amplitude read from each film sample was also recorded and later used to calculate a "maximum value" of cross section.
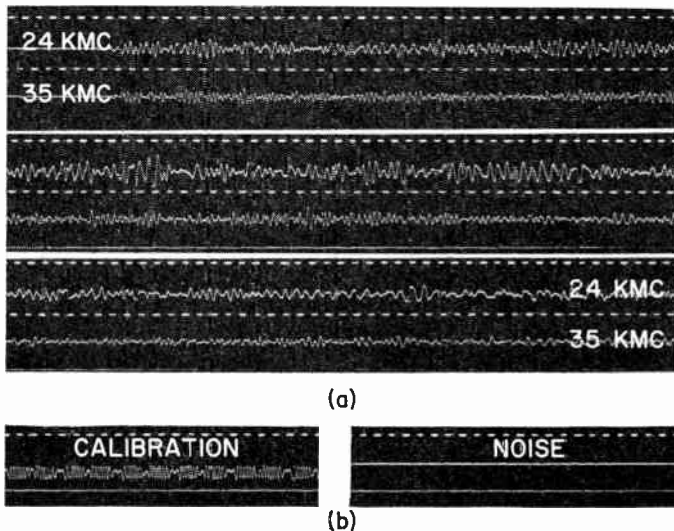


Fig. 3—Film trace samples for (a) simultaneous signals at 24 and 35 kmc obtained at a depression angle of 80°, and (b) the corresponding calibration and noise level for the 24-kmc signal.
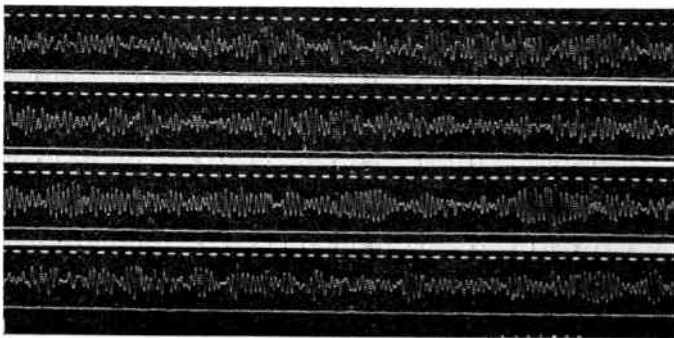


Fig. 4—A typical signal trace obtained at moderate depression angles. In this particular case $\theta = 45°$ and the microwave frequency is 9.6 kmc.

As a check on the film-read values, the amplitudes and Doppler frequencies were also measured by electronic techniques. To measure amplitudes the recorder output was connected to a circuit which integrated signals, calibrations, or noise levels for periods of one to two minutes. Later, these values were normalized to a common time base and the average values obtained. The Doppler frequencies for those runs having a sufficiently high signal-to-noise ratio were obtained by feeding the recorder output into an electronic frequency counter and counting for the full length of the run.

In order to calculate cross section values from the amplitudes measured by either of the above techniques, (3) was applied to both the signal from the sea and the signal from the corner reflector. Since the radar cross section of the latter was known, $\sigma$ (for the sea) could be calculated without knowing the transmitter power or system sensitivity. The signal amplitudes obtained from the electronic integrator generally tended to be slightly higher than the film read values,[7] and because of this difference each set of amplitude readings was used separately to calculate $\sigma$ values.

Once $\sigma$ has been obtained, the value of $\sigma°$ may be calculated from (6). Fig. 5 shows typical curves of $\sigma°$ vs
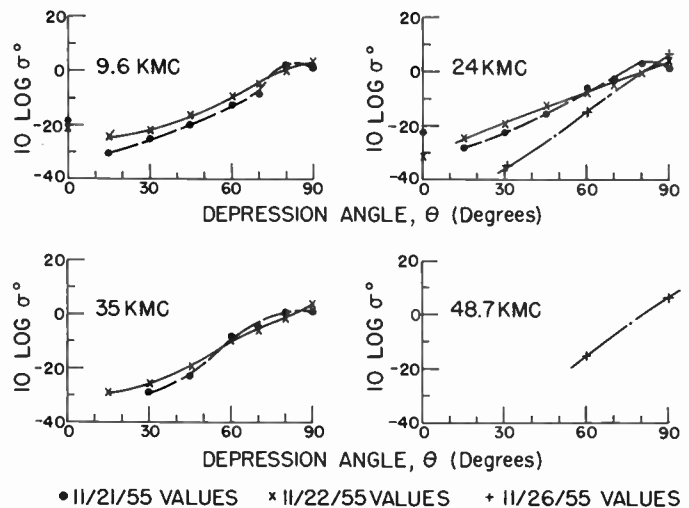


Fig. 5—Plots of 10 log $\sigma°$ vs depression angle for vertical polarization at four different microwave frequencies. The measurements for a particular day and depression angle were made simultaneously at all of the microwave frequencies indicated. (Sea moderate, wind speed 16 knots on November 21; sea moderately heavy, wind speed 26 knots on November 22; sea fairly calm, wind variable 4 to 10 knots on November 26.)

depression angle for vertical polarization.[8] In general, $\sigma°$ increases with depression angle except near normal incidence. These plots allow a comparison to be made of the effect of sea condition on $\sigma°$ since the sea was moderate (occasional whitecaps, average wind speed 16 knots) on November 21, moderately heavy (whitecaps plentiful, average wind speed 26 knots) on November 22, and fairly calm (no whitecaps, wind speed varying from 4 to 10 knots) on November 26. The curves show that rougher sea conditions produce larger values of $\sigma°$ at low angles but slightly smaller values at angles near normal incidence. The dependence of $\sigma°$ on depression angle and sea state is in qualitative agreement with earlier high

[7] This point is discussed further in the concluding section.

[8] It was mentioned earlier that if $\theta$ is close to zero the region from which appreciable signal is returned is more extensive than when $\theta$ is larger. In calculating $\sigma°$ values for $\theta = 0°$ an "effective" 3-db antenna beam width of 25° was used for $\Theta$, an effective depression angle of 15° was used, and the normal H-plane beam width was assumed to be unchanged. In view of these approximations the $\sigma°$ values have been plotted at $\theta = 0°$, but the curves have not been extended to these values.

angle measurements made at 10 kmc[9] and at 1.25 kmc.[10] The actual values given in these earlier references sometimes vary by a factor of 1000 between different experiments, and hence it is difficult to obtain a quantitative comparison.

The polarization dependence of $\sigma°$ is illustrated in Fig. 6, where values of $\sigma°$ calculated from amplitudes measured with the integrator are plotted vs depression angle. Comparison of the horizontal and vertical polarization results shows that the horizontal polarization return is generally lower than that from vertical polarization except at normal incidence. This trend is in general agreement with earlier results at 10 and at 1.25 kmc.[11] The cross-polarization values of $\sigma°$ are much lower than those for any of the other polarizations. The $\sigma°$ values from the circular polarization measurements in which single-bounce signals were accepted are nearly as large as those for vertical polarization, but the $\sigma°$ values from the circular polarization measurements in which double-bounce signals were accepted are considerably less.
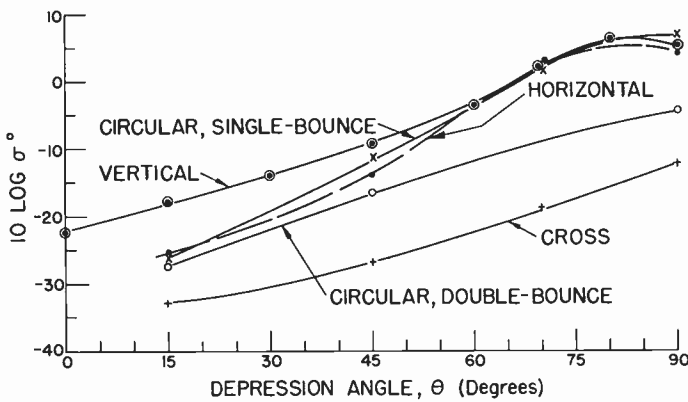


Fig. 6—Plots of 10 log $\sigma°$ vs depression angle for vertical, horizontal, circular, and cross polarization at 24 kmc. The data were taken on November 22, 1955. (Sea moderately heavy, wind speed 26 knots.)

The results of the measurements made to compare the scattering from the upwind, downwind, and crosswind aspects of the waves show that for $\theta = 60°$, $\sigma°$ is almost independent of wave aspect. This constancy is true at all three microwave frequencies. Unfortunately, 60° was the only angle used during this part of the test.

The values of $\sigma°$ for vertical polarization obtained during three separate days at sea have been plotted vs microwave frequency in Fig. 7. It can be seen that for a given sea state $\sigma°$ is nearly independent of frequency at large values of $\theta$. Approximately the same result was found to be true for the other polarizations also.

The expression for the back-scattered power given by (5) states that $P_r \sim (1/R^2)$ if all conditions except range (or height above the surface) are maintained constant.

[9] Kerr, *op. cit.*, p. 505, Figs. 6-13 and 6-14.
[10] MacDonald, *op. cit.*, p. 32, Fig. 2.
[11] Kerr, *op. cit.*, p. 505, Fig. 6-13; and MacDonald, *op. cit.*, p. 31, Table III, and p. 32, Fig. 1.

According to this relation $P_r$ at 26.5 feet should be 5.1 db greater than $P_r$ at 47.5 feet. At depression angles of 60° and 90° the average measured difference in signal level at the two heights was found to be 6 db from the film average readings and 4.2 db from the integrator values, which agree with the $1/R^2$ dependence within our probable experimental error.
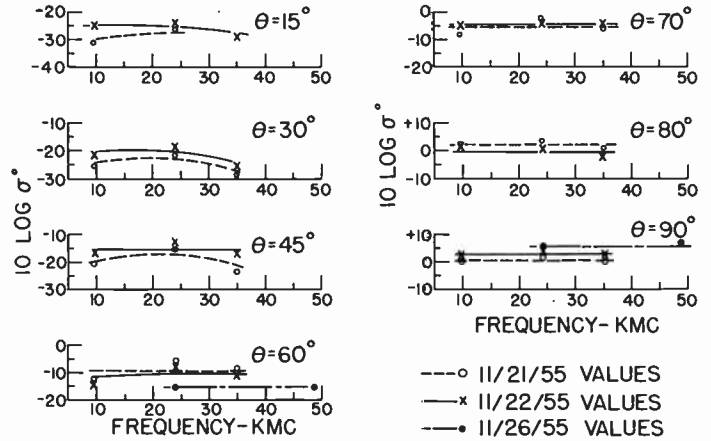


Fig. 7—Variation of 10 log $\sigma°$ with microwave frequency for different values of $\theta$ and for vertical polarization.

Eq. (5) also states that if only the horn gain is varied, then $P_r \sim G^2\Theta\Phi \sim G$. Therefore, $P_r$ with the 28-db gain horns should differ by 13.2 db from the value of $P_r$ with the 14.8-db gain horns. At $\theta = 90°$ the film average readings give a difference of 12.2 db and the integrator values give a difference of 14.6 db, which again are within our probable experimental error. At other depression angles the measurements indicate that the proportionality between $P_r$ and $G$ is valid for $\theta$ values greater than about 50°.

To determine whether any correlation exists among either the instantaneous amplitudes or instantaneous Doppler frequencies of simultaneous signal traces taken at different microwave frequencies, the film records of several runs were analyzed at points spaced one-tenth of a second apart in time. Fig. 8 gives plots of the instantaneous amplitudes at two frequencies for a typical run, and it can be seen that the correlation between the two is very good. The simultaneous Doppler frequency values also show considerable correlation, although in this case the correlation is on the gross pattern; *i.e.*, while points at the same instant of time do not necessarily correlate well, the runs exhibit a pattern made up of similar groups of points lying above or below the mean value. This is illustrated in Fig. 9. These correlations imply that the instantaneous Doppler frequency variation is a function of the large-scale pattern of the ocean waves, whereas the instantaneous amplitude variation is dependent upon the motion and orientation of small facets.

From the instantaneous doppler frequency results, the homogeneity of the Doppler signal as a function of depression angle, microwave frequency, and sea condi-
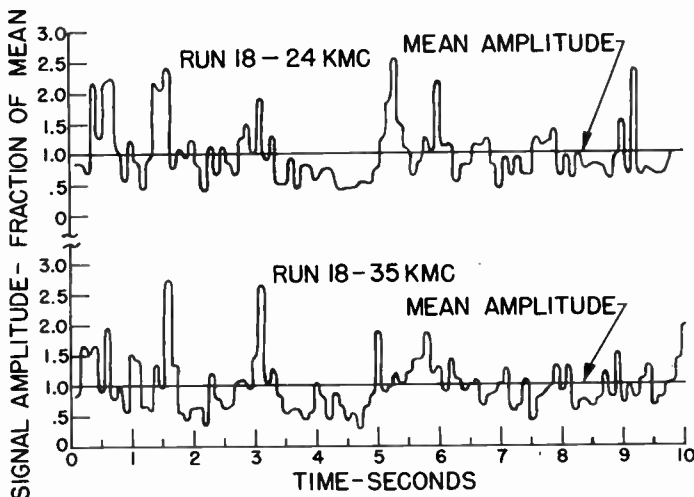
Fig. 8—Comparison of received signal amplitude vs time for simultaneous runs at the same angle, $\theta = 80°$, but for different microwave frequencies. The plots correspond to the film traces of Fig. 3.
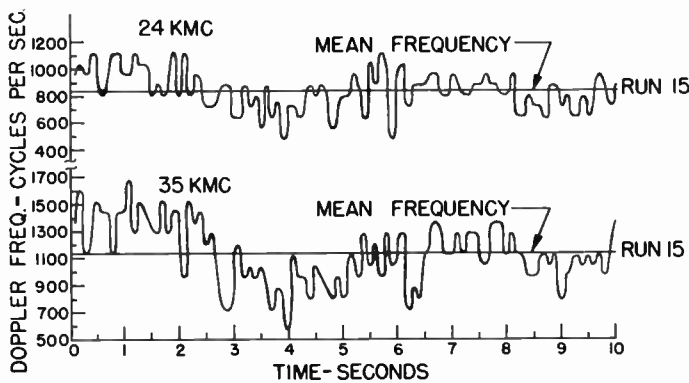


Fig. 9—Comparison of instantaneous Doppler frequencies vs time for two microwave frequencies at $\theta = 45°$.

tion was determined. Fig. 10 shows frequency distribution plots for several runs at different depression angles. The Doppler frequency is remarkably well-defined in view of the area-extensive type of target which the sea presents. For $\theta < 45°$, the instantaneous signal frequencies are contained within a $\pm 12$ per cent bandwidth about the mean for 50 per cent of the time. The percentage bandwidth necessary to contain the frequency spread increases with $\theta$, but there is no appreciable difference in the Doppler frequency homogeneity for different microwave frequencies and for different sea conditions. The scattering measurements made from different aspects of the waves while the ship was traveling in the square showed clearly that the Doppler frequency is relatively independent of sea state. The Doppler frequency changed by less than 10 per cent although the wind velocity changed from 26 knots heading to 26 following.

In order to compare the measured average Doppler frequencies with the values calculated from (1), the quantity $f\lambda/2$, where $f$ is the measured Doppler frequency and $\lambda$ is the wavelength of the microwave radiation, is plotted as a function of $\theta$ in Fig. 11. The quantity

$V \cos \theta$ for the appropriate ship speed, $V$, has also been plotted as a function of $\theta$ in this figure. The agreement between the measured and calculated values is quite good except at angles close to normal incidence.
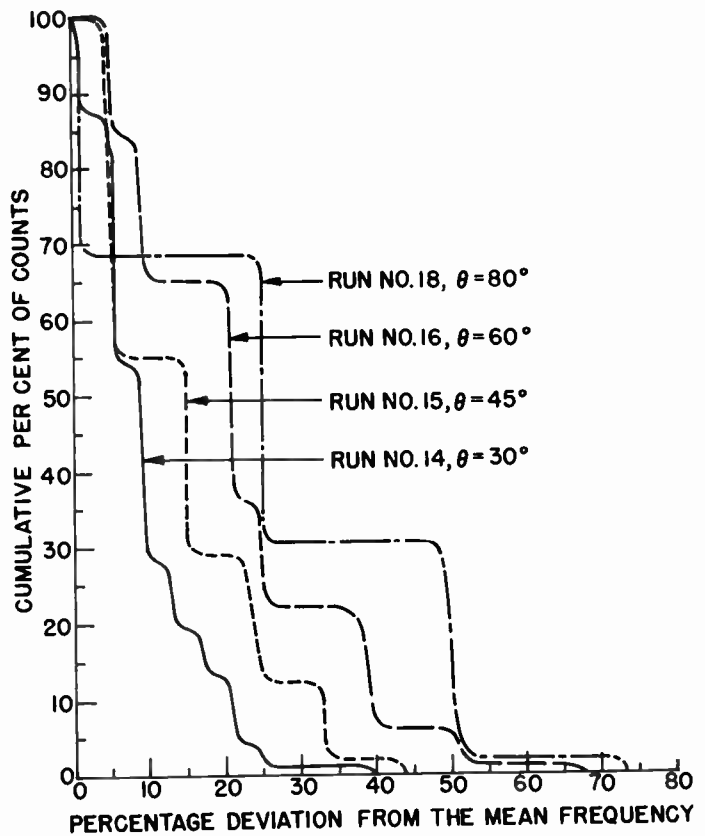


Fig. 10—Effect of depression angle on Doppler frequency distribution at 24 kmc. The cumulative per cent of Doppler frequency counts lying outside of a frequency band about the mean is plotted against the half-width of the band given in per cent of the mean frequency. The data were taken on November 22, 1955.
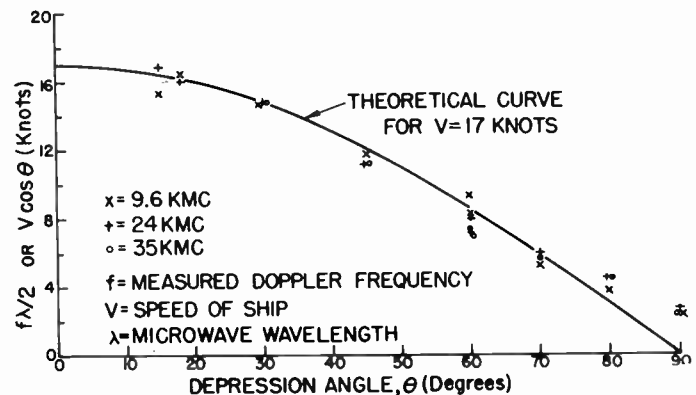


Fig. 11—Normalized Doppler frequency vs depression angle for the vertical polarization data of November 22, 1955.

### Discussion and Conclusion

From the analysis of the test data the following conclusions can be drawn about the back-scattering characteristics of the sea in the region from 9.6 to 48.7 kmc.

1) The cross section per unit area, $\sigma°$, is an increasing function of $\theta$, the slope being greatest near $\theta = 60°$. For calm seas the increase in $\sigma°$ from $\theta = 30°$ to $\theta = 90°$ may be as great as a factor of 5000.

2) Sea state has a very pronounced effect on $\sigma°$. For $\theta < 70°$ rougher sea conditions result in larger values of $\sigma°$, but at angles near vertical incidence the trend is usually reversed.

3) The cross section per unit area is approximately constant with microwave frequency and this approximation is better for large depression angles.

4) For the sea conditions encountered, $\sigma°$ for horizontal polarization is less than that for vertical polarization by several decibels at low depression angles but is nearly the same as that for vertical polarization at large values of $\theta$.

5) The values of $\sigma°$ for cross polarization are less than the vertical polarization values by about 12 db at low depression angles and by about 20 db at $\theta = 90°$.

6) Circular-polarization measurements with the system adjusted to accept single-bounce signals give values of $\sigma°$ at small depression angles which are about 5 to 10 db lower than those obtained with vertical polarization. At $\theta = 90°$, however, the values of $\sigma°$ are nearly the same as for vertical polarization.

7) Circular-polarization measurements with the system adjusted to accept double-bounce signals give values of $\sigma°$ at small depression angles which are about 5 to 10 db lower than those obtained with vertical polarization, and at $\theta = 90°$, 10 to 20 db lower.

8) The percentage fluctuations of the received signal amplitudes about the mean value are greater for large depression angles than for small.

9) There is good correlation of instantaneous amplitudes among signals received simultaneously at different microwave frequencies.

10) For vertical polarization and $\theta = 60°$, $\sigma°$ is independent of azimuth relative to the wave pattern on the ocean surface.

11) The average values of the Doppler frequency measured at depression angles $< 70°$ agree well with the values calculated from (1). The agreement is even better when higher-gain antennas are used.

12) The Doppler frequency is much less homogeneous (well-defined) at large values of $\theta$ than at small values.

13) There is little difference in the Doppler frequency homogeneity for different microwave frequencies.

14) The sea state makes only a small difference in the frequency homogeneity of the signal.

In addition to the foregoing relationships the back-scattered signal from the sea also depends on several system parameters as given in (5). From the measurements involving changes in height and changes in horn gain (5) was found to be valid for narrow beam widths and large depression angles.

Near $\theta = 90°$, the cross-polarization values of $\sigma°$ are about 20 db down from the vertical-polarization values and the circular polarization return is very small when the system accepts only double-bounce signals. Therefore, at large depression angles the back-scattered signal must be principally due to singly-scattered specular reflection. Also at large depression angles, $\sigma°$ becomes much greater than unity, indicating further that specular reflection is the primary mechanism in this region.

At small depression angles the cross-polarized and the double-bounce circular-polarized signals are not as far below the vertical-polarized values as at large angles. Hence, at low angles there is either some depolarization of the signals on reflection from the water surface or some double scattering of the signals or both. The circular polarization data are not as reliable for small depression angles as for large angles, since in the former case a greater portion of the back-scattered radiation is returned from points which are off the center line of the antenna beam; and in this region the ellipticity is generally greater than along the beam center line. Attenuation of the unwanted single-bounce or double-bounce signal, as the case may be, is then correspondingly less.

It was assumed in the beginning that the ocean target as seen by the horn antennas was made up of a large number of random scatterers. However, no assumption was made concerning the scattering pattern of these individual facets. If these facets are assumed to be flat plates or disks,[2] then their pattern will consist of a very narrow main lobe of high gain along the normal, flanked by small lobes of oscillating amplitudes. When the incident radiation strikes a small disk at an angle reasonably far removed from the normal, the amount of back-scattered energy is relatively small. If a uniform distribution of slopes up to some limiting angle is assumed for the many facets seen by the horn antennas, then the return signal is almost constant for depression angles that are less than the complement of the limiting slope angle since the oscillating side lobe intensity averages out for many facets. As the depression angle is increased, contributions from the main lobe of the facets, or specular reflection, tends to increase the back-scattered signal at a rapid rate.

From the foregoing assumptions $\sigma°$ at large values of $\theta$ should be given by

$$\sigma°(\theta) = \frac{4\pi n(\theta)A^2}{\lambda^2} = \frac{4\pi A\epsilon(\theta)}{\lambda^2} \qquad (7)$$

where there are $n$ flat surfaces of area, $A$, per unit area of sea surface, or the fraction of the surface occupied by flat reflecting surfaces is given by $\epsilon = nA$. It appears from this relation that $\sigma° \sim 1/\lambda^2$, but in order to be a

specular reflector the facets must be flat to within a certain fraction of a wavelength, say λ/16. Whatever the criterion selected, if it is maintained at all the microwave frequencies, the linear dimensions of the area increments will be proportional to the wavelength, but the total fraction of the surface occupied by the flat facets, $\epsilon$, remains approximately constant at all wavelengths. If these relations are substituted in (7), $\sigma°$ becomes independent of wavelength, which agrees with the measured results given in Fig. 7.

That the total fraction of the surface occupied at a given time and at a particular angle by flat reflecting facets is the same at all microwave frequencies is further substantiated by the good correlation shown at large depression angles among the instantaneous signal amplitudes obtained simultaneously at different microwave frequencies.

In these sea-scattering measurements the several possible sources of error and their magnitudes are as follows. At each microwave frequency the cross section measurement of the standard corner reflector contributes an uncertainty of ±0.6 db; misalignment of the calibrator in the field, oscillator drift, uncertainty in the height measurements, and film shrinkage make up an additional probable error of ±1.0 db; the calibrated attenuators contribute an uncertainty of ±0.4 db to the calibration level; and finally, errors in measuring the

signal amplitude on the microfilm reader probably do not exceed ±0.5 db. Thus a probable error of ±2.5 db seems reasonable for the cross section values obtained from the film averages. The signal amplitude values obtained from the electronic integration are free of some of the above uncertainties but are probably in error on the high side because the integrator responded somewhat slower to sudden decreases in signal level than to sudden increases in signal level.

As mentioned before, the film samples were only ten seconds in duration, whereas the integrator values represent nearly a two-minute sample from a run. Individual differences between values obtained by the two methods are thus partly due to the different samples taken. In general the film-read values are probably the more accurate of the two; but, since the sample for each run was short in duration, there is a possibility that the average value obtained by this method is not completely representative of the entire run.

### ACKNOWLEDGMENT

# Radar Terrain Return at Near-Vertical Incidence*

R. K. MOORE†, SENIOR MEMBER, IRE, AND C. S. WILLIAMS, JR.†, ASSOCIATE MEMBER, IRE

*Summary*—Measurements have demonstrated that, with occasional exceptions, radar return from the ground is largely due to area scatter, even at angles of incidence near the vertical.

An expression is derived here for a power superposition integral expressing the mean pulse envelope for the pulse returned from the ground to a radar. This integral is the convolution of the transmitted pulse form in power units with a function which includes effects of antenna pattern, ground properties, and distance. This function is generalized to include the effects of specular reflection and large isolated scatterers, as well as the more prevalent area scatter.

While beam-width-limited illumination always results in inverse-square altitude variation for area scatter, it is shown that the variation with altitude for pulse-length-limited illumination varies from inverse-square to inverse-cube, and is a function of altitude as well as of ground properties and antenna pattern. Mean returned pulses are presented for various grounds and antenna orientations.

## I. INTRODUCTION

RADAR ALTIMETERS depend upon return of signals radiated from an aircraft to the ground. In order to understand the operation of such altimeters, it is necessary to understand the processes by which radar energy is returned from the ground.

A great deal of work has been done on radar return from aircraft and ship targets, and some fair amount has been done on radar return from the ground and the sea at ranges such that the angle of incidence is near grazing. Very little work has been done in the past on angles of incidence near vertical; and, so far as the authors can determine, none of the work which has been done has been published in readily available places. An early classic on the general field of radar return from the ground was that of R. E. Clapp.[1]

[1] R. E. Clapp, "A Theoretical and Experimental Study of Radar Ground Return," M.I.T. Radiation Lab. Rep. 1024; 1946.

Scattering is the principal process by which radar energy is returned from the ground to an altimeter. This process is supplemented on some occasions by specular (mirror-type) reflection. In references to the performance of radar altimeters which the authors have encountered, it has usually been assumed that the principal process was reflection. Measurements made at Sandia Corporation have indicated that these assumptions are not correct and that, in fact, reflection is a rather uncommon mode of radar return, even for vertical incidence.[2]

The usual type of scattering process is one in which large numbers of individual scatterers located on the ground contribute more or less equally to the total scattered signal. Occasionally, some scatterer is found within the pattern of illumination whose signal return stands out appreciably above all of the rest. Measurements have shown that very rarely will more than one such target be located within the area illuminated by a radar altimeter using a broad-beam antenna. It should be noted that this is contrary to the situation encountered with search radar equipment carried in aircraft where the desired information is the distinction between the strong scattering target and the general background. The principal reason for the difference is that a radar altimeter illuminates a much larger patch of ground than a properly designed search radar since it must operate when the aircraft performs various maneuvers and, therefore, must have a wide-beam antenna. This effect is compounded since the geometry of the situation near vertical incidence is such that even very short pulses spread out to cover a large area on the ground by comparison with the area which they cover at grazing angles of incidence for a reasonable beam width.

In this paper an integral is developed which is the mean scattered radar return from a collection of small scatterers located on a plane ground. It should be noted that the results here are for a mean returned pulse for a pulse-radar altimeter; this mean is taken over a large number of returned pulses. *Any individual pulse is likely to look quite different from the mean described by the integral developed in this paper.* It is interesting to note that the integral which results is a convolution of the waveform of the envelope of the transmitted pulse in *power* units with another function which includes the effects of antenna gain, distance, and scattering properties of the ground.

Examples have been calculated here to show the effect on the return from the ground of variations of different parameters of a pulse-radar system. In particular, it is pointed out for a beam-width-limited situation (one in which the leading edge of transmitted pulse passes the outer edge of antenna beam before the trailing edge of the pulse reaches the ground) that altitude signal varies inversely as the square of altitude.

Another example considered is that of a pulse-length-limited system (one in which the leading edge of the pulse has not extended to the limit of the antenna beam by the time the trailing edge reaches the ground); here the variation of the signal with altitude is inverse cube (if the variation of the scattering properties of the terrain with angle of incidence is slow). It is shown in other examples that the variation with altitude is between inverse square and inverse cube where the scattering capability of the ground drops off more rapidly with change in angle of incidence. With any scattering variation, the signal variation is inverse cube at higher altitudes and inverse square at low altitudes.

The above examples all assume that the antenna is pointed vertically. In the case where the antenna is not pointed vertically, either because the aircraft is tilted, or because the antenna's main function has to do with radar mapping or ranging on targets at a distance, the variation with altitude and the shape of the mean return pulses have been calculated for several examples.

In order that one may evaluate the integral for mean returned signal, it is necessary to know something about the nature of the ground as a scatterer. If the signal from the ground is made up of the resultant of signals from a large number of scattering elements, one may express its properties, on the average, by a quantity known as the scattering cross section per unit area, $\sigma_0$. This is a number which expresses the average amount of power radiated back from a unit area *on the ground*, provided the incident power density is known.[3] The quantity $\sigma_0$ used here is essentially that defined by Herbert Goldstein.[4]

It must be emphasized again that the expressions calculated here for returned pulse shapes for a radar altimeter are averages over many pulses. Individual pulses will look quite different from the average pulse as the returns from the various scatterers may combine as phasors in such a way as to produce a wide range of amplitudes for individual points in individual pulses. The problem of combination of large numbers of signals of similar amplitude but random phase has been treated statistically by Rayleigh and others.[5,6] This is a problem known to statisticians as the two-dimensional random-walk problem; Rayleigh's treatment is for the random walk with an infinite number of steps. At any given range, the different phasor sums for different pulses result in fading of the signal. The signals at any two ranges corresponding to different illuminated areas fade independently. For a Rayleigh distribution the range of fading from the signal level exceeded 5 per cent of the time down to that exceeded 95 per cent of the time is

---

[2] These experiments have not yet been reported in the literature.

[3] Note that some other writers use slightly different definitions of $\sigma_0$.

[4] D. E. Kerr (ed.), "Propagation of Short Radio Waves," McGraw-Hill Book Company, Inc., New York, N. Y., p. 483; 1951.

[5] Rayleigh, "Scientific Papers of Lord Rayleigh," Cambridge, London, vol. I, p. 491, vol. IV, p. 370; 1899–1920.

[6] J. L. Lawson and G. E. Uhlenbeck, "Threshold Signals," McGraw-Hill Book Co., Inc., New York, N. Y., p. 54; 1950.

18 db. Hence, the average pulse may be quite different from any single pulse.

In the situation where specular reflection and scatter are present together, the specular signal may be much greater than any of the individual scattered signals. If it, or one of the scattered signals, stands out above the rest so that the total power in the one large signal is comparable with that due to all small scatterers, the distribution of the returned signal amplitude is appreciably altered. This problem is the same as that of a large sinusoidal signal in noise and has been treated extensively by Rice;[7] it will not be discussed here.

## II. Derivation of Expression for Scattered Return

To determine the return from a scattering ground, it will first be necessary to derive the formula for the return from a single scatterer. It will then be demonstrated that the return from two scatterers is such that the total return power is the sum of the powers in the two components, *on the average.* This will be generalized to the case of large numbers of scatterers, and it will be shown that the total returned power is the sum of the powers returned from the individual scatterers. Next we will show that, for practical purposes, the power sum may be represented as an integral over the illuminated area on the ground by use of the concept of average scattering cross section per unit area. Finally, it will be shown that this integral is actually a convolution integral involving the shape of the pulse envelope of transmitted power and a function including the effect of the ground, antenna, and distance.

### A) Return from a Single Scatterer

Let us consider a pulse radar which at periodic or quasi-periodic intervals delivers a voltage pulse to its antenna given by

$$v_D(d) = \mathrm{Re}\{V_D(d)e^{i\omega d}\}, \qquad (1)$$

where

$$V_D(d) = 0, \qquad (d < 0).$$

Lower case letters will be used to represent instantaneous voltages and powers, while upper case letters will be used to represent voltage envelopes or power averaged over an rf cycle. Hence, $V_D$ is the envelope of the transmitted voltage pulse; $d$ represents delay from the start of the transmitted pulse—that is, it is not assumed that phase is preserved from pulse to pulse.

The power averaged over an rf cycle is, then, for a real impedance of one ohm, given by

$$P_D(d) = \frac{V_D{}^2(d)}{2} \cdot \qquad (2)$$

[7] S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 23, p. 282; July, 1944; vol. 24, p. 46; January, 1945.

The average rf power of the signal returned by the $m$th scatter is

$$P_{Rm}(d) = \frac{P_D\left(d - \dfrac{2r_m}{c}\right)G_m}{4\pi r_m{}^2} \cdot \sigma_m \cdot \frac{1}{4\pi r_m{}^2} \cdot \frac{G_m\lambda^2}{4\pi}$$

$$= \frac{P_D\left(d - \dfrac{2r_m}{c}\right)G_m{}^2\lambda^2\sigma_m}{(4\pi)^3 r_m{}^4} \cdot \qquad (3)$$

New quantities introduced in this expression are:

$r_m$ = the range to the $m$th scatterer from the radar,
$c$ = the velocity of light,
$\sigma_m$ = the scattering cross section of the $m$th scatterer,
$G_m$ = the gain of antenna in the direction of the $m$th, scatterer (assuming the same antenna for transmitting and receiving),
$\lambda$ = the wavelength of the carrier radiation.

The equation has been written first to show the way in which it is built up and then in a more compact form. The first factor of the first expression shows the transmitted power radiated in the appropriate direction as a power density at the receiving point (except that argument would be $d - r_m/c$). The second factor is the scattering cross section, a quantity which determines the portion of the incident energy which is reradiated toward the radar. The third factor shows the effect of the dispersion with distance on the reradiated power density, and the fourth factor shows the receiving antenna aperture.

Utilizing the results of (3), we may write for the received voltage

$$v_{Rm}(d) = \mathrm{Re}\left\{\sqrt{2P_{Rm}(d)}e^{i\omega_m(d-2r_m/c)}e^{i\alpha_m}\right\}$$

$$= \mathrm{Re}\left\{V_{Rm}(d)e^{i\theta_m}\right\}. \qquad (4)$$

Here the phase shift has been taken into account, both that due to the travel time and the phase shift $\alpha_m$ on reflection. It should be realized that the frequency $\omega_m = \omega + \Delta\omega_m$ used here is not just the carrier frequency $\omega$, but is the carrier frequency as modified by the Doppler shift. Evidently,

$$\theta_m = (\omega + \Delta\omega_m)\left(d - \frac{2r_m}{c}\right) + \alpha_m$$

and

$$V_{Rm}(d) = \sqrt{2P_{Rm}(d)}.$$

### B) Return from Two Scatterers

The return from two scatterers is given by

$$v_R(d) = v_{R1}(d) + v_{R2}(d)$$

$$= \text{Re}\left\{ V_{R1}(d)e^{j\theta_1} + V_{R2}(d)e^{j\theta_2} \right\}. \qquad (5)$$

The product of the complex voltage above with its conjugate gives the square of the envelope of $v_R(d)$:

$$V_R^2(d) = V_{R1}^2(d) + V_{R2}^2(d)$$

$$+ 2V_{R1}(d)V_{R2}(d)\cos(\theta_1 - \theta_2). \qquad (6)$$

This expression gives the envelope of a single return from two scatterers. The average over many returns will be designated by a bar and is

$$\overline{V_R^2(d)} = \overline{V_{R1}^2(d)} + \overline{V_{R2}^2(d)} \qquad (7)$$

provided $\theta_m$ has a uniform statistical distribution, since $V_{R1}$, $V_{R2}$, $\theta_1$, and $\theta_2$ are statistically independent, and $\overline{\cos(\theta_1 - \theta_2)} = 0$. The statistical distribution of $\theta_m$ is, of course, a result of the geometry.

In taking the average above, an average is taken at each position $d$ in the return pulse. For that reason, the average $\overline{V_R^2(d)}$ remains a function of $d$. Further, in order for $d$ to represent the same position from return to return in the signals from both scatterers, it is necessary to assume that the two scatterers remain in the same position with respect to the radar during the averaging period. Actually, of course, the range varies from pulse to pulse and it is this very variation of range which causes the angle $\theta$ to vary. However, there can be a large phase variation for a small range variation and this is the required assumption.

From another point of view the result (7) may be supposed to be due to the effect of Doppler frequencies. Since the returns from the two scatterers are sinusoids at different frequencies as indicated by (4), the application of the theorem which allows superposition of power will give the result.

Pulse-to-pulse variation may be explained in terms of sampling a pattern in space. For each point in space which may be occupied by the radar, there is a specific combination of relative phase and amplitude of the return from the two scatterers, the phase being determined by the round-trip distance and phase-shift on scattering and the amplitude by the directivity of the scatterers and antenna as well as by distance. The radar samples this pattern at the points where pulses are transmitted.

### C) Return from Many Scatterers

Eq. (7) may be readily generalized to the case of many scatterers. The result is

$$\overline{V_R^2(d)} = \sum_{m=1}^{M} \overline{V_{Rm}^2(d)}$$

$$= \sum_{m=1}^{M} 2\overline{P_{Rm}(d)}. \qquad (8)$$

Then

$$\overline{P_R(d)} \triangleq \frac{1}{2}\overline{V_R^2(d)}$$

$$= \sum_{m=1}^{M} \overline{P_{Rm}(d)}$$

$$= \sum_{m=1}^{M} \overline{\left[ \frac{P_D\left(d - \dfrac{2r_m}{c}\right)G_m^2\lambda^2\sigma_m}{(4\pi)^3 r_m^4} \right]}$$

$$= \frac{\lambda^2}{(4\pi)^3} \sum_{m=1}^{M} \left[ P_D\left(d - \frac{2r_m}{c}\right)\frac{G_m^2\overline{\sigma_m}}{r_m^4} \right]. \qquad (9)$$

In the last expression above, the only quantity which must be averaged is $\sigma_m$; $\sigma_m$ for a given scatterer may vary from pulse to pulse due to the changing orientation. The other quantities are all constants with the assumptions that $r_m^4$ and $P_D[d - (2r_m/c)]$ vary negligibly from return to return for a fixed $m$. Note that this can only be so if each scatterer remains *almost* fixed in position in the return pulse during the averaging; the slight change in position must provide the phase variation.

If the phase variation is not sufficient from pulse to pulse, the samples obtained are not independent. Thus, at low frequencies, it is difficult to obtain a sufficiently large sample to determine the statistical properties of the terrain. In fact, it is almost impossible to meet both assumptions of negligible change in $P_D[d - (2r_m/c)]$ during averaging and of large phase change from pulse to pulse at low frequencies.

### D) Signal Returned from Scatterers on a Plane Surface

In many cases, the earth may be considered as plane for the purposes of calculating radar return. In such cases the development below may be used in its entirety. In some situations, as in flying over steep mountains or valleys, such an assumption is not justified and a summation is required over a more complex surface than a plane to obtain the mean power returned. This case will not be considered here but the approach used may be generalized to cover it.

Fig. 1 illustrates the geometry involved in this discussion. Consider a small region $\Delta A_m$ (to become $dA$ in the limit) containing $N$ scatterers and which is within the total illuminated area $A$. According to (9) the mean power returned from such a region is

$$\overline{[\Delta P_R(d)]}_m = \frac{\lambda^2}{(4\pi)^3} \sum_{n=1}^{N} \frac{P_D\left(d - \dfrac{2r_n}{c}\right)G_n^2\overline{\sigma_n}\Delta A_m}{r_n^4\Delta A_m} \qquad (10)$$

where the area $\Delta A_m$ has been placed in both numerator and denominator. Now if the variation of $r$ is small enough over $\Delta A_m$ so that $P_D$, $G$, and $r^4$ can be considered constant over the area, the mean power can be written
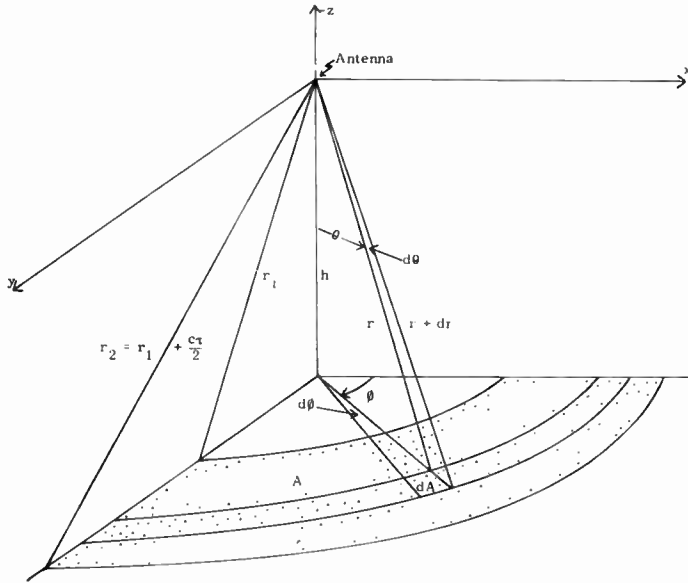
Fig. 1—Illustration of illuminated area $A$ in one quadrant.

$$\overline{[\Delta P_R(d)]_m}$$

$$\doteq \frac{\lambda^2}{(4\pi)^3} \frac{P_D\left(d - \frac{2r_m}{c}\right)}{r_m^4} G_{m2} \sum_{n=1}^{N} \frac{\overline{\sigma_n}}{\Delta A_m} \Delta A_m \quad (11)$$

where $r_m$ is the average distance to $\Delta A_m$ and $G_m$ is the antenna gain in the direction of $\Delta A_m$.

One may now *define* an average scattering cross section per unit area over this small area by

$$\sigma_{0m} \stackrel{\Delta}{=} \sum_{n=1}^{N} \frac{\overline{\sigma_n}}{\Delta A_m}. \quad (12)$$

This is a very important quantity in the theory of radar terrain return. It is not, in general, easy to determine the characteristics of individual scatterers on the ground; in fact, it is not even easy to determine what the individual ground scatterers are. Hence, the use of some sort of average scattering cross section per unit area is almost a necessity. Of course, it is possible for the cross section to be quite different in different parts of the illuminated area simply because of the fact that homogeneous regions on the ground are hard to find. One would expect that, if the entire illuminated area were a forest, $\sigma_0$ would be essentially constant within the entire illuminated region. The same might hold true for a wheatfield, or perhaps even for a city, in which case the individual scatterers would be represented by buildings. On the other hand, one can also conceive of the situation where the average scattering cross section per unit area is quite different in different parts of the illuminated region. For example, over farm land one has different types of fields, roads, fences, small streams, and farmyards, all illuminated simultaneously. Over cities one may have industrial, residential, park, and boulevard areas illuminated simultaneously.

If we use the definition of (12), we obtain from (11) the following expression for the incremental average power:

$$\overline{[\Delta P_R(d)]_m} = \frac{\lambda^2}{(4\pi)^3} \frac{P_D\left(d - \frac{2r_m}{c}\right)G_m^2 \sigma_{0m} \Delta A_m}{r_m^4}. \quad (13)$$

This is the power returned from the incremental area $\Delta A_m$. The same argument which allowed us to combine the many returned signals in (10) or (8) permits us to add the average powers returned from the $M$ different incremental areas illuminated:

$$\overline{P_R(d)} = \sum_{m=1}^{M} \overline{[\Delta P_R(d)]_m} = \frac{\lambda^2}{(4\pi)^3}$$

$$\cdot \sum_{m=1}^{M} \frac{P_D\left(d - \frac{2r_m}{c}\right)G_m^2 \sigma_{0m} \Delta A_m}{r_m^4}. \quad (14)$$

Note that the sum carried out in this case must include the transmitted pulse shape and its delay within the summation because of the fact that different delays will occur over different parts of the illuminated area; and, therefore, different incremental areas will be illuminated with different intensities. In addition, it should be noted that the gain function and the distance function are included within the summation. This is because both of these may vary appreciably over a total illuminated area. The antenna gain $G_m$ is a function of the coordinates of the particular incremental area $\Delta A_m$ in terms of a spherical coordinate system centered on the aircraft (see Fig. 1). The scattering cross section is a function of these coordinates because of the fact that a wave incident upon the scatterer at one angle will reradiate differently from a wave incident upon the scatterer at a different angle; thus, its position relative to the altimeter is material. We write then,

$$G_m = G(\theta_m, \phi_m), \quad \sigma_{0m} = \sigma_0(\theta_m, \phi_m) \quad (15)$$

where $\theta_m$ is angle from vertical to area element $\Delta A_m$ and $\phi_m$ is azimuth to $\Delta A_m$. (See Fig. 1.)

Usually one may choose a value $\Delta A_m$ which is small enough so that all the radar parameters are reasonably constant across this area. It is then possible to pass to the limit of increasingly small incremental area and write (15) as an integral:

$$\overline{P_R(d)} \doteq \frac{\lambda^2}{(4\pi)^3} \int_{A(d)} \frac{P_D\left(d - \frac{2r}{c}\right)G^2(\theta, \phi)\sigma_0(\theta, \phi)dA}{r^4}. \quad (16)$$

It should be noted that this is possible even though the concept of an average cross section per unit breaks down if the differential area is made small enough. This is because the only thing which varies with small changes of

the size of the incremental area is the product of $\sigma_0$ and the area.

The integral of (16) is over $A(d)$, the area illuminated, which depends on $d$ as indicated. At the beginning of this section reference was made to Fig. 1 as an illustration of the illuminated area. It should be noted that, in this coordinate system and for a plane, the area element may be taken as $r dr d\phi$. Thus, we may rewrite (16) for this particular case as

$$\overline{P_R(d)}$$
$$\doteq \frac{\lambda^2}{(4\pi^3)} \int_0^{2\pi} \int_h^{cd/2} \frac{P_D\left(d - \frac{2r}{c}\right) G^2(\theta, \phi)\sigma_0(\theta, \phi) dr d\phi}{r^3}. \quad (17)$$

Reference to Fig. 1 indicates that the limits in the integration on $r$ should be $r_1$ to $r_2 = r_1 + c\tau/2$, where $\tau$ is the transmitted pulse width. However, the limits shown are correct since $P_D(\mu) = 0$ for $\mu < 0$ and for $\mu > \tau$, and since $r_1 \geq h$. In other words, the exact area of illumination is taken care of by the transmitted pulse function $P_D$.

The integral of (16) is valid even though the surface is not a plane. If the surface is nonplanar, the area illuminated will not have a simple shape as in the planar case, but integration over this area by numerical means should be quite feasible.

### E) Mean Return Signal as a Superposition Integral

It will be recognized that the integral of (17) is a superposition integral. This may be shown more readily if one converts this integral into a somewhat simpler appearing form and changes the variable of integration. Let us introduce the new variable $T$ given by

$$T \triangleq \frac{2(r - h)}{c}. \quad (18)$$

$T$ is the radar delay time for the difference between range $r$ and the altitude $h$. Now, we define a function which includes all the effects of antenna, distance, and ground:

$$\left.\begin{array}{ll} B_s(T) \triangleq \dfrac{c\lambda^2}{2(4\pi)^3 r^3} \displaystyle\int_0^{2\pi} G^2(\theta, \phi)\sigma_0(\theta, \phi) d\phi, & (T \geq 0) \\[2mm] \triangleq 0, & (T < 0) \end{array}\right\}. \quad (19)$$

This integral is a function of $T$ since $r$ is a function of $T$ through (18) and $\theta$ is a function of $T$ through (18) and

$$r = h \sec \theta. \quad (20)$$

Now, (17) may be written

$$\overline{P_R(d)} = \int_0^{d - 2h/c} P_D\left(d - T - \frac{2h}{c}\right) B_s(T) dT. \quad (21)$$

If we change from a function of delay time to a function of a modified delay time, *i.e.*, replace $d - (2h/c)$ by $d$, this may be written

$$\overline{P_R\left(d + \frac{2h}{c}\right)} = \int_0^d P_D(d - T) B_s(T) dT. \quad (22)$$

The lower limit is still zero since $B_s(T) = 0$ for $T < 0$. This is recognizable as a convolution or a superposition integral. $B_s(T)$ can be seen to represent the return signal which would be received if the transmitted pulse were an impulse or delta function. Thus, the signal received at any given time is representable as the sum of the signals received from a set of different impulses having weight given by the shape of the transmitted pulse and being transmitted at times corresponding to the times of delay of the various parts of the pulse from its leading edge.

That this is valid with power is a situation which is unique to the case in which the average signal is not dependent upon cross products of the various components. Normally, such a convolution integral can be used only with voltages or currents in linear systems.

### III. Specular Reflections and Large Targets

#### A) Specular Reflections

As stated in the introduction, it is occasionally possible to find a surface which is smooth over a large enough area so that specular (mirror-type) reflections can take place. The mechanism of specular reflection is different from that of scattering and must be treated separately. It is not expected that any specularly reflecting surface will occur anywhere except directly beneath the radar because of the fact that its size must be so great that a surface which would be perpendicular to the incident radiation at any other point would not likely be of sufficient size. An exception may occur where corner reflector action between the ground and a wall of a large building is involved.

Specular reflection is normally treated on the image basis. In such a case, the received power is given by

$$P_R(d) = \frac{P_D\left(d - \frac{2h}{c}\right) G^2(0, \phi)\lambda^2}{(4\pi)^2(2h)^2} K \quad (23)$$

where $K$ is the reflection coefficient of the surface and $G(0, \phi)$ the antenna gain straight down. It is convenient to express the specularly reflected wave in the same notation used for the scattered waves. In order to do this, we define a space-time function $B(T)$ in a similar manner to that defined by (19):

$$B_m(T) = \frac{G^2(\theta, \phi)\lambda^2 K}{(4\pi)^2(2r)^2} \delta(T). \quad (24)$$

This is the $B$ function for mirror-like reflection; this expression uses the Dirac delta function to indicate that the only range at which a return is expected is that corresponding to vertical incidence where $T = 0$, $\theta = 0$ and $r = h$. One may apply this ground function just as in the case of the ground function for scattering to obtain

$$P_R(d) = \int_0^{d-2h/c} P_D\left(d - T - \frac{2h}{c}\right) B_m(T)dT \quad (25)$$

or

$$P_R\left(d + \frac{2h}{c}\right) = \int_0^d P_D(d - T) B_m(T)dT. \quad (26)$$

It should be noted, of course, that (25) is directly equivalent to (23) because of the property of the delta function that

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = f(0)$$

for any analytic function $f(x)$.

### B) Large Targets

It was pointed out in Section II that there may occasionally be individual targets that stand out appreciably above the other scatterers. Such targets, while uncommon, may be present, and should be treated separately since only one is likely to be present in any given incremental area, and it is unlikely that one will be present in more than two or three of the incremental areas involved in summing up the total scattered return. Furthermore, such a target may be physically larger than an incremental area of the type described in the second section. Also, the statistics of signal returns when all the signals have roughly comparable amplitudes are the same as noise statistics, whereas the statistics for one large target and a background of smaller ones correspond to a cw signal in noise.

It was shown in Section IIA that the radar return from one target is given by

$$P_R(d) = \frac{P_D\left(d - \frac{2r}{c}\right)G^2\lambda^2}{(4\pi)^3 r^4} \sigma$$

where $G$ is the antenna gain in the direction of the target and $\sigma$ is its scattering cross section. Now if there are $L$ large targets in the illuminated area, we define the $B$ function for them by

$$B_l(T) \stackrel{\Delta}{=} \frac{G^2(\theta, \phi)\lambda^2}{(4\pi)^3 r^4} \sum_{i=1}^{L} \sigma_i\delta(T - T_i, \phi - \phi_i) \quad (27)$$

where $\sigma_i$ is the scattering cross section of the $i$th large target and where the two-dimensional delta function is zero everywhere except at values $T = T_i$ and $\phi = \phi_i$ which correspond to the position of the $i$th large target. Then, the return due to the $L$ large targets may be written

$$\overline{P_R(d)} = \int_0^{d-2h/c} P_D\left(d - T - \frac{2h}{c}\right) B_l(T)dT \quad (28)$$

or

$$\overline{P_D\left(d + \frac{2h}{c}\right)} = \int_0^d P_D(d - T) B_l(T)dT. \quad (29)$$

It should be noted that (28) yields a sum of $L$ terms of the form of (3), one for each of the $L$ targets.

### C) Combined Return

In Sections IID, IIIA, and IIIB, we have shown how to set up a power impulse response function for the ground which takes into account antenna gain, scattering cross section, and distance. The function has been set forth for the three cases of area scatter, mirror-like reflection, and large individual targets. Because of the additive properties of power functions taken with random phases as described in Section II, the total return power is the sum of the powers due to the three cases. This leads to the definition of a combined $B$ function:

$$B(T) \stackrel{\Delta}{=} B_s(T) + B_m(T) + B_l(T)$$

$$= \frac{\lambda^2}{(4\pi)^2}\left\{\frac{c}{2(4\pi r^3)}\int_0^{2\pi} G^2(\theta, \phi)\sigma_0(\theta, \phi)d\phi\right.$$

$$+ \frac{KG^2(\theta, \phi)}{(2r)^2}\delta(T)$$

$$\left. + \frac{G^2(\theta, \phi)}{4\pi r^4}\sum_{i=1}^{L} \sigma_i\delta(T - T_i, \phi - \phi_i)\right\}. \quad (30)$$

The total mean power due to the three types of return is, then,

$$\overline{P_R\left(d + \frac{2h}{c}\right)} = \int_0^d P_D(d - T)B(T)dT. \quad (31)$$

It should be noted that this power is averaged over many separate pulses. The component associated with mirror-type reflection is a steady one and does not fluctuate from pulse to pulse. The component associated with a large target does not by itself fluctuate. However, if two large targets are present at the same position in the return, there is fading between them so that averaging is required for this result. In addition, the scattered components add phasorially to the other two components and cause fading of the resultant. In computing the statistical variations, it is necessary to consider the distribution function of the resulting voltage. Fortunately, because of the properties described in Section II, it is possible to utilize the power superposition integral of (31) to obtain the average pulse return. Note, however, that this average pulse is bound to be far different in appearance from any individual pulse.

### IV. EXAMPLES

In this section, examples of the various types of limitation of ground illumination and their effects on the return signal will be discussed. With cw or long-pulse systems, it is possible to have the illuminated region on the ground determined by the beam width of the antenna pattern, even though the pattern may be quite broad. With narrow-beam antennas, even short-pulse systems may find their illuminated areas deter-

mined by the beam width. On the other hand, with pulse systems, and fm systems which are equivalent to pulse systems, one finds frequently that the return is limited by the pulse length rather than the beam width. The variation of the illuminated area with height and range for the beam-width-limited case is different from the variation for the pulse-length-limited case.

Of course, with nonsquare pulses and nonsymmetrical antenna patterns, the situation is somewhat more complicated, and some examples are quoted here which show this effect also. This situation becomes particularly interesting when the antenna is pointed nearly horizontally instead of vertically.

## A) Beam-Width-Limited Illumination

For radars operating with long pulses or narrow beam antennas, or for certain types of cw systems, the illumination may be considered limited by the antenna beam. The simplest case to be considered here is one in which we assume the scattering cross section per unit area to

---

$$\overline{P_R(d)} = 0,$$

$$\overline{P_R(d)} = \frac{\lambda^2 G_0{}^2 \hat{\sigma}_0 P_0}{4(4\pi)^2 h^2}\left[1 - \left(\frac{2h}{cd}\right)^2\right],$$

$$\overline{P_R(d)} = \frac{\lambda^2 G_0{}^2 \hat{\sigma}_0 P_0}{4(4\pi)^2 h^2}\sin^2\theta_0,$$

$$\overline{P_R(d)} = \frac{\lambda^2 G_0{}^2 \hat{\sigma}_0 P_0}{4(4\pi)^2 h^2}\left[\left(\frac{2h}{c(d-\tau)}\right)^2 - \cos^2\theta_0\right],$$

$$\overline{P_R(d)} = 0,$$

---

be constant, the gain to be constant (antenna isotropic in region of interest, gain zero elsewhere), and the transmitted pulse to be square but of sufficient length so that it could illuminate a region bigger than the antenna will allow it to. Mathematically, these conditions are expressed as

$$\sigma_0(\theta, \phi) = \hat{\sigma}_0, \quad \text{a constant,} \tag{32}$$

$$G(\theta, \phi) = G_0, \quad (0 \leq \theta \leq \theta_0),$$

$$= 0, \quad (\theta_0 < \theta \leq \pi), \tag{33}$$

$$P_D(d) = P_0, \quad (0 < d < \tau),$$

$$= 0, \quad \text{(otherwise)}. \tag{34}$$

The pulse length is assumed to be $\tau$ and the other quantities are self-explanatory. The condition that the limitation be due to beam width is given by the inequality

$$1 + \frac{c\tau}{2h} > \sec\theta_0. \tag{35}$$

In (35) $\theta_0$ is the angle defined by (33).

In order to determine the mean return power and consequently the mean pulse shape of the return, (17)

will be used rather than those involving the convolution integral as described later in Section II. In dealing with symmetrical beams and square pulses, the original form of the integral is somewhat easier to use since it is difficult to express $B(T)$ compactly for these cases. For convenience, we repeat

$$P_R(d)$$

$$\doteq \frac{\lambda^2}{(4\pi)^3}\int_0^{2\pi}\int_h^{cd/2}\frac{P_D\left(d - \dfrac{2r}{c}\right)G^2(\theta, \phi)\sigma_0(\theta, \phi)dr d\phi}{r^3}. \tag{17}$$

The result, when stated in full, requires three separate nonzero expressions. The first corresponds to the time before the trailing edge of the transmitted pulse reaches the ground while the illuminated circle expands on the ground. The second represents the period during which the illuminated circle remains constant; its outer boundary is determined by the antenna. The third is the period during which the trailing edge of the pulse moves outward to the edge of the antenna beam. The result is

$$\left(d < \frac{2h}{c}\right) \tag{36}$$

$$\left(\frac{2h}{c} < d < \frac{2h}{c}\sec\theta_0\right) \tag{37}$$

$$\left(\frac{2h}{c}\sec\theta_0 < d < \frac{2h}{c} + \tau\right) \tag{38}$$

$$\left(\frac{2h}{c} + \tau < d < \frac{2h}{c}\sec\theta_0 + \tau\right) \tag{39}$$

$$\left(\frac{2h}{c}\sec\theta_0 + \tau < d\right). \tag{40}$$

These results are shown for a particular example in Fig. 2.
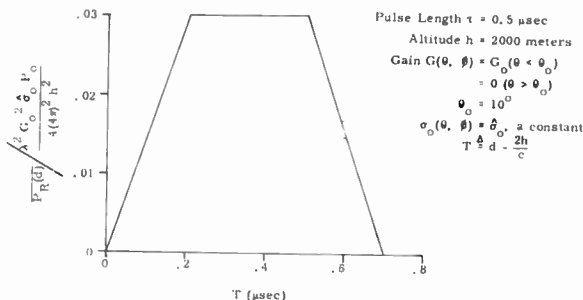


Fig. 2—Example of mean returned power for beam-limited scatter.

Note in (38) that the variation of the signal with altitude $h$ is inverse square. (This statement applies to the peak return for the angle-limited case but does not, in general, apply to the peak return for the pulse-length-limited case.) Specular reflection also yields an inverse-square altitude variation. It is possible then, to express the *scattering* properties of the ground for this type of

limitation in terms of an *equivalent reflection* coefficient. One must note, however, that in doing so the implication is that there is no fading, whereas, in point of fact, the fading is just as severe for beam-angle-limited scatter as for any other type of scatter, while a true specularly-reflected signal does not fade.

It should also be noted that the distinction between beam-angle-limited and pulse-length-limited cases is itself a function of altitude. For altitudes below those given by

$$1 + \frac{c\tau}{2h} = \sec\theta_0 \tag{41}$$

the return will be beam-angle limited. On the other hand, for altitudes above that obtained by solving this equation, the illumination will be pulse-length limited.

### B) Pulse-Length-Limited Illumination, $\sigma_0$ Constant

As stated above, the condition for pulse length limitation of illumination is that

$$1 + \frac{c\tau}{2h} < \sec\theta_0. \tag{42}$$

In this first example, we shall assume that all conditions are the same as in the preceding example with beam-width limitation except for (42); that is, the conditions of (32), (33), and (34) apply here.

In this case, two separate nonzero expressions are needed to state the result. The first corresponds to the interval before the trailing edge of the transmitted pulse reaches the ground. The second is the period after the trailing edge reaches the ground and during which the illuminated annulus on the ground spreads out. The result is, for $\theta_0 = 90°$,



Fig. 3—Example of mean returned power for pulse-limited scatter.

in space is considerably shorter than twice the altitude, *i.e.*, that

$$c\tau \ll 2h. \tag{48}$$

Then, (45) becomes

$$\widehat{\overline{P_R(d)}} = \frac{\lambda^2 G_0{}^2 \hat{\sigma}_0 P_0 c\tau}{4(4\pi)^2 h^3} \tag{49}$$

for the peak of the mean return. This equation states that the peak of the mean return power varies as the inverse cube of altitude $h$. The inverse cube relationship here should be contrasted with the inverse square relationship for beam-width limitation of illumination. Again it should be pointed out that these are merely mean pulse shapes and that there will be much fading about the mean.

Because the variation of signal strength with altitude is different for the pulse-length-limited case from that for the beam-width-limited scatter or specular reflection cases, it is not possible to describe the properties of a scattering ground, insofar as a pulse-width-limited radar is concerned, by an effective reflection coefficient. Hence, measurements made at one altitude and interpreted in

$$\overline{P_R(d)} = 0, \qquad \left(d < \frac{2h}{c}\right) \tag{43}$$

$$\overline{P_R(d)} = \frac{\lambda^2 G_0{}^2 \hat{\sigma}_0 P_0}{4(4\pi)^2 h^2}\left[1 - \left(\frac{2h}{cd}\right)^2\right], \qquad \left(\frac{2h}{c} < d < \frac{2h}{c} + \tau\right) \tag{44}$$

$$\widehat{\overline{P_R(d)}} = \frac{\lambda^2 G_0{}^2 \hat{\sigma}_0 P_0}{4(4\pi)^2 h^2}\left[1 - \frac{1}{\left(1 + \frac{c\tau}{2h}\right)^2}\right], \qquad \left(d = \frac{2h}{c} + \tau\right) \tag{45}$$

$$\overline{P_R(d)} = \frac{\lambda^2 G_0{}^2 \hat{\sigma}_0 P_0}{(4\pi)^2 c^2 d^2}\left[\frac{1}{(1 - \tau/d)^2} - 1\right], \qquad \left(\frac{2h}{c} + \tau < d\right) \tag{46}$$

$$\overline{P_R(d)} \doteq \frac{\lambda^2 G_0{}^2 \hat{\sigma}_0 P_0}{(4\pi)^2 c^2}\frac{2\tau}{d^3}, \qquad \left(\frac{2h}{c} + \tau < d, \tau \ll d\right) \cdot \tag{47}$$

Eq. (45) is included since this gives the peak of the mean return. Eq. (47) follows directly from (46). Eq. (47) states that at any fixed altitude $h$ the mean returned pulse decays as $1/d^3$, *i.e.*, as the inverse cube of range. An example which behaves according to (43) through (47) is shown in Fig. 3.

We now make the assumption that the pulse length

terms of signals at another altitude must specify the mechanism involved in the radar return so that the proper type of extrapolation may be used in going from one altitude to another. Assumption of either specular reflection or constant-$\sigma_0$, pulse-limited scatter may lead to difficulties in extrapolating measurements to different altitudes.

## C) Pulse-Length-Limited Illumination, $\sigma_0$ Variable

The example of Section IVB, assumed that $\sigma_0$ was independent of depression angle. This assumption is not, in general, justified, although types of ground have been seen for which it seems fairly reasonable, particularly at the higher frequencies. Most types of ground which have been observed in the Sandia Corporation experimental program have a $\sigma_0$ which decreases as the angle with the vertical is increased. Some examples have been calculated for different types of variation of $\sigma_0$ with angle and it has been ascertained that neither an inverse square or an inverse cube altitude variation applies to all such situations.

Variations of the exponent associated with the change of signal with altitude have been determined for the cases where

$$\sigma_0 = \hat{\sigma}_0 e^{-\theta/15°}$$

and

$$\sigma_0 = \hat{\sigma}_0 e^{-\theta/5°}$$

for a rectangular transmitted pulse of duration $\tau$ and for an antenna gain which is constant over the region of interest. The results are shown in Fig. 4, where
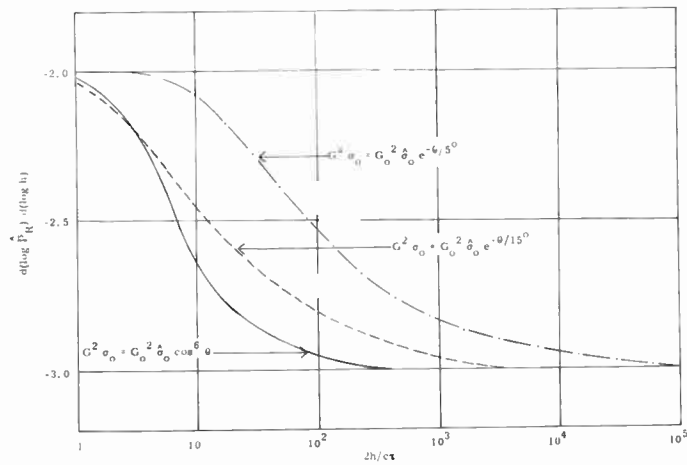


Fig. 4—Variation of peak mean return $\widehat{\overline{P}}_R$ with altitude $h$; rectangular transmitted pulse of duration $\tau$.

$d(\log \widehat{\overline{P}}_R)/d(\log h)$ is plotted vs $2h/c\tau$. $\widehat{\overline{P}}_R$ is the maximum of the mean return. This expression does give the exponent, since, if

$$\widehat{\overline{P}}_R = kh^n,$$

then

$$\log \widehat{\overline{P}}_R = \log k + n \log h,$$

and

$$\frac{d(\log \widehat{\overline{P}}_R)}{d(\log h)} = n.$$

It is readily apparent that the inverse cube law prevails at high altitudes and the inverse square at low altitudes. This might be expected since a very small change in $\sigma_0$ across the region illuminated by the pulse approaches the case of paragraph B) of this section, where $\sigma_0$ is independent of angle. This small change occurs at high altitudes since for this case the angle that defines the illuminated region corresponding to peak return is quite small. On the other hand, for low altitudes, the area illuminated by a pulse corresponds to a very large angle and the limitation of return is due to decrease of $\sigma_0$ with angle rather than pulse length. Hence, an inverse square variation occurs.

Another example of the variation of maximum mean return with altitude has been calculated for

$$\sigma_0 = \hat{\sigma}_0 \cos^{6-n} \theta$$

and

$$G^2 = G_0^2 \cos^n \theta$$

(*i.e.*, for $\sigma_0 G^2 = \hat{\sigma}_0 G_0^2 \cos^6 \theta$) and for a rectangular transmitted pulse of duration $\tau$. The results are also shown in Fig. 4.

In this case, as in the two previous cases, the variation becomes proportional to $1/h^3$ as $2h/c\tau \to \infty$, and becomes proportional to $1/h^2$ as $2h/c\tau \to 0$. The latter case corresponds to $2h \ll c\tau$ and is of little interest for a pulse altimeter since the minimum altitude for which it is useful is $h = c\tau/2$. However, the case is of general interest.

## D) Mean Return for Nonsquare Transmitted Pulse; Effect of Antenna Orientation

The pulse shown in Fig. 5 is taken as the transmitted
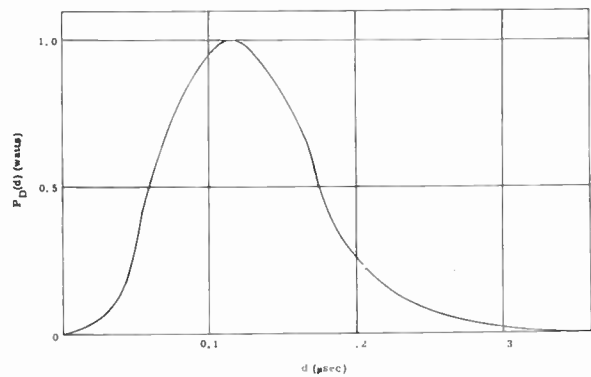


Fig. 5. Envelope of transmitted pulse in power units.

pulse. The antenna gain $G$ and $\sigma_0$ are assumed to be

$$G = 6 \sin^2 \gamma \cos^2 \delta \quad \text{(one lobe only)}$$

where $\gamma$ and $\delta$ are longitude and colatitude in a coordinate system based on the antenna, and

$$\sigma_0 = \hat{\sigma}_0 e^{-\theta/15°}.$$

The mean returned pulse (normalized) for four different

Fig. 6—Example of mean returned power.

$h = 0.122$ km
$\sigma_o = \hat{\sigma}_o \, e^{-\theta/15^\circ}$
$G = 6 \sin^2\gamma \cos^2 \phi$ (One lobe only)
$P_D$ in Figure 5
$\hat{P}_R = 1.66 \times 10^{10} \, \hat{\sigma}_o/f^2$ watts ($\beta = 90^\circ$)
$= 1.96 \times 10^9 \, \hat{\sigma}_o/f^2$ watts ($\beta = 30^\circ$)
Antenna orientation given by $\beta$, angle from horizontal

Fig. 8—Example of mean returned power.

$h = 3.05$ km
$\sigma_o = \hat{\sigma}_o \, e^{-\theta/15^\circ}$
$G = 6 \sin^2\gamma \cos^2 \phi$ (One lobe only)
$P_D$ in Figure 5
$\hat{P}_R = 4.3 \times 10^6 \, \hat{\sigma}_o/f^2$ watts ($\beta = 90^\circ$)
$= 2.9 \times 10^5 \, \hat{\sigma}_o/f^2$ watts ($\beta = 30^\circ$)
Antenna orientation given by $\beta$, angle from vertical

Fig. 7—Example of mean returned power.

$h = 0.305$ km
$\sigma_o = \hat{\sigma}_o \, e^{-\theta/15^\circ}$
$G = 6 \sin^2\gamma \cos^2 \phi$ (One lobe only)
$P_D$ in Figure 5
$\hat{P}_R = 1.89 \times 10^9 \, \hat{\sigma}_o/f^2$ watts ($\beta = 90^\circ$)
$= 1.82 \times 10^8 \, \hat{\sigma}_o/f^2$ watts ($\beta = 30^\circ$)
Antenna orientation given by $\beta$ angle from horizontal

Fig. 9—Example of mean returned power.

$h = 1.22$ km
$\sigma_o = \hat{\sigma}_o \, e^{-\theta/15^\circ}$
$G = 6 \sin^2\gamma \cos^2 \phi$ (One lobe only)
$P_D$ in Figure 5
$\hat{P}_R = 5.3 \times 10^7 \, \hat{\sigma}_o/f^2$ watts ($\beta = 90^\circ$)
$= 4.0 \times 10^6 \, \hat{\sigma}_o/f^2$ watts ($\beta = 30^\circ$)
Antenna orientation given by $\beta$, angle from horizontal

altitudes, 0.122, 0.305, 1.22, and 3.05 km, is then given by Figs. 6, 7, 8, and 9. In each figure, two values of angle $\beta$ are used. This angle represents the antenna orientation; $\beta = 90^\circ$ represents the case where the antenna is directed vertically downward and $\beta = 0^\circ$ corresponds to a horizontally directed antenna.

## V. CONCLUSION

Radar return from the ground at near-vertical incidence is usually due to area scattering; but it may be due, at least in part, to specular reflection and scattering from individual large targets. It has been shown that the mean return to a pulse radar can be expressed by a power superposition integral involving the transmitted pulse envelope and a function including effects of ground properties, antenna gain, and distance.

With beam-width limitation of illumination, the mean peak signal varies inversely with the square of altitude. With pulse-length limitation of illumination, the mean peak signal varies inversely as the square of altitude for specular reflection, between inverse square and inverse cube for area scatter, and inversely as the fourth power for large target scatter. Curves have been presented showing altitude variation for various area scattering coefficients, and examples showing the effect of antenna tilt have been included.

## VI. ACKNOWLEDGMENT

# Operation of a Cold Cathode Gas Triode in a High Impedance Self-Biasing Circuit*

MARVIN SILVER†

*Summary*—It has been found that when a cold cathode gas triode operates in a high impedance grid circuit in the absence of light or any other external radiation, these tubes have a tendency to fire without any triggering signal applied. This unsatisfactory operation is shown to be a function of the rate of rise of voltage and the time delays found in the breakdown of gaseous discharges. It is further shown that operation can be made satisfactory by making the pre-breakdown current larger than a certain minimum value.

## INTRODUCTION

IN MANY applications using a cold cathode gas triode as a trigger tube, it is desirable to keep the triggering power at a minimum. The operation of a cold cathode gas triode is such that for a given plate voltage, the tube will not fire until a certain minimum grid current flows. Therefore, in order to minimize the firing power, the grid should be biased as close to its breakdown voltage as possible. L. E. Richtmyer[1] of the Diamond Ordnance Fuze Laboratory has devised a circuit in which the grid is biased at its breakdown voltage. This is accomplished by applying a voltage greater than the breakdown value through a very large resistor to the grid. The grid current is limited to a very small value by the resistor thus keeping the discharge in the Townsend region. This arrangement should provide self-biasing because the Townsend discharge would be self-maintained at the grid breakdown voltage at such low values of current. Fig. 1 is an example of such a self-biasing circuit. Firing of the tube takes place by applying a pulse across resistor $R_s$. This pulse need only be a few volts in amplitude because the voltage appearing between the grid and cathode is the sum of the biasing voltage and the pulse voltage. Satisfactory operation has been observed when the tube is operated in light. In many instances when the tube was operated in the dark, the tube fired even though no external triggering signal had been applied to the grid. Observations with a fiber electrometer connected between grid and cathode revealed that the grid voltage was fluctuating about its breakdown value. The tube was seen to fire when the amplitude of any of these fluctuations exceeded a certain value. Simultaneous with these voltage fluctuations, there are pulses of current. These current pulses were generally too small to be observed directly; however, they did produce enough light to be observed with a photomultiplier. Fig. 2 is a graph of the photomultiplier

[1] L. E. Richtmyer, unpublished work.
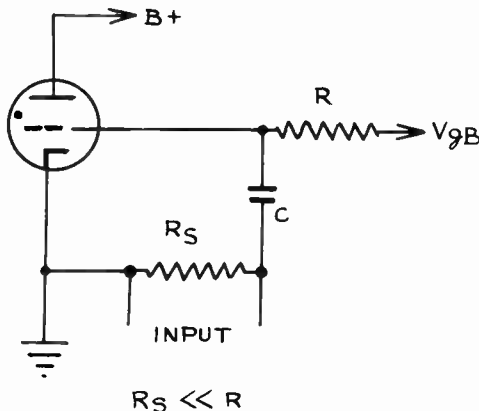


Fig. 1—Typical high impedance self-biasing circuit.
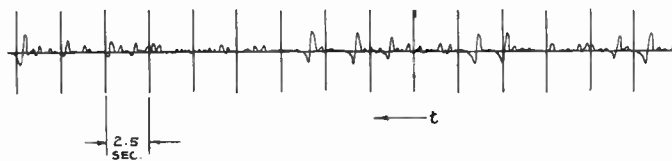


Fig. 2—Graph of photomultiplier current taken with a recording milliammeter.

current vs time taken with a recording milliammeter. As can be seen from the graph, these current pulses vary in amplitude, and accompanying the large pulses was complete breakdown of the tube. There were no current pulses that could be observed when the tube was operated in room light or when there were no voltage fluctuations. This unsatisfactory operation in the dark is a serious disadvantage since most applications require that the tube operate in the dark.

Experiments with Chatham 395A's as well as other 395A's in a circuit like that shown in Fig. 1 with

$$V_{gb} = 150 \text{ v} \qquad B^+ = 150 \text{ v}$$
$$R = 10^{11} \text{ ohms} \qquad C = 10 \text{ mmf}$$
$$R_s = 10^5 \text{ ohms}$$

showed that the self-priming pulses could be made negligibly small only by making the prebreakdown current larger than about $5 \times 10^{14}$ amperes. Therefore it was believed that these pulses were inherent in the breakdown process of a Townsend like discharge device when operating in a high impedance circuit. It is the purpose of this paper to present a theory of operation of a cold cathode gas triode operating in a high impedance circuit based upon the fundamental concepts of Townsend gas discharge theory in order to explain the difference be-

tween light and dark operation and to determine what can be done in order to make tubes function satisfactorily in the dark.

## Theory

Under normal operating conditions, a cold cathode gas triode will become conducting only after the grid current has exceeded the minimum value necessary to transfer the discharge to the plate. Therefore, the analysis of the operation of such a tube need concern itself only with what takes place in the grid circuit. In a high impedance circuit like that shown in Fig. 1, the relatively large value of transfer current, about $10^{-6}$ amperes, cannot come from the grid bias battery because the series resistor would limit the current to a value well below that required for transfer. The only path available for a current as large as the transfer current is through the capacitor $C$. This large value of current may be produced if for some reason the voltage on the capacitor exceeds the breakdown voltage of the grid, or if a pulse is applied through the capacitor to the grid which again increases the grid voltage above breakdown. The second possibility is what takes place in normal stable operation of the tube. The first possibility causes self-firing. The charging current for the capacitor comes from the bias battery through the resistor $R$. If the grid current is less than the charging current, the capacitor voltage will increase. If the grid current is greater than the charging current, the capacitor voltage will decrease. In stable operation, the grid current is just equal to the current available for charging, and the voltage remains fixed at the grid breakdown voltage. The grid current will generally be less than the charging current so long as the grid voltage is less than breakdown. Therefore, when the bias voltage is first applied, the capacitor will charge up until the breakdown voltage is reached. Because there is a delay in the build-up of the discharge, the voltage will continue to increase above breakdown until the grid current is just equal to the charging current. The grid is "over-volted" and the tube will fire or not fire depending upon whether the charge available in the capacitor $C$ due to this over-voltage is great enough to cause a sufficiently large current to flow which can effect transfer. Therefore, the tube will be stable or unstable depending upon the delays in the build-up of the grid current and the rate of rise of voltage during these delays.

If the grid voltage is less than the breakdown value, the grid current consists of the collection of ion pairs created in the gas by radiation from a radium source in the tube and the photo-electric current emitted by the cathode. This prebreakdown current is generally very small and if the series grid resistor is less than $10^{12}$ ohms, the prebreakdown current, $i_0$, is many orders of magnitude less than the resistor current. Therefore, when the battery voltage is first applied, the grid voltage will rise as in an ordinary RC series circuit. In the neighborhood of the grid breakdown voltage, the linear

approximation may be used for the expression for the instantaneous value of over-voltage. Therefore,

$$\Delta V = V - V_s = \frac{V_{gb} - V_s}{RC}(t - t_0) = \frac{i_b t'}{C}. \quad (1)$$

Where $V$ is the instantaneous value of voltage, $V_s$ is the grid breakdown voltage, $t_0$ is the time at which $V = V_s$, $t' = t - t_0$, $V_{gb}$ is the bias battery voltage and $i_b$ is the current from the battery available for charging the capacitor $C$. The voltage will continue to rise in this fashion until the grid current increases and becomes an appreciable part of $i_b$. It is well known from gas discharge theory, that the delay in the build up of the current can be divided into two parts, the formative time lag and the statistical time lag. During the formative time, the instantaneous value of current is given by[2]

$$i = \frac{i_0[M^{((t-t_f)/t_i)} - 1]}{M - 1}. \quad (2)$$

Where $M = \gamma[e^{\eta(V-V_0)} - 1]$, $\eta$ is the number of ionizing collisions per volt per electron, $\gamma$ is the Townsend second coefficient, $V_0$ is approximately equal to the ionization potential of the gas, $t_f$ is the start of the formative time, and $t_i$ is the time between successive Townsend Avalanches during the build-up process. (In the case where $i_0$ is due to ionization of neutral gas molecules by radiation from radioactive material, $i$ must be divided by $\eta(V - V_0)$ to obtain the correct value. Fortunately, however, for oxide coated cathode discharges, as found in the 395A, $\eta(V - V_0) \sim 1$. Therefore, for the purposes of the development of the theory, (2) may be used and still give results which are good to within an order of magnitude.) The following approximations will be made:

$$\gamma e^{\eta(V_s - V_0)} = 1 \quad \text{and} \quad \gamma e^{\eta(V - V_0)} = M.$$

In a 395A where $\gamma \approx 0.25$ these approximations introduce an error of about 20 per cent but simplify the algebra a great deal. There is no essential change in the analysis or the results if these approximations are not made. Eq. (2) does not represent the most accurate expression for the current, but for the over-voltage of interest here it reduces to the same thing.

During the statistical time lag, which precedes the formative period, the probability that a discharge will start between $t'$ and $t' + dt'$ is given by[3]

$$f(t')dt' = \frac{i_0}{e} W \exp\left[-\int \frac{i_0}{e} W dt'\right] dt' \quad (3)$$

where the value of $W$ is obtained from[3]

$$M = \frac{\gamma \ln(1 - W)}{\ln(1 - \gamma W)}. \quad (4)$$

[2] M. J. Druyvesteyn and F. M. Penning, "The mechanism of electrical discharges in gases of low pressure," *Rev. Mod. Phys*, vol. 12, p. 117; April, 1940.
[3] *Ibid.*, p. 118.

From (1), $M$ can be expanded in terms of $t'$; remembering that $M = 1$ at $V = V_s$

$$M = \gamma e^{\eta(V_s - V_0 + i_b t'/C)} \approx 1 + \frac{\eta i_b t'}{C} = 1 + \eta \Delta V. \quad (5)$$

For values of $M$ not much greater than unity, $W$ will be very small, and since $\gamma$ is also small compared with one, a simpler expression for $W$ may be obtained by expanding the log terms in (4) and taking the first two terms of the numerator and ignoring terms smaller than $\gamma W$ in the denominator

$$W \approx 2(M - 1) = \frac{2\eta i_b t'}{C} = 2\eta \Delta V. \quad (6)$$

Putting this result in (3) gives

$$f(t')dt' = \frac{2\eta i_b i_0 t'}{eC} \exp\left[-\frac{\eta i_b i_0 t'^2}{eC}\right] dt'. \quad (7)$$

Or in terms of the grid over-voltage

$$f(\Delta V)d(\Delta V) = \frac{2\eta i_0 C \Delta V}{e i_b} \exp\left[-\frac{\eta i_0 C(\Delta V)^2}{e i_b}\right] d(\Delta V). \quad (8)$$

With this expression for the probability that the voltage will rise to at least $\Delta V$, the average value of over-voltage is obtained by integrating $\Delta V f(\Delta V)d(\Delta V)$, and it is

$$\overline{\Delta V} = \left[\frac{\pi e i_b}{4C\eta i_0}\right]^{1/2}. \quad (9)$$

During the formative time, the voltage will continue to rise until the grid current is equal to the current through the resistor at which point there will be no further charging of $C$. The additional time of rise, $\Delta t$, can be found from (2) and for $t_b - t' \gg t_i$:[4]

$$t_b - t' = \Delta t = \frac{t_i}{\ln \overline{M}} \ln\left[\frac{i_b}{i_0}(\overline{M} - 1)\right] \quad (10)$$

where $t_b$ is the time at which $i = i_b$ and $\overline{M}$ is the average value of $M$ during the formative period. Since the current build-up is exponential, it is only during the two or three avalanches just prior to $i = i_b$ that $i$ is of the order of magnitude of $i_b$. Therefore, the average value of $M$ is approximately the value of $M$ at $\Delta V + (\Delta V_f/2)$ ($\Delta V_f$ is the additional over-voltage during the formative time). The maximum over-voltage attained by the grid is $\Delta V_m$

$$\Delta V_m = \Delta V + \Delta V_f \quad (11)$$

where

$$\Delta V_f = \frac{i_b}{C} \Delta t. \quad (12)$$

After $t_b$, the voltage will decrease and will continue to

do so as long as $i > i_b$. The current on the other hand will continue to increase as long as $V > V_s$. Since $i_0 \ll i_b$, the expression for the instantaneous value of current during the period of voltage decrease is

$$i = i_b \overline{M}^{((t-t_b)/t_i)} \quad (13)$$

where $\overline{M}$ is approximately the value of $M$ at $V = V_m$. The current $i_s$ at $V = V_s$ can be found from

$$\Delta V_m = \frac{1}{C} \int_{t_b}^{t_s} i_b \overline{M}^{((t-t_b)/t_i)} - \frac{i_b(t_s - t_b)}{C} \quad (14)$$

or

$$i_s = \frac{C\eta(\Delta V_m)^2}{t_i}. \quad (15)$$

This calculation will be shown in Appendix I.

Now the condition for stability can be established. In order for a tube not to fire, it is necessary that $\Delta V_m$ be sufficiently small such that $i_s$ is less than the minimum current necessary to fire the tube. If $i_t$ is the critical value of grid current ($i_t$ should not be confused with the static transfer current but is a function of the grid cathode capacity as well as the properties of the gas)[5] and $\Delta V_t$ is the value of over-voltage necessary for $i_s = i_t$, the stability condition is

$$\Delta V_m < \Delta V_t. \quad (16)$$

If $\Delta V_m$ is less than $\Delta V_t$, the current will decay after the voltage drops below $V_s$. The voltage will continue to decay until $i = i_b$ again. At this time, the voltage will increase and the process will start again.

## DISCUSSION

From this analysis, it is seen that the voltage fluctuates between $V_s + \Delta V_m$ and some value of voltage below $V_s$. Since $\Delta V_m$ depends upon $\Delta V$ (the statistical over-voltage) the theory predicts these fluctuations to be statistical. This is of course what is observed experimentally.

That $i_0$ should be the only reasonable variable for reducing these fluctuations is seen from (8). The parameter $\eta$ cannot be altered without affecting many of the other characteristics of the tube which are required. In addition, for Neon-Argon mixtures as used in the 395A, $\eta$ is a very slowly varying function of pressure and electrode spacing, and therefore a very large change in tube construction would be required to materially affect the value of $\eta$. A change in $i_b$ or $C$ would not be too effective from a practical point of view for several reasons: since $i_b/C$ is essentially $dV/dt$, one would not want to make that ratio too small or it would take too long for the condenser to charge. Also, just an increase in $C$ would not be effective because this would tend to decrease the value $\Delta V_t$. It is true that decreasing $V_{sb}$ to a

[4] F. G. Heymann, "Breakdown in cold cathode tubes at low pressure," *Proc. Phys. Soc. London*, vol. 63, pp. 25–41; January, 1950.

[5] E. Meili, "Steverung von glimmtrioden mit kleinsten stromen," *Helv. Phys. Acta.*, vol. 26, pp. 574–577; Fasciculus VI, 1953.

value very close to $V_s$ would still permit fairly rapid charging of $C$ and have the advantage of making $i_b/C$ very small in the neighborhood of $V_s$. However, there is a limitation on how close $V_{gb}$ can be made to $V_s$ and some provision would have to be made for the variation of $V_s$ with time in a tube as well as the variation of $V_s$ from tube to tube which can be as much as 6 volts. This leaves only $i_0$ to be altered which will materially reduce the probability that the tube will fire.

If at some time $\Delta V_m$ is very small, then the voltage will not drop very much below $V_s$ when $i = i_b$. Therefore, it is possible that when the voltage rises again to $V_s$, the current is greater than $i_0$. This would have the effect of shortening the statistical time lag, and $\Delta V_m$ will again be small. A tube with a small $i_0$ may after many cycles stabilize itself if it happens to have a very short statistical lag. This has been observed many times in 395A's. To depend upon this process for stabilization is not very satisfactory because the tube will probably fire many times before it becomes stable. In addition, should the discharge be interrupted for some reason, such as a temporary removal of voltage, operation of the tube will again be governed by its $i_0$.

Since $\Delta V_m$ is not the same for each cycle, there is always a small probability that the tube will fire. One can, by making $i_0$ sufficiently large, make the probability that the tube will fire extremely small and self-firing would be unlikely. In a 395A, for example, with a $10^{11}$ ohm grid resistor and a $10^{-11}$ farad grid capacitor, an $i_0$ of about $2.4 \times 10^{-14}$ amperes makes the probability $10^{-6}$ that the tube will fire.

The fact that this value agrees reasonably well with the experimental value of $i_0$ for the elimination of the self-priming pulses is further verification of the validity of the theory.

When a 395A is operated in normal room light, the dark current is about $10^{-12}$ amperes and therefore, operation under these conditions would make the tubes stable over very long periods of time since its probability for firing would be so small. To make the tubes operate satisfactorily in the dark, it is necessary to increase the amount of radium in the tube. About 0.1 microgram of radium suitably placed within a 395A is necessary for $i_0$ to be about $10^{-13}$ amperes. A 395A with this amount of radium has been found to be stable for several months.

## Appendix I

The derivation of (15)

$$\Delta V_m = \frac{1}{C} \int_t^{t_s} i \, dt - \frac{i_b}{C}(t_s - t_b)$$

$$= \frac{1}{C} \int_{t_b}^{t_s} i_b \overline{M}^{((t-t_b)/t_i)} dt - \frac{i_b(t_s - t_b)}{C}$$

$$= \frac{1}{C} \left\{ \frac{i_b t_i}{\ln \overline{M}} \left[ \overline{M}^{((t_s - t_b)/t_i)} - 1 \right] - i_b(t_s - t_b) \right\}$$

now since $\overline{M} = 1 + \eta \Delta V_m$ and $\eta \Delta V_m < 1$

$$\ln \overline{M} = \eta \Delta V_m$$

$$i_s = i_b \overline{M}^{((t_s - t_b)/t_i)}$$

$$t_s - t_b = \frac{t_i \ln \left( \frac{i_s}{i_b} \right)}{\ln \overline{M}}$$

and

$$\Delta V_m = \frac{t_i}{C \eta \Delta V_m} \left[ i_s - i_b - i_b \ln \left( \frac{i_s}{i_b} \right) \right]$$

and in general $i_s \gg i_b$

$$i_s = \frac{C \eta (\Delta V_m)^2}{t_i} .$$

## Appendix II

*Calculations of $i_0$ for a $10^{-6}$ Probability that the 395 A Will Fire*

| Necessary Data | | |
|---|---|---|
| $R = 10^{11}$ ohms | $V_{gb} = 150$ v | $\eta = 0.028$ l/v |
| $C = 10$ mmf | $V_s = 75$ v | $i_b = 7.5 \times 10^{-10}$ amperes |
| $e = 1.6 \times 10^{-19}$ coulombs | | |

Experiments performed on the 395A (see above) with those circuit constants showed that the tube will fire if $\Delta V_m$ exceeded 0.5 volts. In addition, formative time lags as measured at 0.5 volts over-voltage were approximately 200 microseconds. Therefore, $\Delta V_f$ would be approximately 0.015 volts. Thus $\Delta V_f$ can be neglected compared with $\Delta V_m$ and so $\Delta V = 0.5$ volts. Integrating (8) gives the probability that a discharge will occur after $\Delta V$. This quantity then is what must be made equal to $10^{-6}$. Therefore
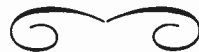
$$10^{-6} = \exp \left[ - \frac{\eta i_0 C (\Delta V)^2}{e i_b} \right]$$

or

$$6 \times 2.3 \approx \frac{\eta i_0 C (\Delta V)^2}{e i_b}$$

and inserting 0.5 for $\Delta V$ and the values of the other quantities given above gives

$$i_0 \approx 2.4 \times 10^{-14}.$$

# Correspondence

## A Simplified Procedure for Finding Fourier Coefficients*

The usual method for finding the Fourier Coefficients of a function is to utilize the orthogonal property of cos $mx$ anc cos $nx$ over the interval $-\pi \leq X \leq \pi$, i.e.,

$$\int_{-\pi}^{\pi} [\cos mx][\cos nx]dx = \begin{cases} \pi \text{ if } m = n \\ 0 \text{ if } m \neq n. \end{cases} \quad (1)$$

Thus, if $f(x)$ (assumed even) is written as

$$f(x) = \sum_{n=0}^{\infty} a_n \cos nx \quad (2)$$

one can multiply both sides by cos $mx$ and integrate, obtaining

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx. \quad (3)$$

When $f(x)$ cannot be expressed by a single relatively simple formula over the whole interval, this method of finding the $a_n$ can become tedious, even for simple-looking functions. This note describes a procedure for finding the $a_n$ which eliminates the complicated integrations if $f(x)$ is of the proper form, and with practice, allows one to write down the coefficients by inspection for many cases. The method depends on the sampling property of delta functions, namely that

$$\int_a^b f(x)\delta(x - \theta)dx = \begin{cases} f(\theta) \text{ if } a \leq \theta \leq b \\ 0 \text{ otherwise.} \end{cases} \quad (4)$$

An example, worked out both ways, will illustrate the simplicity of the method, and its superiority over the conventional method in those cases where the method can be applied. For this purpose, consider Fig. 1.

Proceeding in the usual way, one has

$$a_n = \frac{1}{\pi} \Bigg[ \int_{-\pi}^{-\theta_2} 0 \cdot \cos nx dx$$
$$+ \int_{-\theta_2}^{-\theta_1} \frac{x + \theta_2}{\theta_2 - \theta_1} \cos nx dx + \int_{-\theta_1}^{\theta_1} \cos nx dx$$
$$+ \int_{\theta_1}^{\theta_2} \left[ 1 - \frac{x - \theta_1}{\theta_2 - \theta_1} \right] [\cos nx] dx$$
$$+ \int_{\theta_2}^{\pi} 0 \cdot \cos nx dx \Bigg]$$
$$= \frac{1}{\pi} \Bigg[ \frac{\theta_2}{\theta_2 - \theta_1} \frac{\sin nx}{n} \Big|_{-\theta_2}^{-\theta_1}$$
$$+ \frac{x}{\theta_2 - \theta_1} \frac{\sin nx}{n} \Big|_{-\theta_2}^{-\theta_1} + \frac{\cos nx}{n^2(\theta_2 - \theta_1)} \Big|_{-\theta_1}^{\theta_1}$$
$$+ \frac{\sin nx}{n} \Big|_{-\theta_1}^{\theta_1} + \frac{\sin nx}{n} \Big|_{\theta_1}^{\theta_2} - \frac{x \sin nx}{n(\theta_2 - \theta_1)} \Big|_{\theta_1}^{\theta_2}$$
$$- \frac{\cos nx}{n^2(\theta_2 - \theta_1)} \Big|_{\theta_1}^{\theta_2} + \frac{\theta_1 \sin nx}{n(\theta_2 - \theta_1)} \Big|_{\theta_1}^{\theta_2} \Bigg].$$

Collecting terms, one finds

$$a_n = \frac{2}{n^2\pi(\theta_2 - \theta_1)} (\cos n\theta_1 - \cos n\theta_2).$$

Suppose, however, instead of proceeding directly to find the $a_n$, that one first finds the derivative of $f(x)$. Formally, one has
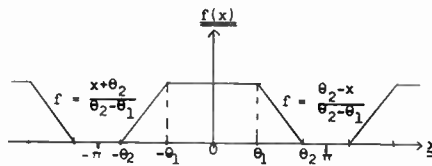
Fig. 1
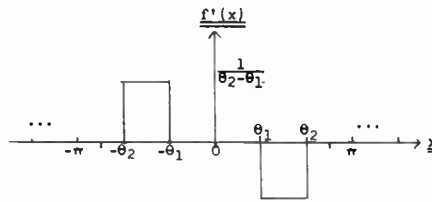


Fig. 2
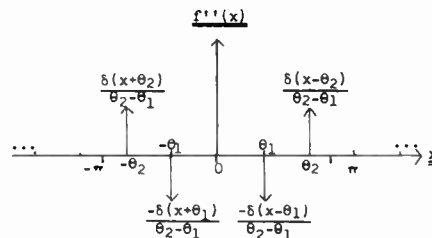


Fig. 3

$$f'(x) = - \sum_{n=0}^{\infty} na_n \sin nx.$$

Performing the differentiation on $f(x)$, one obtains the function shown in Fig. 2. It will be recognized at once that this is an easier function to deal with than the original one; even more simplicity can be obtained by taking the second derivative. Here,

$$f''(x) = - \sum_{n=0}^{\infty} n^2 a_n \cos nx.$$

Graphical differentiation of Fig. 2 yields the description of $f''(x)$ shown in Fig. 3. The Fourier Coefficients for $f''(x)$ are

$$- n^2 a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f''(x) \cos nx dx$$
$$= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{\theta_2 - \theta_1} [\cos nx][\delta(x - \theta_2)$$
$$- \delta(x - \theta_1) - \delta(x + \theta_1) + \delta(x + \theta_2)]dx$$

or

$$a_n = \frac{2}{n^2\pi(\theta_2 - \theta_1)} (\cos n\theta_2 - \cos n\theta_1),$$

by the property of delta functions given by (4). The reader can quickly convince himself that, stripped of the verbiage, this method of finding the $a_n$ is considerably simpler than the standard method.

The method described above can be applied to any $f(x)$ such that the interval $-\pi \leq X \leq \pi$ can be decomposed into subintervals on each of which $f(x)$ can be expressed as a polynomial in $x$, since if this is the case, successive differentiations will ultimately reduce the derived function to merely a sum of delta functions. Of course, it is not necessary that all the delta functions appear at once, as in the example above. For instance, if the function shown in Fig. 1 is modified by drawing parabolic arcs according to the formulas

$$f(x) = 1 - \frac{(x - \theta_1)^2}{(\theta_2 - \theta_1)^2} \quad \theta_1 \leq x \leq \theta_2$$

$$= \frac{(x + \theta_2)^2}{(\theta_2 - \theta_1)^2} \quad -\theta_2 \leq x \leq -\theta_1$$

then after two differentiations, the graphical description of $f''(x)$ is a pair of positive $\delta$ functions at $x = \pm \theta_2$ plus two square waves defined by

$$g(x) = - 2k \quad \begin{matrix} -\theta_2 \leq x \leq -\theta_1 \\ \theta_1 \leq x \leq -2. \end{matrix}$$

For this function, one may find $-n^2 a_n$, the coefficient of cos $nx$ in the formal expression for $f''(x)$, by first dividing $n^2 a_n$ into two parts, say $c_n$ and $d_n$. If now one sets

$$\sum c_n \cos nx = \delta \text{ functions}$$

and

$$\sum d_n \cos nx = g(x),$$

the $c_n$ can be evaluated immediately, and the $d_n$ can be found by one differentiation of $g(x)$.

It is often important to know how rapidly the coefficients of a particular series are decreasing for approximation purposes. There are various theorems in the theory of Fourier Series which give this information. These theorems become exceedingly clear with the insight provided by the above method for finding the $a_n$.

For example, one such theorem is that, if $f(x)$ is a continuous function, then the $a_n$ must fall off at least as fast as $K/n^2$. Now, a continuous function must have at least a piecewise continuous derivative. Hence, one must proceed at least to the second derivative before delta functions are encountered. In the formal differentiation, the second derivative is the point where $-n^2 a_n$ is the coefficient. If the second derivative is composed solely of delta functions, then the $a_n$ can be evaluated immediately; the coefficients obviously decrease as $K/n^2$. If the second derivative contains something besides $\delta$ functions, then there will be other terms in the composition of $a_n$ which will fall off faster than $K/n^2$, since more differentiations will be required to reduce this remainder to delta functions; each component of $a_n$ falls off at least as fast as $K/n^2$, however.

The mathematical basis for the method described above rests in the "theory of distributions," which provides an elegant foundation for all justifiable operations with $\delta$ functions.

J. F. GIBBONS
Bell Telephone Labs., Inc.
Murray Hill, N. J.

# Contributors___

Arthur G. Anderson (AM'51) was born on November 22, 1926 in Evanston, Ill. He served in the Navy from 1944 to 1946.

In 1949 he graduated with the B.S. degree in physics from the University of San Francisco, and in 1951 he received the M.S. degree in mathematics from Northwestern University.

In 1952 he joined the International Business Machines Corporation at Poughkeepsie where he worked on laboratory instrumentation for the measurement of machine operation and on analytical mechanics problems.

A. G. ANDERSON

Since 1953 he has been engaged in digital computer research at the IBM Watson Laboratory at Columbia University. During this time he has worked on magnetic resonance phenomena with particular application to spin-echo storage techniques and on high-speed pulse techniques with applications to fast computer elements.

Since 1952 he has been doing graduate study in the physics department at New York University.

Mr. Anderson is a member of the American Physical Society.

❖

For a photograph and biography of Seymour B. Cohn, see page 1067 of the August, 1956 issue of PROCEEDINGS OF THE IRE.

❖

Raymond C. Cumming (S'50–A'55–SM'56) was born at Missoula, Mont., on January 6, 1927. He recceived the B.S. degree in engineering physics from Montana State College in 1950, and the M.S. and Ph.D. degrees from Stanford University, Stanford, Calif., in 1951 and 1955, respectively.

During World War II, he was an instructor of electronics in the U. S. Navy. Since 1951, he has been em-

R. C. CUMMING

ployed as a full-time research associate by the Stanford Electronics Laboratories at Stanford University. He has done research on radar systems, on a number of microwave problems, and on radar echoes from meteors, and is presently engaged in exploratory research on electronic systems techniques.

Dr. Cumming is a member of Phi Kappa Phi, Tau Beta Pi, and Sigma Xi.

Stephen Doba, Jr. (A'45–SM'54) was born in New York, N. Y. on May 27, 1907. He attended the College of the City of New York and Cooper Union. He joined the Bell Telephone Laboratories in 1926 and was engaged in circuit research on voice operated devices such as Vogads, compandors, and echo suppressors.

Since 1938 Mr. Doba has been with the Transmission Development Department of the laboratories concerned with development of television transmission systems. During the war years he was occupied in the development of airborne radar equipment. In 1954 he was named Transmission Systems Development Engineer in charge of groups engaged in development of television transmission systems, submarine cable systems, and Bell System field test equipment.

S. DOBA, JR.

❖

Robert M. Frazier was born in Winchendon, Mass. on January 13, 1927. From 1944 to 1949 he served in the U. S. Navy as a communications technician.

Since 1950, he has been attending the Massachusetts Institute of Technology where he is enrolled in the cooperative course in electrical engineering and received the B.S. and M.S. degrees in September, 1956. During this time his coopera-

R. M. FRAZIER

tive and part-time work at the General Radio Company, Cambridge, Mass., has been concerned with development and design of variable delay lines.

❖

Charles M. Johnson was born in Nashville, Tenn. on May 31, 1923. He received the B.E. degree in civil engineering from Vanderbilt University in 1944 and was employed as a structural engineer until 1948 when he returned to graduate school at Duke University. He received the Ph.D. degree in physics, in 1951 for microwave spectroscopy research in the millimeter region.

C. M. JOHNSON

From July, 1951 until October, 1956 Dr. Johnson was a staff member of the Radiation Laboratory of The Johns Hopkins University, as research associate until 1953 and research scientist until 1956. In October, 1956 he joined in the formation of Electronic Communications, Inc.

Dr. Johnson is a member of the American Physical Society.

❖

F. D. Lewis (S'36–A'38–VA'39–SM'50) was born in Liberty, Mo. on July 2, 1911. He received the A.B. degree at Central College, Fayette, Mo., in 1933, and the B.S. and M.S. degrees from Massachusetts Institute of Technology in 1937 and 1940. From 1937 to 1940 he was engaged in experimental work on uhf receivers and electromagnetic horn radiators at M.I.T.

F. D. LEWIS

During the summer of 1940 he worked on Doppler-effect radar at Loomis Laboratory, subsequently going to the M.I.T. Radiation Laboratory when it opened in November, 1940.

In 1941 he went to England as scientific liaison officer for the NDRC-OSRD. Returning in 1942, he became an expert consultant in the Office of the Secretary of War on radar countermeasures and allied problems. He has been with the General Radio Company since 1945, and has been working on frequency measurement since 1949.

Mr. Lewis holds the President's Certificate of Merit. He is a member of Sigma Xi and the AAAS.

❖

Richard K. Moore (S'43–A'46–M'50–SM'54) was born November 13, 1923 in St. Louis, Mo. He graduated from Washington University in St. Louis with the B.S.E.E. degree in 1943. After graduate study at Cornell University, Washington University, and again at Cornell University, he received the Ph.D. in 1951.

After receiving the B.S.E.E., Mr. Moore was employed by RCA Victor Divi-

R. K. MOORE

sion as a test equipment engineer for about a year. He spent two years in the U. S. Navy during which time he attended the Bowdoin and M.I.T. radar schools and served as electronics and communications officer for the USS Rehoboth (AVP-50). He was em-

ployed by Washington University from 1947 to 1949 as a research engineer, doing work on diversity reception, and as an instructor of electrical engineering. From 1949 to 1951, he was a research associate at Cornell University, working principally in the field of ionospheric propagation and aurora. In 1951, Dr. Moore joined Sandia Corporation where he was a research engineer and section supervisor in charge of a program of radar terrain return, until September, 1955. He was then appointed associate professor and chairman of the electrical engineering department at the University of New Mexico, where he had been teaching night graduate courses for two years while employed by Sandia.

Dr. Moore is a member of Tau Beta Pi, Sigma Xi, and Pi Mu Epsilon. He is a member of the Wave Propagation Committee of IRE and Commission II of the USA National Committee of URSI.

❖

Norris S. Nahman (S'50–A'52) was born on November 9, 1925 in San Francisco, Calif. He received the B.S. degree in electronics from California State Polytechnic College in 1951 and the M.S. degree in electrical engineering from Stanford University in 1952. During his stay at Stanford Mr. Nahman directed a research project concerned with the design of analog computer components for the solutions of nonlinear differential equations.

N. S. NAHMAN

In late 1952 Mr. Nahman entered active service with the U. S. Army where he served as an electronic scientist assigned to the National Security Agency until 1955. His research work at NSA was concerned with semiconductor devices, square loop hysteresis devices and millimicrosecond techniques.

Since September, 1955, Mr. Nahman has been a staff member of the University of Kansas, electrical engineering department, where he is presently directing a research program in millimicrosecond techniques.

❖

Richard F. Post was born in Pomona, Calif., on November 14, 1918. He received the B.A. degree in physics from Pomona College in Claremont, Calif., in 1940 and remained there to teach physics until February, 1942, when he joined the Underwater Sound Division of the Naval Research Laboratory in Washington, D. C. There he did field and laboratory work in mine countermeasures and submarine sonar. Following the war he undertook electron accelerator studies.

R. F. POST

In 1946 he entered Stanford University, and began theoretical and experimental work on high energy electron linear accelerators under the late Professor W. W. Hansen, receiving the Ph.D. degree in physics in 1950. From 1947 to 1951 he held the position of research associate at Stanford.

In 1951, Dr. Post joined the staff of the University of California Radiation Laboratory at Berkeley. In 1952 he transferred to the newly-founded Livermore branch of the Radiation Laboratory. He was at that time assigned responsibility for the direction of some studies in controlled fusion reactions, with which he is still concerned.

❖

Donald E. Rosenheim (SM'56) was born on March 23, 1926 in New York, N. Y. His education was interrupted by service in the Navy where he attended the Radio Materiel School, and served two years as an electronic technician.

He received the degree of B.E.E. magna cum laude from Polytechnic Institute of Brooklyn in 1949. From 1949 to 1951 he did development work on "Servo-Synchronized" transmitting equipment for the Servo Corporation of America.

D. E. ROSENHEIM

In 1951 Mr. Rosenheim joined the engineering laboratories of International Business Machines Corporation at Poughkeepsie, N. Y. From 1951 to 1953 he was concerned with development work and instruction on the IBM 701 digital computer.

Since 1953 he has been at the IBM Watson Laboratory at Columbia University where he has been engaged in digital computer research, with emphasis on high-speed logic and pulse techniques. During this time he has been doing graduate study in the department of electrical engineering at Columbia University.

Mr. Rosenheim is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.

❖

S. Perry Schlesinger (M'55) was born in New York, N. Y., on October 9, 1918. He received the B.A. degree from Michigan State University in 1941, the M.S.E. degree from Union College in 1950, and is at present completing his work toward a Doctorate at The Johns Hopkins University.
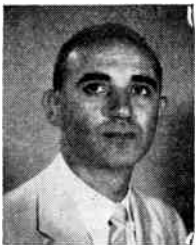
In 1946, following war service as a destroyer engineer officer, Mr. Schlesinger

S. P. SCHLESINGER

joined the General Electric Company, subsequently serving as an assistant professor of electrical engineering at Union College (1947–1950) and the U. S. Naval Academy (1950–1953). From 1953 through 1956 he was a research associate at the Radiation Laboratory of The Johns Hopkins University working on back-scattering studies, millimeter techniques and, most recently, problems in dielectric image-line transmission. At the present time he is a member of the faculty of the department of electrical engineering at Columbia University.

Mr. Schlesinger is a member of Sigma Xi and Phi Kappa Phi.

❖

Marvin Silver was born in New York, N. Y., in 1924. He received the B.E.E. from Rensselaer Polytechnic Institute in 1945. After a tour of duty in the Navy, he worked at RCA, Harrison, N. J., in 1947 on receiving tube problems and at the Franklin Institute in 1948 and 1949 on ultra-sonic generators. Mr. Silver studied physics at New York University and received the M.S. in 1951. In 1952 he went to Chatham Electronics to engage in research on problems of breakdown phenomena in gas filled cold cathode tubes. It was here that the work for the present paper was performed. He left Chatham in 1955 to resume studies for the Ph.D. in physics. Presently engaged in solid state research at New York University, Mr. Silver is also retained by the Kuthe Laboratories of Newark as a consultant on problems in cold cathode gas discharge devices.

M. SILVER

❖

Harold A. Thomas (SM'51) was born in Newport, Ore., in 1911. He received the B.S. degree in electrical engineering in 1936 and the M.S. degree in physics from Oregon State College in 1937. After taking graduate work at both Purdue University in Indiana and at Texas A&M College he received the Ph.D. degree in 1947. Dr. Thomas spent seven years in both research and teaching, first at Purdue and subsequently at Texas A&M.

H. A. THOMAS

For two years during the war Dr. Thomas served as a research engineer at the Westinghouse Research Laboratories working in the field of mass spectrometry. He joined the National Bureau of Standards in 1947 work

ing on the measurement of atomic constants. For this work he received the Department of Commerce Meritorious Award. In 1951 he was made head of the electronic proximity fuse work at the Corona Laboratories of the National Bureau of Standards, and in 1954 he was appointed Chief of the Radio Standards Division. He recently joined the General Atomic Division of General Dynamics.

Dr. Thomas is a member of the American Physical Society, RESA, Eta Kappa Nu, Phi Kappa Phi, and Sigma Pi Sigma.

❖

Ronald L. Wigington was born in Topeka, Kan. on May 11, 1932. He received the B.S. degree in engineering physics in 1953 from the University of Kansas. From graduation until the summer of 1954 he was employed by Bell Telephone Laboratories in the Electron Tube Development Department. After being inducted into the Army, he was assigned to electronics research with the National Security Agency. In June, 1956 he was released from active duty, and he remained with the National Security Agency as a civilian.

R. L. WIGINGTON

Mr. Wigington is a member of Tau Beta Pi, Sigma Pi Sigma, Pi Mu Epsilon, and an associate member of Sigma Xi.

Charles S. Williams, Jr. (S'48–A'50) was born February 20, 1920, in Altus, Okla. He attended Princeton University and the University of Oklahoma prior to four years of military service in World War II, during which he served as meteorologist and interceptor controller in the southwest Pacific.

C. S. WILLIAMS, JR.

After the war, Mr. Williams returned to school and received the B.S.E.E. degree in 1948 and the M.A. degree in mathematics in 1949, both from the University of Oklahoma. In 1956, he received the M.S. degree in electrical engineering from the University of New Mexico.

From 1949 until his recall to military service in 1951, he taught mathematics at Oklahoma University and also pursued further studies in that subject. He served as project officer in the development of an air defense system at the Rome Air Development Center, Rome, N. Y. until 1952. Since 1952, Mr. Williams has worked at the Sandia Corporation and has taught electrical engineering at the University of New Mexico. His work at Sandia has been concerned with radar terrain return studies and more recently the evaluation of the operation of certain special weapon systems in an electronic interference environment. He is now an assistant professor at the University of New Mexico and a part-time staff member of the Sandia Corporation.

Mr. Williams is a member of Sigma Xi, Phi Kappa Phi, Tau Beta Pi, and Pi Mu Epsilon.

❖

James C. Wiltse (S'48–A'53) was born on March 16, 1926, in Tannersville, N. Y. He received the B.E.E. degree in 1946 from Rensselaer Polytechnic Institute. After a year of shipboard duty in the U. S. Navy as an electronics officer he was employed by the General Electric Company for approximately a year, and from 1948 until 1951 taught in the department of electrical engineering at
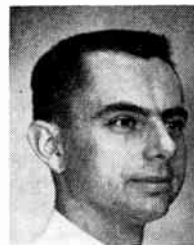
J. C. WILTSE

Rensselaer Polytechnic Institute. He took graduate work concurrently and received the M.E.E. degree in 1951.

Since 1951 he has been associated with The Johns Hopkins University as a graduate student, instructor in electrical engineering, and for the past two and a half years, as a research associate with the Radiation Laboratory, where he has been involved in research dealing with back-scattering problems and with microwave transmission lines. He is at present a candidate for the Dr. Eng. degree at Johns Hopkins.

Mr. Wiltse is a member of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu.

# IRE News and Radio Notes

## EARLY RESERVATIONS URGED FOR IRE NATIONAL CONVENTION

Plans are all but completed for the forty-fifth annual national convention of the IRE, scheduled for the Waldorf-Astoria Hotel and the New York Coliseum, March 18–21. A comprehensive program of fifty-five technical sessions is being set up by the Technical Program Committee with the assistance of all the IRE Professional Groups. Thirty-three sessions will be held at the Waldorf; the remainder, at the Coliseum. Full program details will be published in the March issue of the PROCEEDINGS OF THE IRE.

The popular social events will open with a "get-together" cocktail party on the afternoon of March 18. The annual banquet, to be held in the Grand Ballroom of the Waldorf during the evening of March 20, will feature Dr. J. A. Hannah, President of Michigan State University, as guest speaker.

Since a convention attendance of 50,000 is expected, all members are urged to place

their reservations with IRE Headquarters as soon as possible. Cocktail party tickets may be purchased at $4.00 each, and banquet tickets at $15.00.

Registration fees are $1.00 for each IRE member, $3.00 for each non-member.

## Panama City Subsection Receives Charter
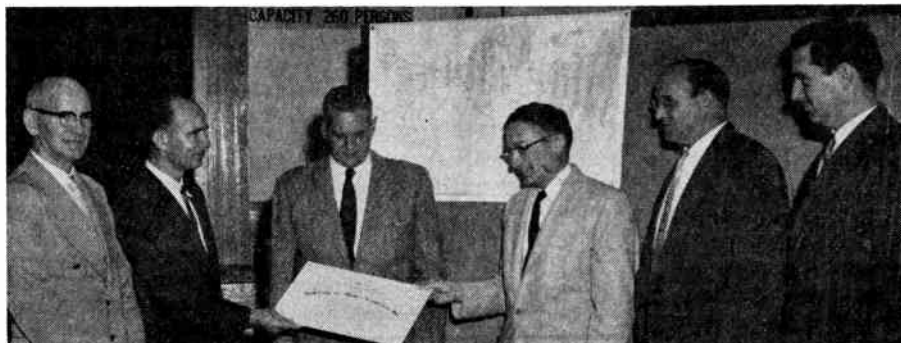


Pictured above are officers of the Northwest Florida Section presenting a charter to the newly-formed Panama City Subsection officers. *Left to right:* George Fleming, Northwest Florida Section Vice-Chairman; Fred Howard, Section Chairman; Capt. J. C. Myers, Comm. Officer of the Navy Mine Defense Lab.; C. B. Koesy, subsection chairman; Austin Haulman, subsection vice-chairman; and R. C. Aucremann, subsection secretary-treasurer.

## 1957 WESCON Officers Named

Officers of the 1957 Western Electronic Show and Convention Board of Directors were named for the event, scheduled for San Francisco's Cow Palace, August 20–23, 1957.


D. B. HARRIS

Elected to the post of Chairman for the 1957 WESCON is D. B. Harris (SM'45–F'56), manager of electron tube research at General Electric, Microwave Laboratory, Palo Alto, Calif. He took over the gavel from the 1956 Chairman, C F. Wolcott.

Vice-chairman in charge of exhibit operations for the 1957 WESCON is N. H. Moore, vice-president of Litton Industries and managing director of that company's Electron Tube Division, San Carlos, Calif. Mr. Moore has served on the WESCON Board for the past two years.

The vice-chairman in charge of convention activities for the 1957 WESCON is B. M. Oliver, director of research and development at Hewlett-Packard Company, Palo Alto, Calif. This is Mr. Oliver's first year on the Board of Directors. The IRE appointed him to a four-year term.

H. M. Stearns, vice-president and general manager of Varian Associates, Palo Alto, Calif., was elected Secretary-Treasurer for the 1957 WESCON. Mr. Stearns was named to the Board by the West Coast Electronic Manufacturers' Association.

The additional four WESCON directors are B. S. Angwin of General Electric Tube Division, E. P. Gertsch of Gertsch Products, Inc., C. F. Wolcott of Gilfillan Brothers, Inc., and Gramer Yarbrough of Elgin Electronics Division.

WESCON is the joint activity of the IRE San Francisco and Los Angeles Sections and the West Coast Electronic Manufacturers' Association.

## West Virginia Will Have a Radio Astronomy Observatory Soon

A. T. Waterman, Director of the National Science Foundation has announced that the Foundation has entered into a contract with Associated Universities, Inc., for the conduct of basic research activities in the field of radio astronomy. The contract provides for the establishment of a radio astronomy observatory at Green Bank, Pocahontas County, West Virginia. Under terms of the contract, AUI will construct the facility and provide for the "management, operation and maintenance of the observatory with the primary purpose of making its facilities available for visiting scientists."

In announcing the contract, Dr. Waterman stated that the Foundation will earmark $4 million for the fiscal year 1957 to provide the facilities and equipment needed for the conduct of research in radio astronomy. The equipment will include a precision radio telescope with a diameter of approximately 140 feet.

## International Symposium on Color TV To Be Held in Paris

An International Symposium on Physical Problems of Color Television will be held at the National Conservatory of Arts and Crafts, Paris, France, July 2–6, 1957. The International Union of Pure and Applied Physics, the French Physical Society, the French Society of Radioelectricians, and the French Society of Vacuum Engineers and Technicians are sponsoring this event.

The provisional agenda will cover properties and behavior of the human eye in color television, image analysis and restitution, assessment and measurement of picture quality, and coding procedures for transmission of color signals.

Persons intending to present papers at the symposium should send titles and abstracts of their papers, according to the UNESCO "Guide for the Preparation of Authors' Synopses," before May 1, 1957. Thirty minutes will be allotted to the reading of each paper; a fifteen-minute discussion will follow after each paper is read.

Titles and abstracts of papers to be presented, and registration fees of 1500 francs per person should be sent to the Secretary: Colloque International sur les Problèmes de la Télévision en Couleurs, Conservatoire National des Arts et Métiers, 292, rue Saint-Martin, Paris 3, France.

## PIB Sponsors International Symposium on Solid State

The seventh of a series of International Symposia sponsored by the Microwave Research Institute of the Polytechnic Institute of Booklyn will be held on April 23–25, 1957 in the auditorium of the Engineering Societies Building, 33 West 39th Street, New York City.

This symposium will cover the more recent developments in applications of electrical circuits or systems.

Emphasis will be on the phenomenological description of new or unexploited effects which may be useful in electrical circuits and the consideration of the basic limitations of these effects, as well as the application of these phenomena to electric circuits.

The cooperation of the IRE Professional Groups on Circuit Theory and Electron Devices and the co-sponsorship of the Office of Naval Research, the Air Force Office of Scientific Research and the Signal Corps permits this symposium to be held without admission charge or registration fee. Volume VII of the MRI Symposia Series, "Proceedings of the Symposium on the Role of Solid State Phenomena in Electric Circuits," will be published by October, 1957, at a cost of $5.00 per copy, clothbound. Orders for the "Proceedings," accompanied by check or money order, made out to Treasurer, Symposium Committee, will be accepted in advance. Copies of the detailed program, hotel accommodation information and registration forms are available on request. All correspondence should be addressed to: Polytechnic Institute of Brooklyn, Microwave Research Institute, 55 Johnson Street, Brooklyn 1, New York.

## Calendar of Coming Events

# Books

### Television Engineering, Volume II by S. W. Amos and D. C. Birkinshaw

Published (1956) by Philosophical Library, In·., 15 E. 40 St., N. Y. 16, N. Y. 265 pages +2 index pages +2 pages of bibliography. 156 figures. 8¼×5¼. $15.00.

This is an excellent little book on video amplifiers. It is based on an internal BBC manual written in collaboration with J. L. Bliss of the Designs Department, and edited by L. F. Ostler of the Engineering Training Department, for the engineering staff of BBC. As might be expected, it is designed to teach the reader the principles and practices of video amplifier design, rather than to dazzle him with the authors' erudition, and is therefore admirably suited for study by the better technician or engineering aide as well as by the average engineer. No mathematics beyond complex algebra and vectors is employed.

After an introductory chapter on amplitude and phase characteristics of video amplifiers in general, the book takes up high-frequency design considerations, and shows how the response of an ordinary resistance-coupled stage may be improved by the various types of peaking circuits available. The conditions for maximal flatness are given in each case, as well as compromise values for the circuit components to obtain optimum phase and amplitude characteristics. Following this, cascaded stages and distributed amplifiers and phase equalizers are discussed.

Next, the low-frequency response is treated, and the various decoupling circuits are analyzed. In connection with this, an error appears to be present in Fig. 107, where $g_s V_{sgk}$ should be used instead of $g_m V_{sgk}$, and on page 207, the term transconductance is used for $g_s$, and mutual conductance for $g_m$, which makes the reader wonder whether there is some basic difference between the two terms other than that they refer to different grids. Also on the same page, $G_s$ should be defined as a screen rather than the mutual conductance.

A further point is that the conditions for obtaining a bowed output square wave might have been discussed, as well as the methods for avoiding this effect. Also, it might have been brought out that when two tubes are used in parallel, the ratio of $g_m$ to total circuit capacity is somewhat improved because the stray wiring capacity does not double, although the improvement is considerably less than that obtained by distributed amplification.

The last part of the book deals with feedback considerations, as well as an all-too-brief discussion of noise and the design of a camera-head amplifier. However, even here there is sufficient practical information to be worth while to the design engineer, and the brief bibliography should be of further aid.

In short, this is a concise, practical, and authoritative book, and should be of great value to the average engineer not only in television, but in allied arts. There is, however, one drawback to its purchase, and that is the price: fifteen dollars!

ALBERT PREISMAN
Capitol Radio Engineering Institute
Washington, D. C.

### Photoconductivity Conference ed. by R. G. Breckenridge, B. R. Russell and E. E. Hahn

Published (1956) by John Wiley & Sons, Inc., 440 Fourth Avenue, N. Y. 16, N. Y. 640 pages +13 index pages +xiii pages. Illus. 9¼×6¼. $13.50.

This volume contains most of the papers presented at the Conference on Photoconductivity held at Atlantic City in November, 1954. The thirty papers in the book are grouped in five sections: I. Phenomenological Theory of Photoconductivity; II. Photon Absorption Processes; III. Electron Processes; IV. Photoconducting Materials; V. Current Topics.

In Section I some of the general properties and limitations on performance of photoconductors are discussed. The papers in Section II discuss the production of extra charge carriers by the interaction of light with the lattice of the photoconductor. Section III is concerned with the interactions of the optically produced charge carriers with the crystal lattice. Sections IV and V consist of recent experimental results on a variety of photoconducting materials: germanium. silicon and tellurium; the intermetallic compounds such as InSb; lead sulfide and related compounds; and zinc oxide. Some of the comments and discussion which followed the oral presentation of the papers appears in the book.

This book will be of interest primarily to active research workers in the fields covered. Much of the material has appeared in the scientific literature but it is a convenience to have such a collection in a single volume. The book does not serve as a guide to immediate practical applications of photoconductivity by the engineer nor as a textbook of photoconductivity. It is an excellent summary of the most recent work in the field and is evidence that much progress has been made in gaining a basic understanding of photoconductive processes.

D. T. STEVENSON
Lincoln Laboratory, M.I.T.
Lexington 73, Massachusetts

### Handbook of Semiconductor Electronics ed. by L. P. Hunter

Published (1956) by McGraw-Hill Book Co., Inc., 330 W. 42 St., N. Y. 36, N. Y. 132 pages +13 index pages +68 pages of bibliography +xxviii pages. Illus. 9¼×6¼. $12.00.

This book, which includes sections by fourteen contributors, is the result of an effort to collect together the major principles in the semiconductor electronics field. An idea of its scope can be obtained from the table of contents. There are twenty sections distributed amoung four main divisions: physics of transistors, diodes, and photocells; technology of transistors, diodes, and photocells; circuit design and application for transistors, diodes, and photocells; and reference material. Section headings are: 1. Transistor Characteristics; 2. Electronic Conduction in Solids; 3. Rectification in Solids; 4. Transistor Action; 5. Photoconductivity and Photovoltaic Cells; 6. Preparation of Semiconductor Materials; Methods of Preparing *PN* Junctions; 8. Metal-Semiconductor Contacts; 9. Encapsulation; 10. Device Design Considerations; 11. Low-

Frequency Amplifiers; 12. High-Frequency and Video Amplifiers; 13. Directly Coupled Amplifiers; 14. Transistor Oscillators; 15. Transistor Switching Circuits; 16. Circuits Using Special Semiconductor Devices; 17. Graphical Analysis of Nonlinear Circuits; 18. Matrix Methods of Circuit Analysis; 19. Measurement of Semiconductor-Device Parameters; 20. Measurement of Material Parameters. An extensive bibliography of the transistor, semiconductor, and rectifier fields covering the period 1940-1955 inclusive closes the book; this bibliography is the most comprehensive listing in print for these fields.

As its title indicates, the book is organized as a handbook, simplifying the location of desired information but carrying with it the disadvantage of small type, the same size as that of this review. There are very few typographical errors. The excellent selection and balance of topics, and the lack of appreciable duplication of material in sections written by different contributors are a tribute to the editor's skill and patience.

All the contributions are authoritative and reflect the wide experience of their authors. Especially noteworthy, however, are the initial sections on the physics of semiconducting devices, the section on low-frequency amplification, and that on transistor switching circuits. Many of the sections depart from being strictly a collection and discussion of past work and in addition include valuable material unavailable elsewhere. It will be a rare worker in the semiconductor field who will not find much to learn from the book. Nevertheless, its topics are so well developed from basic principles that the newcomer to semiconductors may use it to familiarize himself with almost any subject in the area. The book fulfills its aims very well, may be highly recommended, and is likely to remain for some time the standard reference in this rapidly growing field.

J. R. MacDONALD
Texas Instruments Inc.
Dallas, Texas

### Analog Computer Techniques by C. L. Johnson

Published (1956) by McGraw-Hill Book Co., Inc., 330 W. 42 St., N. Y. 36, N. Y. 246 pages +6 index pages +11 appendix pages +xi pages. Illus. 9¼×6¼. $6.00.

The material covered in this book should be a valuable aid to users and operators of dc electronic analog computing equipment. As the author points out, the analog computer and the digital computer become much more effective tools for the engineer, the scientist, and the mathematician when he understands the principles of operation of such equipment and has a working knowledge of the techniques involved in its utilization.

Brief, but in general, adequate descriptions are given of the principles of operation and design of the computing components which comprise present day analog computers. In view of the rapid changes occurring in the state of the art, references made by the author to the characteristics of com-

ponents in commercially available computers must be accepted by the reader largely as examples for the point being made by the author, rather than as information on the capability of the latest available equipment.

Good discussions are presented on how to program problems including the considerations of time scale, amplitude scale, linear and nonlinear function representation, problem stability and computer stability. One chapter presents the techniques involved in using an electronic differential analyzer for special applications such as solving algebraic equations, partial differential equations, and eigenvalue problems.

Control and checking features included in present-day computers are summarized The use of numerical analysis techniques and digital computers as auxiliary tools for analog computer checking is also covered.

Two chapters at the end of the book are devoted to related-type computing equipment; there is a discussion of the principles of operation of the repetitive type analog computer and of the digital differential analyzer.

It should be noted that two or three areas of computer operation are not included in the book. Although the drift and bandwidth problems are touched on in the discussion of time scale, certain mathematical and graphical information should be available to the operator so that he may more readily adapt his problem to the computer and also chose a computer for his problem. The latter consideration requires a discussion of using analog computers for real time simulation. More information should have been provided regarding coordinate systems and vector transformations. It would have been helpful also for the book to include a discussion of a-c electro-mechanical differential analyzers and passive network analyzers.

The author is to be congratulated for his accomplishment of assembling in one place the techniques which are being used by most of the large analog computer operating groups, in some cases, as written instructions, but most often as operator trade secrets. Although the book is written as a text for an introductory course on using analog computers, it should prove valuable to computer operators and users out of the class room.

C. M. EDWARDS
Bendix Aviation Corp.
Detroit 35. Mich.

### The Theory of Linear Antennas by R. W. P. King

Published (1956) by Harvard University Press, 44 Francis Avenue. Cambridge 38. Mass. 855 pages +18 index pages +23 pages of appendix +11 pages of bibliography +xxi pages. Illus. 10¼×7. $20.00.

The era of the linear antenna, which began with the Hertzian dipole, has reached a new pinnacle with the appearance of this 965-page treatise by Ronold W. P. King of Harvard University. It presents the advanced mathematical theory of this antenna with a thoroughness and detail never before attempted while at the same time presenting many experimental measurements for comparison.

As stated by the author, the book "is directed neither to the mathematician whose primary interest is rigor, nor to the practical engineer who desires only a final working

formula or a set of charts." Rather, it "seeks to provide a bridge from the mathematician to the practical engineer. As such, it is addressed to the applied scientist who is concerned with physical phenomena of practical importance and their mathematical representation in a form that provides both an insight into the physical aspects of the problem and reasonable quantitative accuracy in their numerical evaluation." The extensive charts and tables of both theoretical and experimental data should appeal also to many engineers, although some charts require extensive study of the text to interpret and apply them.

The various methods and theories for analyzing the linear antenna, including Hallén's integral equation theory, are treated in great detail. Arrays of linear antennas in numerous configurations, parasitic antennas, receiving antennas, base driven antennas, and dipoles above a conducting plane are some of the many specific types treated. The book brings together essential portions from the many contributions by King and his students on the subject of linear antennas over a period of many years. The volume is, indeed, a *magnum opus* that marks a milestone in the progress of antenna theory.

J. D. KRAUS
Ohio State University
Columbus. Ohio

### Introduction to Solid State Physics, 2nd ed. by Charles Kittel

Published (1956) by John Wiley and Sons, 440 Fourth Avenue, New York 16, N. Y. 570 pages +19 index pages +28 appendix pages +xvii pages. Illus. 9¼×6¼. $12.00.

This book, is an introductory textbook intended for students of physics, chemistry, engineering, and metallurgy. The presentation is designed for a reader having a moderate familiarity with modern atomic physics, and perhaps a slight acquaintance with elementary quantum mechanics. Since many topics in solid state physics cannot be treated adequately without recourse to detailed quantum mechanical and statistical mechanical reasoning, it follows that the subject matter covered is severely restricted by the assumption made regarding the reader's mathematical and physical background. However, solid state physics is so broad a subject that a good deal of ground can be covered even on an elementary level.

Representative topics in various branches of solid state physics form the content of this book. Since lengthy monographs have been written on the subject matter of each chapter, it is clear that a book of six hundred pages can hope to cover no more than a small portion of all the available material. The author has exercised excellent judgment in his choice of topics. In general, he has favored those areas which can be discussed in terms of simple, concrete, and well-developed models. These areas are usually treated quantitatively, while others are treated descriptively and qualitatively.

The present edition is two hundred pages longer than the first. Much of the added material is devoted to fuller explanations of the basic concepts, particularly in the realms of energy band theory and crystal symmetry. The rest serves to round out and bring up to date a variety of subjects, including alloys,

semiconductors, photoconductivity, luminescence, and imperfections in crystals. The revisions made enhance the value of the book considerably.

The early chapters deal with the essentials of crystal symmetry, the concepts and methods relating to wave propagation in periodic structures, and the approximate classification of solids according to their physical properties and the nature of the binding forces. The next few chapters are devoted to the classical theory of ionic crystals, the formal phenomenology of elasticity theory, and the thermal and dielectric properties of crystals. There is a good account of the theory of vibrations of one-dimensional crystals, but only a bare mention of the corresponding two- and three-dimensional problems. In the opinion of this reviewer, the theory of vibrations of two- and three-dimensional crystals deserves a more extensive treatment than it receives.

The author is at his best in those areas where he has made important original contributions, namely, ferro- and antiferro-electricity, dia-, para-, ferro-, and antiferro-magnetism. Particularly well done are the sections on nuclear and electronic spin resonance, ferromagnetic resonance, and cyclotron resonance.

There are also chapters on the free electron model of metals, the band theory of solids, and the theory of Brillouin zones. The band theory is applied in a descriptive and qualitative way to account for a number of the physical properties of metals and alloys. The central physical features of semiconductor crystals, and the physical principles of rectification and transistor action are amply treated. Briefly surveyed are the basic experimental facts and the current theories of superconductivity.

The remainder of the book is devoted to imperfections in solids. As in other portions of the book, the author is able to do no more than sample the wealth of experimental and theoretical results available. There are introductory discussions of lattice vacancies, diffusion, color centers, excitons, photoconductivity, and luminescence. The need to sample vast areas often leads to unevenness of treatment, but this is certainly excusable in an introductory textbook. The final chapter, which is concerned primarily with the interpretation of the plastic properties of crystalline solids in terms of the theory of dislocations, is extremely well done, being clear, concise, and up-to-date.

This book can be heartily recommended to the beginning student. He will find it an excellent introduction to a wide and important branch of physics. There are ample references to the original literature and to more advanced accounts. A well chosen set of problems appears at the end of each chapter; this feature enhances the pedagogical value of the book substantially.

Professor Kittel is well known in the solid state field not only for his numerous and varied original contributions, but also for his great gifts as a teacher and lecturer. His abilities as a creative and versatile thinker, and as a lucid and enthusiastic expositor are manifest on nearly every page.

FRANK HERMAN
RCA Laboratories
Princeton, N. J.

## Comment on Review of "Reliability Factors for Ground Electronic Equipment"

On page 1073 of the August issue of the PROCEEDINGS, Mr. C. R. Knight criticizes rather forcefully a book, "Reliability Factors for Ground Electronic Equipment,"[1] which the writer of this letter edited. Mr. Knight offers several examples of why he thinks the book "falls far short of the excellent reputation of (the) publisher," McGraw-Hill, who produced the book under an Air Force contract. Mr. Knight states that the examples he cites are illustrative only and that the "list should not be considered comprehensive." The implication is, of course, that Mr. Knight found other examples of poor editing and publishing and by not citing any examples of what he considered as worthwhile material, he implies that there is nothing worthwhile in the book.

After spending over thirty years writing and editing his own books and reading, reviewing and criticizing the works of others, the writer has come to certain conclusions on the purpose of a book review.

As a purchaser and reader of other people's books, this writer looks upon a book review as a free look at what the book is supposed to do and how well it does it. He would like to know the purpose of the author or editor and the audience aimed at. He would like to know what is good and bad about the book so he can judge for himself the overall merit for his own purpose. He would like to know what is in the book, and perhaps something about the author or editor so he can determine if the opus comes from experience or was merely a "library book" with long lists of references which the author never saw. In short—what the author tried to do, what he actually did, and what the book has to offer the reader.

A book review should not be primarily a spring board from which the reviewer can take off on his own to vaunt his own knowledge, or lack of it. After all, a book takes at least a year of a man's life to prepare and many thousands of dollars of a publisher's funds to produce. All this can be lost by an unfavorable review written in an hour by someone who may not have judged the contents in terms of the author's intent. In other words, the review should be fair.

Mr. Knight is correct in parts of his review.

The book is out of date. No technical book in a rapidly moving field like electronics can be up to date. If it takes a year to gather the material, digest it, make the drawings and write the manuscript, it will take another year to get the final books ready for reviewers to shoot at.

There are, of course, quicker ways of getting from concept to manuscript or even from manuscript to reader, but they do not seem to measure up to what reputable writers and publishers and critical readers feel are desirable.

This particular book was contracted for in December, 1953; actual gathering, digesting, writing and editing the material and

[1] "Reliability Factors for Ground Electronic Equipment" ed. by Keith Henney, reviewed by C. R. Knight, PROC. IRE, vol. 44, pp. 1073–1074; August, 1956.

getting it ok'ed by the Air Force took place within the period December, 1953–March, 1955. The required one hundred copies were delivered in May, 1955, and promptly proved not to be sufficient to satisfy the demand for this, the first, book on reliability. Another nine months were needed to get through the red tape required for the publisher to produce a commercial version of a document for which the U. S. Government had footed the bill.

Thus between the time the first words were written and the time when Mr. Knight probably got his review copy, 28 months had gone by. Of course the book was out of date.

Mr. Knight is also correct on one other matter. The legends for three figures were incomplete. For Mr. Knight's benefit there are other typographical errors which neither he nor the editor has found to date. This is always true.

Other comments of his seem to involve matters of opinion or interpretation and are not subject to yes or no discussion.

Mr. Knight also critizes several technical matters. Here, the writer must defer to the experts. He has received comments from Ray Miles who wrote the chapter on statistical methods and from N. B. Ritchey and A. L. Wilson who wrote most of the material on vacuum tubes.

Mr. Knight objects that Chapter 4, titled "Mathematical Approach to Reliability," fails to distinguish between population and sample, and that omission of this distinction is dangerous.

Mr. Miles says: "The matter of statistical sampling procedures, interpretation of sample characteristics in terms of probable characteristics of the complete population, etc., is a subject requiring a fairly advanced knowledge of probability and statistics, and was certainly beyond the scope of Chapter 4. The purpose of Chapter 4 was to give the reader, previously unfamiliar with probability and statistics, a basic understanding of the mathematical nature of reliability, not to equip him for a career as a professional statistician.

"However, Chapter 4 does not by any means overlook the distinction between population and sample, and it certainly does not ' . . . infer . . . to the uninitiated that one failure in two observations has precisely the same meaning as ten failures in twenty observations.' On the contrary, Chapter 4 specifically cautions the reader to beware of drawing false conclusions on the basis of data that is either insufficient in quantity or unrepresentative in kind. Thus,

' . . . a large number of such tests must be made to arrive at any accurate result expressing the reliability of the equipment type.'

'A blind statistical approach, without regard for the dictates of common sense, is particularly fruitful in its possibilities for arriving at incorrect conclusions.

'In statistical treatment of reliability, the most serious possible sources of error include . . . basing a firm conclusion on a quantity of data so small as to have little or no statistical significance.

'In all likelihood, . . . 10 radars is not a sufficient number for an accurate conclusion.

'To be useful for the quantitative predic-

tion of . . . reliability, reliability data must . . . be statistically significant. That is, it must represent the results of a sufficient number of tests to be a valid basis for statistical prediction, and the test samples must be truly representative of the component parts (population) to which the prediction is to be applied.' The parenthesis has been added to point up the very clear distinction that this statement makes between the sample and the population.

'A report by the U. S. Naval Air Missile Test Center includes a discussion of the relation between the number of samples (tested) and the degree to which the test results are indicative of the characteristics of all the component parts in the sampled lot.' (The report in question is specifically identified in Chapter 4's list of references.)"

Mr. Miles cites other statements in his chapter which, if read, would prevent the reader from making invalid assumptions, and the present writer will be glad to point them out to any reader who may wish this information.

On page 8-5 of the book the authors stated that the theory of electron emission is not yet completely developed nor is the mechanism fully developed. Mr. Knight says "Anyone familiar with electron tube physics would recognize (this) as an inexcusable misstatement of fact." On this matter Messrs. Wilson and Ritchey say "We in the tube industry know that emission 'busts' continue to occur. As recently as 1955, one of the major tube companies experienced an emission 'bust,' and all of the 'electron tube physicists' on the payroll with all their most recent theories did not solve the problem any faster than in the good old days when we got out of a bust by witchcraft. See J. J. Coupling (J. R. Pierce), *Science Fiction*, November, 1946."

Messrs. Ritchey and Wilson also say:
"Mr. Knight is unfamiliar with the fact that more than one major paper a month has been published since 1946 on the subject of emission, and that work is continuing in many laboratories the world over to try and put a firm foundation under the theory of oxide-coated cathodes. A partial bibliography of 166 papers on work done in the field from 1946 to 1954 is attached plus a few items from a more recent bibliography dating from 1954 and including 1956. As Mr. Knight can see, the last word has not yet been written on the subject of emission." (Mr. Knight and any reader may have a copy of these bibliographies if he wishes to read up on the subject. They include eleven items published in 1955 and six in 1956.)

"Mr. Knight found fault with Fig. 8-20 on page 8-23. It is his contention that the curve of ionic grid current is in error and should keep increasing as the bias goes positive. He states his criticism in these words: 'That ionic grid current is directly proportional to plate current has been an established fact of physical electronics for many decades.' This statement could not be more in error.

"There are several mechanisms within a tube that will cause the ionic grid current curve to follow that shown in Fig. 8-20. Mr. Knight is obviously unfamiliar with the fact

# Abstracts of IRE Transactions

## Broadcast & TV Receivers

### Vol. BTR-2, No. 3, October, 1956

**Stagger-Tuned Transistor Video Amplifiers**
—V. H. Grinich (p. 53)

A method of analyzing common-emitter transistor video amplifiers has been given that is sufficiently simple and accurate for design work. This paper discusses the use of the above design method for the purpose of obtaining RC-coupled stagger-tuned video amplifiers. For a poles-cancelling-zeros scheme, the design equations are given relating external circuit parameters to desired response shape and transistor parameters. By means of stagger tuning, the gain-times-bandwidth factor of a transistor amplifier is kept at unity as in the case of stagger-tuned vacuum tube amplifiers. Experimental results are given for amplifiers that are designed to be maximally flat by this method.

**Analytical Approaches to Local Oscillator Stabilization**—W. Y. Pan and D. J. Carlson (p. 57)

The behavior of local oscillator tubes and associated circuit elements under the complex conditions of heat flow can be treated analytically. By means of such an analytical treatment, local oscillators at frequencies up to and beyond 1000 mc may be stabilized methodically with conventional temperature-sensitive elements. The resultant frequency stability is generally satisfactory for most practical purposes.

The analytical approaches to local oscillator stabilization in general and the special considerations at uhf are fully illustrated with two types of commercial television tuners. The vhf television tuner exhibited a maximum residual frequency deviation of ±50 kc, while that of the uhf television tuner was ±100 kc after either tuner had been energized for a period of two minutes.

It is believed that the same general approaches and considerations can be utilized to stabilize local oscillators for other applications.

**Retrace Driven Deflection Circuit**—W. B. Guggi (p. 65)

A simple, efficient cathode-ray tube deflection circuit suitable for standard television horizontal sweep applications with excellent linearity has been developed. The circuit employs either a single transistor or a tube, operated as a switch. The sweep amplitude may be adjusted continuously without affecting the linearity or general performance of the circuit. The switching operation is such that the inrush current upon closing the switch and the voltage build-up upon opening the switch are delayed sufficiently to prevent excessive momentary dissipation in the switching device. This and other circuit characteristics makes it possible to operate the switching device efficiently, particularly when the circuit is transistorized, and therefore permits exceptionally wide deflection angles with a relatively small transistor.

**The Possibilities of Reduced Television Bandwidth**—Sid Deutsch (p. 69)

A solution is sought to the following problem: Given a wide bandwidth of the order of 50 kc, devise standards such that the entertainment value shall be maximized. Four applications are considered: tape recording at a tape speed of 15 inches per second; telephone conversation accompaniment over high-quality telephone lines; long-distance wire-line transmission over high quality telephone lines; and short-wave transmission over a 120 kc channel. Bandwidth reduction is achieved through reduced resolution, reduced field frequency, increased interlacing ratio, the use of dot interlace, and the use of quantizating and coding.

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* |
|---|---|---|---|---|
| Broadcast & TV Receivers | Vol. BTR-2, No. 3 | $1.05 | $1.55 | $3.15 |
| Circuit Theory | Vol. CT-3, No. 3 | 1.00 | 1.50 | 3.00 |
| Electronic Computers | Vol. EC-5, No. 3 | 1.05 | 1.55 | 3.15 |
| Electron Devices | Vol. ED-3, No. 4 | 1.45 | 2.15 | 4.35 |
| Medical Electronics | PGME-7 | 1.00 | 1.50 | 3.00 |
| Nuclear Science | Vol. NS-3, No. 4 | 4.50 | 6.75 | 13.50 |
| Vehicular Communications | PGVC-7 | 1.20 | 1.80 | 3.60 |

\* Public libraries and colleges may purchase copies at IRE Member rates.

The scanning frequencies and synchronizing signal are derived for each case. In addition, for the short-wave system, a method of combating multipath transmission is suggested.

## Circuit Theory

### Vol. CT-3, No. 3, September, 1956

**Abstract of Papers in This Issue (p. 162)**

**Synthesis of Tchebycheff Impedance-Matching Networks, Filters, and Interstages**—G. L. Matthaei (p. 163)

Work of R. M. Fano on the theoretical limitations of broad-band impedance-matching is utilized to give a straightforward procedure for synthesis of optimum, lossless, Tchebycheff, wide-band impedance-matching networks for various classes of loads. The loads which are treated consist of an inductance, capacitance, and a resistance (all of arbitrary size) connected in series, parallel, or any of a variety of series-parallel combinations. Procedures are described and examples are presented for low-pass, high-pass, and band-pass matching network design for use with specified loads so as to meet prescribed bandwidth and reflection coefficient magnitude specifications (*i.e.*, any specifications that are physically realizable). The design procedure is simplified by the use of tabulated reflection coefficient polynomials, each of which is a function of frequency $p$ and a design parameter. The application of this procedure to Tchebycheff filter and vacuum-tube amplifier, interstage design, is also discussed.

**On Electrical Circuits and Switching Circuits**—Sundaram Seshu (p. 172)

From an abstract point of view, both electric circuits and switching circuits may be considered as weighted nonoriented graphs. There are two main differences between electric and switching circuits from this point of view. The first of these is the algebra to which the weights belong. In electrical network theory the weights belong to the complex number field or to the field of rational functions of a complex variable since they are impedances or admittances. In switching theory they are Boolean functions. The second difference is that one is interested, in electrical network theory, in the circuits or the "loops" in the system, whereas in switching theory one is interested in the paths. This paper seeks to relate the two theories by means of topological considerations. Formulas are derived relating the switching function and the

driving point admittance function of a two-terminal network. Certain relations between dual networks are also established. The paper concludes with a synthesis procedure for a type of switching circuit of academic interest, the single contact switching circuit, and statements of some important unsolved problems.

**On Networks Without Ideal Transformers**—Israel Cederbaum (p. 179)

In the above paper necessary conditions for a matrix to be an impedance or admittance matrix of pure resistance $n$ ports are given. The results are then applied to RLC networks for real positive values of the complex frequency, and to the most general linear reciprocal networks without ideal transformers at zero frequency. Some properties of the matrices fulfilling the above conditions are given in the last section of the paper.

**Linear Pulse-Forming Circuits**—W. C. Gore and T. Larsen (p. 182)

The purpose of this paper is to present a generalized method for the design of linear pulse-forming circuits, *i.e.*, circuits which have pulse-type responses to a step-function input, which would be applicable for a variety of pulse shapes. Although the problem has been treated previously by solving for the circuit impedance necessary to produce a specified pulse shape, practical difficulties are usually encountered when one tries to construct the required impedance from available components. These difficulties are avoided by a different method of setting up the circuit problem. This method, in essence, consists of regarding the circuit as composed of two parts, a pulse-forming network and a load or termination; and the desired pulse shape as the sum of two identically shaped waveforms displaced in time. As a result of this approach, a solution of the problem of determining the load or termination is obtained in terms of the impedance of the pulse-forming network which is easily recognizable. Design equations are given and experimental results are shown for the new method.

**RLC Transfer-Function Synthesis**—E. C. Ho (p. 188)

This paper considers an RLC transfer-function synthesis based on an extension of a real part decomposition technique as applied to the synthesis of a class of driving-point immitances. The basic concept is due to a valuable relationship between the short-circuit parameters existing in the network (connected as a two-terminal-pair network) realized for a driv-

ing-point admittance function by the real part decomposition technique using its simplest decomposition form. The synthesis procedure involves defining a suitable half-admittance matrix from the prescribed transfer function, such that a network can be realized by the real part decomposition technique alone. While the procedure represents merely a further exploitation of the real part decomposition technique, it is, however, extremely simple to apply, and can be considered an alternative to the already existing RLC transfer-function synthesis procedures.

**A Theorem Concerning Noise Figures**—A. G. Bose and S. D. Pezaris (p. 190)

A theorem is formulated and proved which determines the greatest lower bound of the single-frequency noise figure of a general system consisting of $n$ amplifying devices and passive coupling elements. According to the theorem, this lower bound is equal to the noise figure of an optimum system using a selected one of these amplifying devices.

**Reviews of Current Literature—Approximation of a Two-Pole Function, Whose Components are Prescribed Within a Limited Frequency-Band** (Ermittlung der Zweipolfunktion, deren komplexe Werte in einem Teilbereich reeller Frequenzen vorgeschrieben sind)—Werner Krageloh. . . . Reviewed by J. F. Peters (p. 197)

**Procedures to Reduce Noise and to Increase Channel Capacity with Pulse Code Modulation** (Massnahmen zur Störverminderung und Erhöhung der Kanalkapazität bei Impulskodemodulation)—F. Benz. . . . Reviewed by J. F. Peters (p. 197)

**Time Characteristics of Pulsed Signals Which Have Passed Through a Linear System**—G. A. Levin and B. R. Levin (in Russian). . . . Reviewed by J. T. Bangert (p. 198)

**Abstracts of Foreign Language Articles on Circuit Theory—On a Simple Method for the Numerical Calculation of Time Functions for the Case of Fractional Rational Functions in the Subdomain**—R. Hofmann and W. Walcher (in German) (p. 199)

**Simultaneous Approximation of Amplitude and Delay Time of an Ideal Low-Pass Filter by Means of a Flow Analogy**—J. Peters (in German) (p. 199)

**Active Four-Terminal Interstage Networks**—H. Beneking (in German) (p. 199)

**The Operating Parameter Cascade Matrix of Four Poles**—F. L. Bauer (p. 199)

**The Impedance Transformation of Linear 2n Poles**—Hans Kleinwachter (in German) (p. 199)

**The Formation of a Six Terminal Curve, Its Graphical Evaluation and Its Use for Determination of Transfer Properties of Loss Free Six and Eight Terminal Devices**—Heinz Lueg (in German) (p. 200)

**Spectral Density of Cross Modulation Noise**—R. Codelupi (in Italian) (p. 200)

**Transformation Rules for Noise Four Poles**—W. Dahlke (in German) (p. 200)

Correspondence (p. 200)
PGCT News (p. 203)

## Electron Devices

### VOL. ED-3, No. 4, OCTOBER, 1956

**Experimental Investigation of the Transient Behavior of Gold—Germanium Surface Barriers**—O. Curtis, Jr. and B. R. Gossick (p. 163)

Following the application of a current pulse to a semiconductor diode, a transient voltage appears which is caused by deviations in carrier concentrations near the barrier. The high-frequency capabilities of a diode are intimately related to the decay time of this transient voltage, which is treated herein for a variety of experimental conditions.

The open circuit voltage following a small amplitude impulse current was investigated

with different values of fixed bias. The recovery time was observed to decrease with increasing forward bias, and the voltage was seen to decay exponentially with time given sufficiently forward bias.

The effect of the duration of the injection upon the voltage decay was also investigated. It was noted that there was a rapid initial decay which was faster with shorter injection, and a final exponential decay which was insensitive to the duration of the current pulse; however, it was dependent upon the bulk lifetime of the material.

The influence of the surface barrier itself upon the transient behavior was also investigated. In general the recovery time was found to be faster with decreasing injection ratio.

The open-circuit voltage following a large amplitude current pulse was investigated using specimens with near unity injection ratios. The final voltage decreased linearly with time, the slope being determined by the bulk lifetime.

**Effect of Nonlinear Collector Capacitance on Collector Current Rise Time**—T. R. Bashkow (p. 167)

The collector capacity, $C$, of a junction transistor is known to vary as a nonlinear function of the voltage, $V$, across it. A calculation is made of the collector current rise time of a grounded emitter alloy junction transistor for which $C = kV^{-1/2}$. A comparison is then made with linear analyses in which $C$ is assumed to have one of the following constant values. 1) $C = C_{cc}$, where $C_{cc}$ is the small signal capacity measured at the collector supply voltage, ($V_{cc}$. 2) $C = 1.52 C_{cc}$. This capacity is one which displaces the same charge as the nonlinear capacity as the voltage across it changes from 0–90 per cent of its final value. 3) $C = 2C_{cc}$. This capacity is one which displaces the same charge as the nonlinear capacity as the voltage across it changes from 0–100 per cent of its final value. The linear analysis using the latter two capacity values gives 0–90 per cent and 0–100 per cent rise times which are very close to those given by a numerical solution of the nonlinear circuit equation. The usual linear analysis using $C = C_{cc}$, on the other hand, is very much in error for predicting rise time.

Experimental results show that the $2C_{cc}$ value, in a linear analysis, predicts the 0–100 per cent rise time almost exactly. In addition, analog computer solutions of the nonlinear circuit equation give results almost identical with the shape of the experimental curves.

**Breakup of Hollow Cylindrical Electron Beams**—R. L. Kyhl and H. F. Webster (p. 172)

The experimental breakup of a hollow electron beam focused by a longitudinal magnetic field into a discrete set of vortex filaments is described. Agreement with the theory of J. R. Pierce is satisfactory. The theory is extended to include the cases of Harris flow and hollow beams focused by a negative center conductor and a magnetic field. It is predicted that growing waves are possible in Harris flow beams. General considerations of energy and momentum conservation support the detailed analysis.

**Instability of Hollow Beams**—J. R. Pierce (p. 183)

This paper gives a linearized theory for the breakup of magnetically focused hollow beams of electrons. A zero-thickness developed beam is assumed. Growing waves are found to be possible both at zero frequency and at finite frequencies. When there are electrodes close to the beam inside and out, the waves grow more rapidly as the electrodes are moved further from the beam. If the current is increased and the magnetic field is increased just enough to keep the beam together, the rate of growth is increased. The theory predicts a greater rate of growth for a higher number $n$ of cycles of variation around the circumference of the beam, but for actual beams of finite thickness the theory is inaccurate for large values of $n$. In a simple

case, frequency becomes important only when the wavelength of waves along the beam becomes comparable with the wavelength measured around the circumference of the beam. Increasing waves are also found in a zero-thickness beam in crossed electric and magnetic fields, as in a magnetron amplifier or carcinotron, but not at zero frequency.

**Modified Contra-Wound Helix Circuits for High-Power Traveling-Wave Tubes**—C. K. Birdsall and T. E. Everhart (p. 190)

Experiments performed on modified forms of the Chodorow-Chu contra-wound helix circuit are presented with emphasis on those properties useful in the design of high voltage, high power traveling-wave tubes. Velocity and impedance measurements are shown for a fair range of pitches, crossover angles, wire widths, and wire thicknesses. Impedances are compared with those of a circuit having ideal fields (and the same phase and group velocities) and are found to be very good. Loading effects of glass and metal cylinders (envelopes), current paths, a second mode, and transitions from helix to waveguide are described. It is shown that periodic supports (stubs), used to make the structure much more rugged mechanically and capable of large heat dissipation, also increase the (already high) impedance almost as much as they decrease the group velocity, implying that the stubs add but little stored energy.

Experimental Notes and Techniques (p. 205)
Index to IRE Transactions on Electron Devices—Vol. ED-3, 1956 (p. 207)

## Electronic Computers

### VOL. EC-5, No. 3, SEPTEMBER, 1956

**The Representation of Constraints by Means of an Electronic Differential Analyzer**—D. T. Greenwood (p. 111)

The use of high gain amplifiers is shown to be helpful in the representation of constraints. This method enables one to represent constrained systems in a manner such that all coordinates are available for the application of arbitrary forces or displacements. The procedure is explained by means of an example.

**High-Speed Shift Registers Using One Core Per Bit**—V. L. Newhouse and N. S. Prywest (p. 114)

A three, and a two winding per core, high-speed, current driven, one core per bit shift register is presented together with an analysis of the basic circuit involved. An intermediate storage capacitor is used between successive logical elements. The charge and discharge of this capacitor are controlled in a positive manner by voltage blocking pulses. The undesired feedback of energy from one stage to earlier stages is thereby prevented, giving high efficiency of operation. The three winding per core register described is reversible and capable of operating in the megacycle range. The application of the basic shift register element to computer logic applications is described.

**High-Speed Flip-Flops for the Millimicrosecond Region**—Z. Bay and N. T. Grisamore (p. 121)

The problem of designing high-speed flip-flops has been approached by dividing the circuit operation into steady-state and switching functions, the steady-state function being assigned to a slave flip-flop and the switching function to a driving circuit. Circuits, using conventional components and having a resolving time of 10 m$\mu$ sec, are described. Resolving times as low as 2 m$\mu$ sec have been attained by using a special beam-deflection tube as the slave flip-flop. These circuits dissipate considerably less power in the intervals between switching than conventional circuits.

**A Topological Method for the Determination of the Minimal Forms of a Boolean Function**—R. H. Urbano and R. K. Mueller (p. 126)

The topology of the $n$-dimensional cube is

used to reduce the problem of determining the minimal forms of a Boolean function of $n$ variables to that of finding the minimal coverings of the essential vertices of the basic cell system associated with the given function. The proof of this statement is contained in the central Theorem 4. A numerical easily programmed procedure is given with which it is possible to treat problems with a greater number of variables than has heretofore been practical. The procedure bypasses the determination of the basic cells (the prime implicants of W. V. Quine) and locates the essential vertices, from which in turn the irredundant and minimal forms are obtained.

**Logic Circuits for a Transistor Digital Computer**—G. W. Booth and T. P. Bothwell (p. 132)

The reliability and performance which can be achieved in high-speed switching circuits using presently available high-frequency junction transistors suggests the use of transistor circuits in preference to more conventional circuits for many applications. The area of airborne digital computers is one in which the physical characteristics of the transistor as well as of associated low dissipation components can be most favorably exploited. We have presented here a group of circuits which fulfill the requirements of speed, low dissipation, size, and weight for most such applications. The circuits shown will operate over the range $-30°C$ to $+60°C$, and actually have been operated over the range $-50°C$ to $-90°C$. Low dissipation of the circuits places minimum requirements on power supplies, and temperature control may be achieved with only a small amount of cooling. An estimate can be obtained from an example. For a large-scale computer, say 2000 transistors, 800 of which are in bistable circuits, total dissipation will be less than 140 watts.

**Correspondence (p. 139)**

**Symposium on the Impact of Computers on Science and Society (p. 142)**

**Some Automatic Digital Computers in Western Europe**—N. M. Blachman (p. 158)

**Correction (p. 167)**

**Contributors (p. 168)**

**PGEC News (p. 169)**

**Reviews of Current Literature (p. 171)**

## Medical Electronics

### PGME-7, December, 1956

**A Low-Level High-Speed Switching System for Brain Mapping**—Carl Barus (p. 1)

In recent experiments J. C. Lilly has obtained contour maps, in rapid sequence, depicting electrical activity over selected areas of the cortex of an animal's brain. The present work has been an attempt to improve the resolution in both time and space of this mapping technique. Adequate resolution requires more than 200 electrodes covering a square centimeter of the cortex. A complete map must be observed 500 to 1,000 times a second.

In order to handle the required number of information channels, a time-division switching system has been devised. The electrodes are divided into groups of sixteen. Simultaneously operating switching units commutate the information within the respective groups. An amplifier for each group passes a train of samples representing the potentials at the electrodes of that group. These simultaneous sample trains can be recorded on multitrack magnetic tape. Later the sampled information can be displayed in map form, e.g., on a cathode-ray screen.

The electronic switching of the sixteen electrode signals in a group takes place after a single stage of preamplification for each electrode. The switching elements are matched pairs of silicon junction diodes connected, with

resistances, in bridge circuits. A master pulse generator operates all the switching units. The system is sensitive to electrode potentials of about 30 $\mu v$.

**Potentials Produced by an Eccentric Current Dipole in a Finite-Length Circular Conducting Cylinder**—R. H. Okada (p. 14)

The electric potential produced by a current dipole located arbitrarily within a finite length circular conducting cylinder is obtained in the form of a double summation. The impulse function is employed to satisfy the end cap boundary condition and to reduce a definite integral to a summation. Special cases include expressions for potentials on the cylinder surface and in an infinite slab of conducting medium bounded by two parallel planes.

(*Papers Presented at WESCON Convention, Aug. 21–24, 1956, Los Angeles, Calif.*)

**A Projection Microphotometer for Quantitative Microscopy**—R. C. Bostrom (p. 20)

Quantitative measurement of the characteristics of microscopic objects has become an increasingly important technique, particularly in the field of biology. This paper describes an instrument used for obtaining the size and optical density of objects on the specimen slide. Although methods for making these measurements have long been available, they are usually tedious and time-consuming, and little attention has been paid to the adaptation of the system to human requirements for comfort and ease. This is particularly true for cytological measurements where variations in individual cells are such that large numbers of cells must be selected before the results are statistically significant. A projection system overcomes many of the disadvantages of the other methods by not constraining the eyepoint to a fixed position, permitting direct measurement of size or area with a scale or planimeter, and allowing the photometer field to stop to be always under direct visualization. The design and performance of a projection microphotometer and its application to several biological measurement problems is thoroughly discussed.

**Communication by Vibratory Tactile Stimuli**—Joseph Hirsch (p. 29)

**Exposure Hazards from Cosmic Radiation in Flight in Extra-Atmospheric Regions**—H. J. Schaefer (p. 38)

(*WESCON Program Abstracts*)

**Progress in the Field of Xerography**—D. B. Slauson (p. 45)

**Methods of Recording Blood Pressure and Blood Flow**—J. P. Meehan, Jr., and H. I. Jacobs (p. 45)

**Cybernetics, Past, Present and Future**—L. B. Gardner (p. 46)

**Some Philosophical Bearings on Cybernetics**—Abraham Kaplan (p. 46)

**On the Role of Neural Network Models**—N. M. Martin (p. 47)

**Spatial Perception: The Development of Dimensional Awareness**—R. D. Tschirgi (p. 47)

**Studies in the Neurophysiology of Learning**—J. A. Gengerelli (p. 48)

**Correction to Energy Densities of Microwave Radiating Systems**—W. E. Tolles and W. J. Horvath (IRE Transactions PGME-4, February, 1956) (p. 49)

## Nuclear Science

### Vol. NS-3, No. 4, November, 1956

**Introduction**—S. Krasik (p. 2)

(*Proceedings of the Fifth Scintillation Counter Symposium, Washington, D. C., February 28–29, 1956*) (p. 2)

**Survey of Pulse Height Analyzers**—W. A. Higinbotham (p. 3)

An excellent review of the pulse height analyzer field was published in 1952.

The principal new developments have been in analyzers having a large number of channels,

which make use of digital computer techniques. The characteristics of these analyzers will be discussed. The gray wedge analyzers for high speed applications and other new ideas will be described. An effort will be made to suggest the advantages and disadvantages of the different types for various applications.

**Multi-Alkali Photo Cathodes**—A. H. Sommer (p. 8)

The quantum efficiency of a photoemitting material is determined by a combination of physical properties. It has been found that the requirements are met best by semiconductors of low work function surfaces. In the light of these general considerations, a short survey is given of the photocathodes available heretofore. A more detailed description will then be given of the new multi-alkali cathodes, particularly those of the composition of Sb-K-Na and Sb-K-Na-Cs. Their advantages for use in scintillation counters are discussed.

**Techniques and Theory of Fast Coincidence Experiments**—Zoltan Bay (p. 12)

A review of fast coincidence circuits, working in the millimicrosecond region, is presented. The main points covered are: 1) operating characteristics of the various types of coincidence circuits, 2) description of the various circuits, 3) differential coincidence method, 4) analysis of coincidence curves and the determination of time delays, 5) the resolving time, 6) effect of random time lags on coincidence curves, 7) experimental investigation of random time lags, 8) coincidence efficiency, and 9) statistical accuracy of time measurements.

**Transmission Secondary Electron Multiplication for High-Speed Pulse Counting**—E. J. Sternglass and M. M. Wachtel (p. 29)

A new type of high-speed electron multiplier employing transmission secondary electron emission from thin insulating films is described. Electrons from a photosurface are multiplied in a series of plane-parallel dynodes consisting of a thin scattering layer of a heavy metal followed by a layer of a pure crystalline insulating material of high secondary emission yield. The characteristics of such dynodes are described, and the dc and pulse performance of an experimental seven-stage device employing this principle is presented.

**Recent Developments in Multiplier Phototubes**—B. R. Linden (p. 33)

A number of developments on multiplier phototubes have taken place in these laboratories since the last scintillation counter symposium. The most important of these developments are discussed and detailed operating data on various tubes are presented.

One development of interest has been the large area multiplier phototubes. These tubes have been constructed with usable cathode diameters of 11 inches, $14\frac{1}{2}$ inches, and 20 inches. On the other extreme, ten stage multiplier phototubes with an over-all outside diameter of $\frac{3}{4}$ inches have been developed and the characteristics of these tubes will be presented. A program of ruggedization of multiplier phototubes has been carried out on all standard tubes and the results of this program have led to tubes with better mechanical stability in the field.

Work is also described on cathode surfaces, on secondary emission surfaces, and on improvements of the dark current characteristics of multiplier phototubes.

**Gaseous Scintillation**—C. Eggler and C. M. Huddleston (p. 36)

The method of detecting the passage of charged particles through noble gases by analyzing the light of de-excitation will be discussed. A description is given of the mechanism, and various experimental methods are presented. Attention is directed to performance, applications, and some unsolved problems.

**Alkali Halide Scintillators**—Wesley Van Sciver (p. 39)

The spectra of the luminescence from unactivated NaI and thallium activated NaI have been measured in the temperature range $-190°C$ to $+20°C$ using alpha particles, gamma rays, and ultraviolet light as sources of excitation. The characteristic emission from NaI ("pure") is in a band centered at approximately 300 m$\mu$. The intensity is weak at room temperature but increases by a factor of ten in cooling the crystal to $-190°C$. By comparison, the characteristic emission from the commonly used NaI($10^{-3}$Tl) is in a band centered at approximately 420 m$\mu$ and changes in intensity by a factor of less than two over the above temperature range. The intensities are approximately equal at $-190°C$. Another crystal (NaI($10^{-6}$Tl), was weakly luminescent at room temperature, but emitted both bands stronglyy at the lowest temperature, with a total intensity approximately equal to the intensity of NaI($10^{-3}$Tl). Differences in the spectra which result from the different sources of excitation will be discussed.

The results, when combined with those of other investigators, have suggested a model for the scintillation phenomenon in NaI which is at least internally consistent. The model in turn leads to a plausible explanation of the high relative luminescent efficiency of NaI(Tl) and unactivated NaI as compared with some others of the alkali halide family.

The characteristic decay time of the scintillation light from unactivated NaI was found to vary between $10^{-8}$ seconds at $20°C$ and $3 \times 10^{-8}$ seconds at $-190°C$. The pulse height divided by particle energy at $-190°C$ is approximately equal to that of NaI($10^{-3}$Tl) at $+20°C$. The high luminescent efficiency, rapid decay time, and relatively high density of unactivated NaI at $-190°C$ make it a unique scintillation material which may prove to be useful in experimental and practical applications.

**Scintillating Solutions Containing Heavy Elements**—H. P. Kallmann, M. Furst, and F. H. Brown (p. 51)

The use of heavy metals in scintillators (excited by high energy) is limited by two factors: 1) such materials are very often not sufficiently soluble in solvents exhibiting considerable energy transfer, and 2) they generally quench the excitation energy of the solution and thereby decrease the fluorescent output.

The latter effect was found to be due mostly to a quenching of the excited molecules of the solvent which are responsible for transferring energy, and to a lesser extent to the quenching of the fluorescent molecule, although this process also occurs. To circumvent the first difficulty one uses solvents with better solubilities for the metal compounds, but the choice is limited because of poor energy transfer characteristics of many of them. One of the best solvents found up to now is $p$-dioxane, which exhibits considerable energy transfer and shows better solubility properties than most of the other energy transferring solvents. Another means of overcoming difficulties of solubility is to use organic solvents favorable from the standpoint of solubility, but not so favorable from that of energy transfer. One can effectively induce considerable energy transfer in many of these solutions and thus produce high energy excited fluorescence by adding a relatively small amount, of the order of 10 per cent, of an effective liquid solvent (one which shows energy transfer). In many instances such an amount of effective solvent takes over a considerable amount of the excitation energy originally absorbed in the less effective solvent and transfers it to the fluorescent molecule before quenching occurs. This method of adding effective liquid solvents to ineffective ones has been very successful in making even almost completely ineffective solvents usable. One can in many cases find even more favorable results by using, instead of the effective liquids, a suitable material, solid at room temperature, which

can be dissolved in fairly large amounts in the ineffective solvent. It has been observed that some solids are very potent in making ineffective solvents useful for high energy fluorescence, and that these are even more effective than the best energy transferring liquid solvents known so far. Such solids are, e.g., naphthalene and biphenyl.

The second detrimental factor, the quenching of the excitation energy, also can be overcome to a noticeable extent by using as the energy transferring material a molecule which is less quenched than a normal effective solvent molecule such as toluene. It has been found that just those solids which produce energy transfer in ineffective solvents are also those which are less quenched. Thus, using these materials it is possible to make solutions which contain considerable amounts of quenching materials, e.g., metals, and are nevertheless fluorescent under high energy excitation.

There is a last step which can be taken to prepare solutions moderately fluorescent under high energy excitation while containing still greater concentrations of various quenching molecules. One could try to use water as solvent, since many salts are soluble in it; but water has no energy transfer capabilities. This can be circumvented by means of a method similar to that described above; $p$-dioxane is used as the organic fraction of the solvent since it is miscible with water. If about 20 per cent of water is added to a fluorescent dioxane solution, its fluorescence goes down to about 10 per cent of the value of a normal liquid scintillator under high energy excitation. If one of the above mentioned solids is added to such a water-dioxane system, an increase in fluorescent to almost half the value of the scintillator in the most effective solvent is observed; at the same time the relative water content is not changed very much. The fluorescence is in some cases as high as with the same fluorescent solute in a dioxane solution with no other materials present. Such an amount of water is large enough to dissolve various metal salts, and successful experiments with inorganic materials like cupric- and silver-nitrate and sodium iodide dissolved in a $p$-dioxane-water solution have been performed.

**Intrinsic Scintillator Resolution**—G. G. Kelley, P. R. Bell, R. C. Davis, and N. H. Lazar (p. 57)

Measurements of spectrometer resolution using an electron gun and phosphor type of light source yield values considerably better than those obtained from scintillations in a phosphor. Line widths of about 3.4 per cent have been obtained at a pulse height equivalent to 661 kev in NaI. An investigation is being made to determine the causes of this difference. A number of possibilities have been eliminated. Flasher light pulses may be collimated or diffused at the photo surface with no difference in result. They may be mixed with random scintillation pulses giving simultaneously the relatively poor resolution of the scintillator and the good resolution of the flasher. Optical filters do not affect the result when the flasher is used, but do give a relative improvement with a scintillator. It is found that as the light is attenuated in any way from a scintillation crystal, its relative resolution is improved. The improvement is compatible with the concept of an intrinsic crystal resolution. In other words, the difference of the squares of crystal width and flasher width remains constant with varying attenuation. It remains constant also for a given crystal with photomultipliers of different types and resolutions. The intrinsic width indicated for a good crystal is about 6 per cent. Results of experiments attempting to locate the cause of this effect are reported.

**Some Aspects of Gas Scintillation Counters** —J. A. Northrop and R. A. Nobles (p. 59)

A brief study has been made of the dependency of pulse heights obtained from alpha

particles in gas scintillation counters on the gas and wavelength shifter used. Constant and reproducible pulse heights were obtained either by chilling the scintillation cylinder to inhibit evolution of contaminants or by chemically removing them in a uranium furnace. Relative pulse heights have been obtained for various wavelength shifters, gases and gas mixtures, and reflectors. When the waveshifter was vacuum evaporated directly onto the tube face, the optimum thickness was found to be $\sim 20$ $\mu g/cm^2$, with tetraphenylbutadiene giving $\sim \frac{1}{3}$ more light than quaterphenyl. Using a 2-inch long aluminum cylinder with the alpha particle source at one end and the phototube at the other, the pulse height was further increased by a factor of 5 by smoking the cylinder with MgO which was subsequently coated with waveshifter. The relative pulse heights for different gases in such a system were: Xe:Kr:A:Ne:He = 32:16:5:1:10. A NaI(Tl) crystal in similar reflector system gave a pulse height of 72. The rise time of the pulses in a xenon counter using an RCA-6342 tube coated with quaterphenyl was found to be on the order of 3.5 m$\mu$ sec. In such a system the pulse height was found to be a linear function of energy for protons, deuterons, and helium ions and to be independent of the type of particle producing the scintillations.

**Relative Scintillation Intensity of Some Cerenkov Counter Media**—R. Madey and L. Leipuner (p. 61)

Selected transparent media were bombarded with the 40 mev alpha particle beam from the 60-inch cyclotron. The ratio of the photomultiplier output current to the alpha particle beam current was taken as a measure of the relative scintillation intensity. The output current from an RCA type 6342 photomultiplier was read on a 50 microampere (full scale) meter. The alpha particle beam current was monitored by counting alpha particles elastically scattered from a 0.005-inch gold foil.

Preliminary results on the light output from some media relative to that for distilled water are tabulated in a table.

Background appeared to give negligible contribution to the scintillation intensity. The background measurement was made by stopping the alpha beam with a 1/16-inch aluminum plate inserted ahead of the substance studied.

**Temperature Effects in Gas-Free Liquid Scintillators**—H. H. Seliger and C. A. Ziegler (p. 62)

The effects of reduction of temperature on the response of Dumont 6292 and RCA 5819 multiplier phototubes and on the light outputs of the gas-free liquid scintillators DPO, $\alpha$NPO, and PBD, dissolved in xylene, have been measured from room temperature down to $-35°C$. Both 6292 and 5819 phototubes exhibit a small linear decrease in response with decreasing temperature. The light outputs of the liquid scintillators when excited by alpha particles increase markedly with decreasing temperature. The measurements indicate that under proper conditions it may be possible to obtain a liquid scintillator more efficient for alpha particles than crystal anthracene. Previous data using gamma-ray excitation carried out over a limited temperature range are in agreement with the alpha particle results.

**High-Energy Gamma Spectroscopy**—C. E. Swartz (p. 65)

Gamma ray energies from 50 mev to 2 bev have been measured with a total absorption method. The gamma rays produce electron showers in a transparent, high Z material. A photomultiplier views the visible light produced, due either to Cerenkov radiation or scintillation. Limits on the energy resolution of this device are the loss of part of the shower from the sides and end, varying efficiency of light collection from different parts of the material, and paucity of Cerenkov photons

when that technique is used. Factors involved in choosing the material and shape of the shower producer will be discussed. Because of the small size and high efficiency, this type of spectrometer is very useful for experiments with high energy accelerators. Descriptions and characteristics of several versions of this instrument are given.

**Low-Energy Gamma Scintillation Spectrometry**—C. J. Borkowski (p. 71)

Techniques are reviewed which are used to: obtain good resolution with scintillation spectrometers; minimize pulses which fall outside of photo peaks; measure absolute gamma efficiencies of NaI crystals as a function of energy; and measure $K/L$ capture ratios and gamma conversion coefficients.

Some factors affecting energy resolution including associated electronic circuits are discussed.

**Neutron Scintillation Counters**—C. O. Muehlhause (p. 77)

The various types of neutron scintillation detectors are examined with respect to efficiency, background sensitivity, pulse height and time response. Actual performance of certain counters under different source conditions are considered. These are given, and include a few imaginary cases.

**Response of "Total Absorption" Spectrometers to Gamma Rays**—R. C. Davis, P. R. Bell, G. G. Kelley, and N. H. Lazar (p. 82)

A single crystal of sodium iodide $9\frac{1}{4}$ inches in diameter and 9 inches high with the top bevelled at 45° was recently obtained from Harshaw Chemical Company. A half-inch hole was drilled in the bevelled end to the center of the crystal. Three 5-inch diameter Dumont 6364's were optically connected directly to the flat end of the crystal. Sources were placed inside the crystal and $\gamma$ spectra obtained, taking advantage of full geometry.

A $4\frac{3}{4}$ inches diameter sodium iodide crystal $5\frac{1}{2}$ inches high with the top bevelled at 45° has a $\frac{1}{4}$ inch hole drilled to the center. The crystal is optically connected to a Dumont 6364. The crystal is placed inside a 28-inch tank of liquid scintillator which uses four Dumont 6364's to detect pulses in the liquid. Samples are placed inside the crystal and the pulses from the crystal and liquid scintillator are placed in anticoincidence to reduce the Compton portion of the $\gamma$ spectra.

Typical spectra with these counters are shown along with a comparison of data taken with these and smaller crystals.

**Transmission Characteristics of Light Pipers**—C. C. Harris and P. R. Bell (p. 87)

The transmission characteristics of light pipers used in scintillation counting are very important, especially in situations in which the light piper must be long and thin, as in medical probing counters. This paper discusses several of the important factors involved in the transmission of light through pipers. The superiority of specular reflectors as compared with diffuse reflectors, the superiority of dry reflectors as compared with optically coupled reflectors, and the importance of surface polish are discussed. The use of total-internal-reflection couplers between scintillation crystal and light piper is described. Relative pulse heights obtained with certain crystal-piper-reflector combinations are given, and transmissions as functions of piper length and configuration are shown.

**An Analysis of the Background Radiation Detected by NaI Crystals**—C. E. Miller, L. D. Marinelli, R. E. Rowland, and J. E. Rose (p. 90)

The 0.07-2 5 mev scintillation pulse height spectra of several NaI crystals, heavily shielded by Fe and Hg, have been analyzed in coincidence and anticoincidence with cosmic ray counters surrounding the crystals. The residual spectra have been analyzed further and the following sources have been identified: 1) $Ra^{226}$

in Al casing, 2) $K^{40}$ in phototube, crystal, and canning window, and 3) low pulse height spectrum due to the ionizing components of cosmic rays.

The above sources still leave 0.05 c/m/g of NaI unexplained. Activities in stainless steel, quartz, and MO reflectors are too low to account for the bulk activity. Moreover, activation analysis of crystal fragments has eliminated the presence of $Rb^{87}$ as a natural contaminant, and decay studies in neutron free caves preclude the presence of $I^{126}$ from cosmic ray neutron activation.

A possible source of contamination, consistent with the none too well defined shape of the residual spectrum, is $Cs^{137}$, a fission product, in weight concentrations of about $5 \times 10^{-16}$. Studies aimed at proving or disproving its existence are currently being pursued.

**In Vivo Gamma Measurements at Very Low Levels with $4\pi$ Liquid Scintillation Detectors**—E. C. Anderson (p. 96)

A large $4\pi$ liquid scintillation counter for the rapid measurement of total internal gamma contamination of the human body at concentrations 1000 times less than the maximum permissible levels is described. A counting time of only 100 seconds per subject is required to obtain a sensitivity of $5 \times 10^{-10}$ curies for 1.5 mev gamma rays. Because of the high counting efficiency and the great stability of operation, the magnitude of counter background is of comparatively little importance so that elaborate shielding precautions are not required. Energy resolution is sufficient to distinguish between gamma rays of markedly different energy, but not adequate for the identification of individual nuclides.

The operating characteristics of the counter are discussed including the spectrum and stability of background, energy resolution, and the spectra and limits of sensitivity for nuclides of interest in connection with nuclear power stations and uranium mining operations.

**The Physics of Solid-State Light Amplifiers**—D. A. Cusano (p. 102)

An essential component of present light amplifying screens is a layer of inorganic, luminescent material. The ability of this layer to derive energy from an electric field to which it is subjected and convert this energy into visible light (*i.e.*, electroluminescence) is fundamental to, but not alone sufficient for the realization of solid-state light amplification. What must be included is some means whereby the luminescent response to the field can be strongly controlled by incident radiation. In one case this has been obtained through the discovery of a phosphor film with the necessary properties, and in another by the utilization of a photoconducting material in contact with an electro-luminescent layer. In the latter case the photoconductor is the radiation-sensitive element, which serves to modify the potential across the luminescent component.

The main characteristics of these light amplifying screens will be described, with particular emphasis on the interpretation of the phenomenon observed with the single phosphor film. The behavior, as it is relevant to their possible use in radiation detection will be discussed.

**Photomultiplier Counters in High-Energy Physics Experiments**—B. J. Moyer (p. 107)

Techniques in the use of photomultipliers in high-energy nuclear physics are presented by the description of experiments that have been prominent recently. The applications involve the use of photomultipliers with attendant amplification at the limit of present possibilities of time resolution. Such applications include measurements of time of flight, measurement of lifetime of heavy unstable particles, and the elimination of background effects in the pulsed radiation fields of high-energy accelerators. The need for the development of large fast-rising signals from events of small light output,

such as Cerenkov light pulses, is made apparent, and a discussion of the performance desired from electron multipliers in view of contemplated experiments is presented.

**High Output Multiplier Tubes with Accelerator Grids**—J. S. Allen and L. R. Megill (p. 112)

A multiplier tube capable of producing output currents greater than one ampere has been constructed. In order to produce high electrical fields at the surfaces of the dynodes a design utilizing an accelerating grid near the surface of each dynode was chosen. A space charge limited current of 1.2 to 1.3 amperes was obtained using 400 volts between each dynode and accelerating grid.

**Improved Time Response in Scintillation Counting**—Q. A. Kerns (p. 114)

Antiproton observations at the Bevatron are typical of scintillation counting experiments that would benefit by improved time response. The antiproton yield increases as the flight path of the particle is shortened and the corresponding solid angle increased. From the electronics standpoint, the problem of pulse rise time prevents an arbitrary reduction of the flight path. The first limitation appears in the scintillator-photomultiplier combination, the second in the interconnecting cable, and the last in the indicating circuitry, *i.e.*, coincidence devices, scalers, oscilloscopes, amplifiers, etc. In Part I of this paper, transit time data for the RCA type 6810 are presented. These suggest a method for improving tube performance. Figures are given in Part II for pulse distortion in typical coaxial cables. Performance data are given for a low-distortion 125-ohm transmission line. The line, using styrofoam dielectric, is easily made and useful over several hundred feet in the $10^{-10}$ second range.

**Photomultiplier Transit-Time Measurements**—R. V. Smith (p. 120)

Electron transit time has been measured as a function of over-all voltage, and of the position illuminated on the photocathode, for the following photomultipliers: 931-A, 5819, 6342, 6199, 6292. In addition, a few preliminary measurements have been made for the new RCA 6810 and a new Westinghouse transmission-type tube. Light pulses from a spark gap simultaneously illuminated a 935 phototube and, through attenuators, small areas on the various multiplier photocathodes. The two output pulses were displayed on the same cathode-ray tube sweep, recorded photographically, and their time difference measured by reference to a superposed 700-mc sine wave. As a function of voltage, the 5819, 6342, and 6199 transit times range between 21 and 38 $m\mu sec$, and fall on a single curve as expected from their similar geometries. The 931-A shows values between 11 and 18 $m\mu sec$, the 6292, 56–73 $m\mu sec$, and the Westinghouse transmission type of the order of 3 $m\mu sec$. As a small illuminated spot is moved across the photocathode, transit-time variations range from 3 $m\mu sec$ for the 5819 to 8 $m\mu sec$ for the 6292.

**Recent Developments in the Scintillation Counter Field**—G. A. Morton (p. 122)

Rarely does a device having the conceptual simplicity of the scintillation counter gain the wide recognition accorded this detector. It has become an essential in most phases of nuclear research, and its usefulness extends beyond this into the fields of radiochemistry, biology, tracer technology, radiation and contamination protection, and prospecting for oil and fission fuel. This importance rests upon its extremely high sensitivity, its superior time resolution capabilities, and its ability to measure the energy of the incident nuclear particles. Furthermore, the instrument is rugged, reliable, and flexible.

The scintillation counter consists of a clear phosphor scintillation, a multiplier phototube, and the required presentation circuitry. Incident nuclear radiation excites light from the phosphor. This light releases photoelectrons in

# 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.

The number in heavy type at the upper left of each Abstract is its Universal Decimal Classification number and is not to be confused with the Decimal Classification used by the United States National Bureau of Standards. 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

**534.2–14:534.5**      **1**
Distortion and Interaction of Acoustic Waves with Finite Amplitude in a Viscous Medium—G. D. Mikhailov. (*C.R. Acad. Sci. U.R.S.S.*, vol. 109, pp. 68–71; July 1, 1956. In Russian.) A theoretical paper. For a note on an experimental confirmation of the results, see *Zh. Eksp. Teor. Fiz.*, vol. 30, p. 1142; June 1956.

**534.232–14–8**      **2**
The Conversion Efficiency of a Piezoelectric Quartz Crystal: Relation between Input Electrical Power and Frequency—S. Parthasarathy and V. Narasimhan. (*Z. Phys.*, vol. 145, pp. 368–372; May, 1956. In English.) Experiments indicate that the conversion efficiency depends on frequency and on the nature of the liquid surrounding the crystal, but not on the input power. In a separate paper (*ibid.* pp. 373–376) it is reported that the conversion efficiency is higher for the fundamental than for overtones.

**534.26+[538.566:535.42**      **3**
Asymptotic Solution of some Diffraction Problems—Keller, Lewis, and Seckler. (See 90.)

**534.32:534.75**      **4**
Pitch of Inharmonic Signals—E. de Boer. (*Nature Lond.*, vol. 178, pp. 535–536; September 8, 1956.) Brief report of a study of the dependence of the subjectively perceived pitch of a complex tone on the harmonic relation between the af carrier and the modulating signal, the latter comprising components whose frequencies are in arithmetic progression.

**534.6**      **5**
Comparison of Artificial Ears—U. Degano. (*Piccole Note Ist. Super. Poste e Telecomunicazioni*, vol. 5, pp. 195–202; March/April, 1956.) An objective substitution method is described, with test results for the Italian, Swiss, and C.C.I.F. standards on three different Telephone receivers.

**534.61–8**      **6**
Ultrasonic Intensity Meter—C. A. Wiederhielm. (*Rev. Sci. Instrum.*, vol. 27, pp. 540–541; July, 1956.) A calorimetric method is described which does not involve disturbance of the ultrasonic field. Intensity measurements can be obtained for areas as small as 8 mm².

**534.851(083.74)**      **7**
Disk Recording Characteristics—J. D. Smith. (*Wireless World*, vol. 62, pp. 526–528; November, 1956.) British Standard 1928:1955 is discussed and a complete circuit diagram of a suitable replay equalizer is given.

**621.395.61**      **8**
The Design of a High-Quality Commentators' Microphone Insensitive to Ambient Noise—H. D. Harwood. (*B.B.C. Eng. Div. Monographs*, No. 7, pp. 1–19; June, 1956.) Description of the Type-L.2 lip microphone, which was designed in 1951. Special steps are taken to secure correct balance between the sound emanating from the nose and that from the mouth, and to allow for the spectral distribution of the energy at short range. The transmission of low-frequency pulses produced by high-velocity air streams accompanying explosive consonants and of wind noise is avoided. The microphone is highly insensitive to interfering alternating magnetic fields.

**621.395.623.7**      **9**
The Diffaxial Speaker—A. B. Cohen. (*Audio*, vol. 40, pp. 20–23, 58; June, 1956.) A three-range loudspeaker is described consisting of a dual diaphragm for reproduction of the lower frequencies, the two sections being coupled so as to produce a mechanical crossover, together with a coaxially mounted "tweeter."

**621.395.623.7:676.4**      **10**
On a Model of Paper with the Dynamic Dynamic Properties [suitable for loudspeaker cones]—Nimura and Kido. (See 214.)

## ANTENNAS AND TRANSMISSION LINES

**621.315.2**      **11**
Minimum-Resistance High-Frequency Cable —H. J. Hoehnke. (*NachrTech.*, vol. 6, pp. 252–257; June, 1956.) The variation of the hf resistance of stranded tubular cable conductors with the diameter and number of the strands is investigated. The results are shown graphically; from the positions of the minima an estimate can be made of the optimum distribution of the conductor metal.

**621.315.212:621.372.2**      **12**
Distortionless Coaxial Cables—G. Mattson. (*Ericsson Tech.*, vol. 12, pp. 25–29; 1956.) The transmission time and attenuation of coaxial cables for television and multichannel telephony are equalized by using a thin-walled tube as inner conductor, in conjunction with series loading capacitors. Suitable values of capacitance and loading distances are calculated. The characteristic impedance of the cable is equivalent to the series combination of a resistance and a capacitance for all frequencies.

**621.372:538.221:621.318.134**      **13**
Ferrites—(See 203.)

**621.372.2**      **14**
Surface and Space Waves on the Surface-Wave Transmission Line—H. Uchida and S. Nishida. (*Sci. Rep. Res. Inst. Tohoku Univ.*, Ser. B, vol. 6, pp. 217–227; 1955.) The surface and space waves due to a magnetic current encircling a perfectly conducting wire coated with a thin loss-free dielectric film are evaluated by the saddle-point integration method. The surface wave exists only near the line; further out it is cancelled by a part of the space wave.

**621.372.2**      **15**
The Shunt Reactive Element on the Surface-Wave Transmission Line—H. Uchida, S. Nishida, H. Uda, and H. Nagasawa. (*Sci. Rep. Res. Inst. Tohoku Univ.*, Ser. B, vol. 6, pp. 229–238; 1955.) A thin metal annulus or rod positioned symmetrically on a transmission line acts as a shunt capacitance; the analysis is checked by experiment.

**621.372.2+621.372.8]:537.226**      **16**
Propagation of Microwaves along a Solid Conductor Embedded in Three Coaxial Dielectrics—S. K. Chatterjee and R. Chatterjee. (*J. Indian Inst. Sci., Section B*, vol. 38, pp. 157–171; July, 1956.) A boundary-value treatment is presented. The field components in the three dielectric media are derived in terms of the axial power flow. The characteristic equations for the $E_0$ and $H_0$ modes are established and the special cases of 1) a solid conductor embedded in free space, 2) a dielectric rod, and 3) a dielectric tube, are investigated.

**621.372.2:621.318.134**      **17**
A New Ferrite Isolator—B. N. Enander. (PROC. IRE, vol. 44, pp. 1421–1430; October, 1956.) An isolator for a helical transmission line consists basically of ferrite rings or cylinders surrounding the line; these ferrite elements are premagnetized circumferentially, hence no applied magnetic field is required. A backward /forward loss ratio of 25:1 is attainable at 6.5 kmc, with a ratio >15:1 over a 25 per cent frequency band. The effect of an additional axial magnetic field, such as the focusing field of a traveling-wave tube, is discussed. With suitable modifications the arrangement can also be used as a switch or a modulator.

**621.372.8:621.318.134**      **18**
Ferrite Directional Couplers—A. D. Berk and E. Strumwasser. (PROC. IRE, vol. 44, pp. 1439–1445; October, 1956.) Design theory and

curves and tables of limiting values are presented, facilitating rapid calculation of these parameters for any frequency. Power gain is discussed, and a formula is derived for its frequency variation in common-emitter circuits.

**621.375.4.024:621.314.7**     66

Transistor D.C. Amplifier—D. M. Neale and F. Oakes. (*Wireless World*, vol. 62, pp. 529–532; November, 1956.) The amplifier described provides a current gain of 1000 and power gain of 45 db; the equivalent noise fluctuations are about 1 m$\mu$A at the input; input impedance is 5–10 k$\Omega$. The push-pull grounded-emitter input stage is followed by a push-pull grounded-collector stage; a fifth transistor in a negative-feedback loop limits the effects of a collector leakage-current variations and by stabilizing the first-stage collector voltage restricts the effect of transistor noise. A complete circuit diagram of an amplifier using five Type-OC71 transistors is given.

**621.375.4.029.5:621.314.7**     67

Transistor RF Amplifiers—D. D. Jones. (*Wireless World*, vol. 62, pp. 494–496, 544–546; October/November, 1956.) The use of the Type-GET4 *p-n-p* transistor in a 465-kc IF amplifier is discussed, and practical circuit with neutralization and agc is described. The gain of a two-stage amplifier is about 36 db.

**621.372.413 + 621.372.8**     68

Elektromagnetische Wellenleiter [Book Review]—Goubau. (See 32.)

## GENERAL PHYSICS

**535.22:538.566.029.6**     69

Microwave Determinations of the Velocity of "Light"—K. D. Froome. (*J. Brit. IRE*, vol. 16, pp. 497–513; September, 1956.) "A description is given of methods utilizing microwaves for the purpose of measuring the free-space vacuum velocity of electromagnetic waves. The cavity resonator and the microwave interferometer methods are discussed in detail; brief mention is made of the molecular band spectrum method. The values obtained from these methods are compared with optical values."

**535.23**     70

The Maximum of a Distribution or Spectrum Function—W. van der Bijl. (*Nature, Lond.*, vol. 178, p. 691; September 29, 1956.) The general applicability of Bracewell's arguments (398 of 1955) is emphasized and illustrated by examples.

**535.6**     71

Use of an Exponential Function for establishing a Uniform Chromaticity Scale—R. Taguti and M. Sato. (*C.R. Acad. Sci., Paris*, vol. 243, pp. 654–656; August 13, 1956.)

**537/538**     72

What is Permeability?—E. G. Cullwick. (*J. IEE*, vol. 2, pp. 416–417; July, 1956.) Various definitions of permeability are compared. The introduction of the concepts permeability and permittivity of free space is considered to be misleading. Relations between fundamental electric and magnetic constants are presented in a form suitable for use in the M.K.S. system.

**537.2**     73

Charged Right-Circular Cylinder—W. R. Smythe. (*J. Appl. Phys.*, vol. 27, pp. 917–920; August, 1956.) "A new method permits the calculation of the electric field surrounding a charged conducting surface of revolution without the use of orthogonal functions. Detailed formulas show how to find the charge density on a right circular cylinder with any desired precision."

**537.5**     74

Elementary Electron-Recombination Processes in Decaying Discharge Plasma—E. Schulz-DuBois. (*Z. Angew. Phys.*, vol. 8, pp.

267–269; June, 1956.) Volume recombination, ambipolar diffusion, and the formation of negative ions are considered.

**537.5:538.56.029.6**     75

Processes in the Gas Discharge in the Spectrochemical Identification of Nonmetals by Ultra-short-Wave Excitation using Gatterer and Frodl's Method—D. von Bezold. (*Z. Angew. Phys.*, vol. 8, pp. 269–281; June, 1956.)

**537.52**     76

Oscillating Glow-Discharge Plasma—A. B. Stewart. (*J. Appl. Phys.*, vol. 27, pp. 911–916; August, 1956.) Probe measurements are reported of the potential, electron temperature and electron density in the positive column of a dc argon glow discharge with moving striations accompanied by af oscillations.

**537.52**     77

High-Frequency Gas-Discharge Breakdown in Neon/Argon Mixtures—H. J. Oskam. (*J. Appl. Phys.*, vol. 27, pp. 848–853; August, 1956.) Report of measurements at 9.5 kmc of the variation of the breakdown field with the percentage of argon in the mixture.

**537.525**     78

Pre-breakdown Current and Vacuum Breakdown—W. J. R. Calvert. (*Proc. Phys. Soc.*, vol. 69, pp. 651–660; June 1, 1956.) In investigating discharges between un-outgassed electrodes in strong electric fields, at pressures of the order of $10^{-5}$ mm Hg, an initial pre-breakdown stage involving the cold emission of electrons was noted at field strengths of about $10^5$ V/em, followed by two further stages involving current pulses before complete breakdown. This characteristic development is discussed in terms of the particle-exchange theory [3090 of 1947 (Trump and Van de Graaff)].

**537.533:537.534.8**     79

Secondary Emission under the Action of Positive Potassium Ions with Different Charges—I. P. Flaks. (*Zh. Tekh. Fiz.*, vol. 25, pp. 2463–2466; December, 1955.) A report is presented on an experimental investigation into the secondary emission of electrons and ions from a platinum target under the action of $K^+$, $K^{2+}$, and $K^{3+}$ ions accelerated by voltages from 1000 to 20,000 v. The secondary-emission coefficients increase with increasing charge of primary ions. The curves of secondary electron emission for various charges of primary ions are practically parallel. The increase in secondary electron emission under the action of multicharge ions can be regarded as a result of a combination of field and kinetic effects.

**537.533:621.38.032.21**     80

Symposium on Cathode Electronics—V. M. Gavrilyuk. (*Uspekhi Fiz. Nauk*, vol. 59, pp. 363–374; June, 1956.) Report of a symposium held at Kiev in November, 1955. About 40 papers were presented on thermionic, field, secondary, and photoelectric emission and the operation of cathodes under ion bombardment. See also (I. M. Dykman) *Radiotekhnika i Elektronika*, vol. 1, pp. 393–403; March, 1956.

**537.533.8**     81

The Possible Occurrence of Exciton-Enhanced Secondary Emission—A. J. Dekker. (*Physica*, vol. 22, pp. 361–366; May, 1956.) "A theoretical study is made of the possible occurrence of exciton-enhanced secondary emission in insulators containing donor levels. It is assumed that the primary beam produces, besides secondary electrons, excitons; the latter may ionize donors and thereby enhance the secondary yield. The discussion is limited to high primary energies, so that the range of the primaries is large compared to the diffusion length of the excitons. It is concluded that the enhancement may be considerable if the destruction of excitons at the surface is not too large."

**537.56:538.56**     82

Electromagnetic Waves in an Ionized Gas—P. Rosen. (*Phys. Rev.*, vol. 103, pp. 390–394;

July 15, 1956.) The influence of electric field distribution on the complex conductivity of a low-pressure ionized gas is investigated theoretically. In the case of a standing wave the electron diffusion constant varies over the field; in the case of a traveling wave in an infinite ionized medium, the conductivity is independent of position.

**538.221:537.56:538.6**     83

Diffusion of Charged Particles across a Magnetic Field—C. L. Longmire and M. N. Rosenbluth. (*Phys. Rev.*, vol. 103, pp. 507–510; August 1, 1956.) Analysis is presented based on a random-walk treatment. The relative importance of ion/electron and ion/ion collisions depends on the plasma density.

**538.221:621.318.134:621.372**     84

Ferrites—(See 203.)

**538.3**     85

The Magnetic Fields of Isolated Moving Charges—H. L. Armstrong. (*Elect. Eng.*, vol. 75, pp. 554–555; June, 1956.)

**538.3**     86

An Expansion Theorem for Electromagnetic Fields—C. H. Wilcox. (*Commun. Pure Appl. Math.*, vol. 9, pp. 115–134; May, 1956.)

**538.561:621.372.413**     87

Radiation from a Charged Particle moving through Coupled Resonators—A. I. Akhiezer, G. Ya. Lyubarski, and Ya. B. Fainberg. (*Zh. Tekh. Fiz.*, vol. 25, pp. 2526–2534; December, 1955.) Cherenkov radiation from a charge moving uniformly in a periodic structure is considered. This radiation can be interpreted as a resonance between the natural vibrations of the structure and a force associated with the moving particle. A relation is derived determining the spectrum of the radiated frequencies. The Doppler effect in the case of an oscillator moving in a periodic structure is also considered.

**538.561:621.372.413**     88

Interaction of Coupled Electromagnetic Resonators with a Beam of Charged Particles—A. I. Akhiezer and Ya. B. Fainberg. (*Zh. Tekh. Fiz.*, vol. 25, pp. 2516–2525; December, 1955.) If a beam of charged particles of velocity above a critical value passes through a chain of coupled resonators, the fluctuations of charge density and beam velocity are propagated in the form of waves with increasing amplitude. A relation is established between these waves and the Cherenkov radiation for the case of particles moving in a periodic structure or a dielectric.

**538.566:535.4**     89

Electromagnetic Cross-Section of a Small Circular Disk with Unidirectional Conductivity—G. Toraldo di Francia. (*Nuovo Cim.*, vol. 3, pp. 1276–1284; June 1, 1956. In English.) An investigation is made of scattering by an infinitely thin screen with infinite conductivity in one direction and zero conductivity in the perpendicular direction. A physical system closely approximating these conditions is constituted by a grating of closely spaced parallel wires. The analysis is similar to that presented by Bethe (706 of 1945), with modifications to remove certain inaccuracies.

**538.566:535.42] + 534.26**     90

Asymptotic Solution of some Diffraction Problems—J. B. Keller, R. M. Lewis, and B. D. Seckler. (*Commun. Pure Appl. Math.*, vol. 9, pp. 207–265; May, 1956.) An asymptotic expansion with respect to frequency is developed for a periodic solution of wave problems; the numerous cases treated individually include diffraction by wedges, cylinders, cones, paraboloids, and spheres.

**538.566:535.42] + 534.26**     91

Diffraction by a Wide Slit—S. N. Karp and A. Russek. (*J. Appl. Phys.*, vol. 27, pp. 886–894; August, 1956.) Approximate expressions

are derived for the diffracted field resulting from normal incidence of a plane scalar wave on a conducting plane with an infinite slit, and numerical values are calculated for slits of width $0.96 - 2.5\lambda$. The method takes account of the interaction between the edges, and the accuracy of the solution increases with slit width.

**538.566:535.42** 92
The Problem of the Diffraction of Electromagnetic Waves at a Circular Hole in a Plane Screen—N. I. Akhiezer and A. N. Akhiezer. (*C.R. Acad. Sci. U.R.S.S.*, vol. 109, pp. 53–56; July 1, 1956. In Russian.) An application of results obtained by N. I. Akhiezer in a paper on "Some Paired Integral Equations" (*ibid.*, vol. 98, pp. 333–336; 1954.) The screen is assumed to be ideally conducting.

**538.566:535.42** 93
Diffraction of 3.2-cm Electromagnetic Waves by Dielectric Rods: Part 1—Lucite and Tenite 1-in.-Diameter Cylinders—M. K. Subbarao and A. B. McLay. (*Canad. J. Phys.*, vol. 34, pp. 546–554; June, 1956.) Experiments similar to those of Wiles and McLay (3538 of 1954) are reported. The results obtained with the two loss-free dielectric rods are compared with those for conductors and for lossier dielectrics; the differences are explained in terms of transmission and/or reflection effects. Measurements on a lucite $1\frac{1}{2}$-in.-diameter cylinder and a semicylinder are discussed in Part 2 [*ibid.*, pp. 555–562 (McLay and Subbarao)].

**538.566:535.42** 94
Diffraction of a Finite Beam of Electromagnetic Waves by a Cylindrical Obstacle—E. L. Burshtein and L. S. Solov'ev. (*C.R. Acad. Sci. U.R.S.S.*, vol. 109, pp. 473–476; July 21, 1956. In Russian.) Analysis is presented, starting with the solution of the two-dimensional problem of the diffraction of a plane wave incident normally on a cylindrical obstacle and using the method of superposition of plane waves for representing the field.

**538.566:535.42** 95
Diffraction of a Plane Electromagnetic Wave by a Conducting Cylinder—A. S. Goryainov. (*C.R. Acad. Sci. U.R.S.S.*, vol. 109, pp. 477–480; July 21, 1956. In Russian.) Analysis is presented for the case of a cylinder of radius large compared with $\lambda$.

**538.566:535.43** 96
Scattering of Electromagnetic Waves by Coaxial Cylinders—A. W. Adey. (*Canad. J. Phys.*, vol. 34, pp. 510–520; May, 1956.) "A scattering system comprising two coaxial, dielectric cylinders has been studied theoretically and experimentally. Calculations have been made of the forward and back scattered fields for several combinations of inner and outer radii. It has been found that, by covering a metal cylinder with a coaxial dielectric shield, it is possible to eliminate to some extent the deep near-field shadow. Experimental results obtained at a wavelength of 3.275 cm using a parallel-plate transmission line are in good agreement with calculations."

**538.566:535.43:537.226** 97
Scattering of Microwaves by Long Dielectric Cylinders—A. W. Adey. (*Wireless Eng.*, vol. 33, pp. 259–264; November, 1956.) Extension of previous work dealing with metal cylinders (400 of 1956). In the present case, because the field penetrates into the dielectric, resonance can occur. Damping effects due to the dielectric losses are exhibited. Experimental results are given for square and rectangular polystyrene rods.

**538.569.4:537.226** 98
Transmission of Electromagnetic Waves through Inhomogeneous Layers: Part 2—Absorption of Electromagnetic Waves—H. G. Haddenhorst. (*Z. Angew. Phys.*, vol. 8, pp. 264–267; June, 1956.) An investigation of tech-

niques for obtaining surfaces with low-reflection factors over a wide-frequency band by means of absorbing layers of least possible thickness, using wedge or pyramidal projections as described in connection with the previous work on nonabsorbing layers (1035 of 1956). The desired loss can be produced by including graphite in a dielectric such as paraffin. Diffraction effects are avoided by suitable geometric design. Measurements made on absorbers in waveguides are reported. At 8.75 cm $\lambda$, with an angle of incidence 48.8° and the electric vector perpendicular to the plane of incidence, a layer with pyramidal projections can be produced having a reflection factor of 10 per cent with an over-all layer thickness of 7–8 cm.

**538.569.4:539.152.2** 99
Nuclear Magnetic Resonance—J. G. Powles. (*Sci. Progr.*, vol. 44, pp. 449–471; July, 1956.) The basic ideas are presented.

**538.569.4:539.1** 100
Molecular Beams [Book Review]—N. F. Ramsey. Publishers: Clarendon Press, Oxford and Oxford University Press, London, 1956, 466 pp., 75s. (*Nature, Lond.*, vol. 178, pp. 608–609; September 22, 1956.) A monograph presenting descriptions of the techniques of many types of experiment using molecular beams, together with the results obtained and their significance.

## GEOPHYSICAL AND EXTRATERRESTRIAL PHENOMENA

**523.16** 101
Galactic Radio Emission and the Energy Released in Nuclear Collisions of Primary Cosmic-Ray Protons—G. R. Burbridge. (*Phys. Rev.*, vol. 103, pp. 264–265; July 1, 1956.) Approximate calculations are made based on the assumption of a uniform distribution of cosmic-radiation energy over the galactic disk and halo with a density of about $1eV/cm^3$, and of a rf radiation mechanism of synchrotron type. The results suggest that the rf radiation from our galaxy is a natural consequence of the presence of a cosmic-ray flux in the interstellar gas and magnetic field.

**523.16** 102
Non-ionospheric Fluctuations of the Intensity of R.F. Radiation from Nebulae—V. L. Ginzburg. (*C.R. Acad. Sci. U.R.S.S.*, vol. 109, pp. 61–63; July 1, 1956. In Russian.) Observed fluctuations may be explained as an effect of the motion of the earth relative to a diffraction pattern produced by the diffraction of the rf radiation at the edge of a hypothetical phase screen in space, or by a screen with periodic inhomogeneity. The screen is assumed to be very much nearer to the nebula than to the earth.

**523.16** 103
Emission Nebulae as Radio Sources—B. Y. Mills, A. G. Little, and K. V. Sheridan. (*Aust. J. Phys.*, vol. 9, pp. 218–227; June, 1956.) Results of observations on a wavelength of 3.5 m of 14 bright emission nebulas are discussed and values for electron densities and nebular masses are derived. Temperatures appear to be in the neighborhood of 10,000°K. except for the nebula NGC 6357, observed in absorption, for which a temperature of 6500°K. is estimated.

**523.16:523.42** 104
Rotation Period of the Planet Venus as determined by Radio Observations—J. D. Kraus. (*Nature, Lond.*, vol. 178, pp. 687–688; September 29, 1956.) The probable value of the rotation period deduced from the observations reported previously (3357 and 3358 of 1956) is 22 h 17 min, with an uncertainty of about $\pm 10$ min. A tentative explanation of the observed fluctuations of the rf signals is based on the effect of a hypothetical ionosphere surrounding Venus.

**523.16:523.72:621.396.677.3** 105
Interferometric Study of Brightness Dis-

tributions in Radio Astronomy—J. Arsac. (*Rev. d'Optique*, vol. 35, pp. 65–95, 136–165, and 396–413; February, March, and July, 1956.) The resolving power of interferometer antenna arrays is studied, the system being considered as a "spatial-harmonics filter"; the criterion of resolving power is the highest order of harmonics to which the system responds, but the form of the "spatial pass band" must be taken into account. The influence of the resolving power on the observations of some simple brightness distributions is studied, and possible errors in measurements of the apparent diameters of sources are indicated. Analysis is presented for some particular incomplete arrays; with four antennas, six harmonics of equal amplitude can be passed; with a large number $K$ of antennas the number of harmonics is about $K^2/4$. A fuller account is given of observations of the sun's brightness at 9.35 kmc reported previously (1966 of 1955).

**523.16:551.510.535** 106
Regions of the Ionosphere Responsible for Radio Star Scintillations—J. P. Wild and J. A. Roberts. (*Nature, Lond.*, vol. 178, pp. 377–378; August 18, 1956.) Recent observations in Australia of radio-star scintillations have been analyzed. The results indicate that night-time scintillations are correlated with spread F and day-time scintillations with sporadic E. The night-time observations are consistent with a west-to-east component of about 80 m in the motion of the scintillation patterns across the ground; such a motion would be caused by the earth's rotation if the irregularities were stationary and at a height of about 500 km.

**523.16:621.396.677** 107
Strip Integration in Radio Astronomy— R. N. Bracewell. (*Aust. J. Phys.*, vol. 9, pp. 198–217; June, 1956.) The resolution obtainable when the sky is canned by a strip-shaped antenna beam is discussed as a special case of two-dimensional antenna smoothing; it is shown that there is a principal solution for determining the true distribution in the scanned region, and a method is presented for reconstructing it from the observed data.

**523.5:621.396.11.029.62** 108
The Duration of Forward-Scattered Signals from Meteor Trails—Forsyth and Vogan. (See 257.)

**523.7:538.12** 109
The General Magnetic Field of the Sun— R. G. Conway. (*Observatory*, vol. 76, pp. 106–108; June, 1956.) From observations using a high-resolving-power interferometer system at a wavelength of 60 cm it is deduced that the magnetic field in the lower corona does not exceed a value corresponding to a surface polar field strength of 2.5 g.

**523.7:538.12** 110
The Sun's General Magnetic Field—H. W. Babcock. (*Nature, Lond.*, vol. 178, p. 533; September 8, 1956.) Arguments are presented in opposition to the views expressed by Alfvén (2714 of 1956).

**55:551.5:061.3** 111
Proceedings of the Washington Conference on Theoretical Geophysics, 1956. (*J. Geophys. Res.*, vol. 61, pp. 317–414; June, 1956.) Summaries are given of the papers presented, including several on the ionosphere, aurora, and atmospheric physics.

**550.372:621.396.945** 112
Determination of the Electrical Properties of Rock Strata from the Attenuation of Radio Waves—Tarkhov. (See 258.)

**550.382** 113
Numerical Integration of Geomagnetic Field Lines—L. Block and N. Herlofson. (*Tellus*, vol. 8, pp. 210–214; May, 1956.)

**551.510.41** 114
Infrared Evidence for the Presence of

535.37:537.226:537.311.33 135

The Photodielectric Effect in Zinc Sulphide and Oxide—J. Roux. (*Ann. Phys., Paris,* vol. 1, pp. 493–545; May/June, 1956.) A detailed investigation has been made of the variation of the electrical properties of ZnS and ZnO phosphors under the action of phosphorescence-producing radiation. Three different effects are distinguished, one of which corresponds to a modification of the dielectric properties, associated with trapping, while the other two correspond to photoconductivity variations operating in conjunction with barrier potentials and crystal shape factors respectively. 79 references.

535.37:546.472.21 136

Activator Systems in Zinc Sulphide Phosphors—J. S. Prener and F. E. Williams. (*J. Electrochem. Soc.,* vol. 103, pp. 342–346; June, 1956.) Measurements on ZnS containing Cu at random Zn sites show that such isolated impurities do not contribute to luminescence. A covalent model of the compound is proposed in which the activator and coactivator impurities act as acceptors and donors respectively, each located at substitutional sites.

535.37:546.472.21 137

Electroluminescence and Thermoluminescence of ZnS Single Crystals—G. F. Neumark. (*Phys. Rev.,* vol. 103, pp. 41–46; July 1, 1956.) Measurements have been made of the electroluminescence and thermoluminescence of ZnS separately and in combination. A degree of correlation was observed between the current passed by the crystal and the increment of luminous emission observed on applying a field during thermoluminescence. The results suggest that some of the electrons released from traps contribute to the current; electrons from deeper traps appear to be more effective than those from shallower traps.

535.376:546.472.21 138

Voltage Distribution inside Electroluminescent ZnS Crystals—G. Diemer and P. Zalm. (*Physics,* vol. 22, pp. 561–562; June, 1956.) Experimental results obtained by Frankl (1432 of 1956) are discussed and a theoretical analysis is made of the voltage distribution resulting after the build up of local voltage barriers.

537.226/.228:546.431.824-31 139

Growing Mechanism of Barium Metatitanate by the Firing Process—G. Ōhara. (*Sci. Rep. Res. Inst. Tohoku Univ., Ser. B,* vol. 6, pp. 117–136; 1954.) Report of an investigation of methods of preparing BaTiO₃ for use as a ceramic in electrical applications. Chemical analysis of fired specimens indicates that the material is best prepared by firing a stoichiometric mixture of $BaCO_3$ and $TiO_2$ for one hour at 1300° C. in the first place; moulding pressure is not important.

537.226/.228.1 140

Effect of Hydrostatic Pressure on the Hysteresis Loop of Guanidine Aluminium Sulphate Hexahydrate—W. J. Merz. (*Phys. Rev.,* vol. 103, pp. 565–566; August, 1956.) Measurements at a frequency of 60 cps indicate that both the spontaneous polarization and the coercive field strength of this material increase with pressure up to about 500 atm.

537.226/.227 141

Ferroelectricity in Ammonium Sulphate—B. T. Matthias and J. P. Remeika. (*Phys. Rev.,* vol. 103, p. 262; July 1, 1956.) A brief note reporting that $(NH_4)_2SO_4$ becomes ferroelectric below its transition point at −49.5° C.

537.226/.227 142

Dielectric Behavior of Lead Titanate at Low Temperature—J. Kobayashi, S. Okamoto, and R. Ueda. (*Phys. Rev.,* vol. 103, pp. 830–831; August 1, 1956.) Measurements made using a frequency of 1 mcs indicate the presence of anomalies in the dielectric-constant/temperature curve at about −100° and −150° C.

537.226/.227:546.431.824-31 143

Switching Time in Ferroelectric BaTiO₃ and its Dependence on Crystal Thickness—W. J. Merz. (*J. Appl. Phys.,* vol. 27, pp. 938–943; August, 1956.) "The switching time $t_s$ and the switching current $i_{max}$ have been measured as a function of applied field $E$ and of the size of the sample. It has been observed that the "activation field" $\alpha$ for the nucleation of new domains is inversely proportional to the thickness of the sample. This behavior can be explained by assuming a surface layer. The thickness of this layer has been calculated to be of the order of $10^{-4}$ cm. The same way we can explain the thickness dependence of the 60-cps coercive field strength. Furthermore, it has been found that the switching time depends to a first approximation linearly on the thickness of the sample if the field $E$ is kept constant. This can be explained by assuming a domain wall motion primarily in the forward direction or by assuming a nucleation mechanism. The maximum velocity of the domain growth was found to be of the order of the velocity of sound. The switching time does not depend on electrode area."

537.226/.227:546.431.824-31:539.234 144

Time Changes in Thin Films of BaTiO₃—C. Feldman. (*J. Appl. Phys.,* vol. 27, pp. 870–873; August, 1956.) "The decrease of dielectric constant with time under an applied alternating field, less than the coercive field, has been studied in films of BaTiO₃ between 1 and 3 μ thick. The phenomenon may be interpreted as being associated with the process of switching the domains to a position more nearly parallel to the applied field."

537.226/.227:621.315.612.4 145

Infrared Reflection Spectra of some Titanates—P. Turlier, L. Eyraud, and C. Eyraud. (*C. R. Acad. Sci., Paris,* vol. 243, pp. 659–660; August 13, 1956.)

537.226+538.22]:538.569.4.029.6 146

Dielectric Constants and Permeability of Artificial Dielectrics at 3 cm Wavelength—E. Meyer, H. J. Schmitt, and H. Severin. (*Z. Angew. Phys.,* vol. 8, pp. 257–263; June, 1956.) An account is given of an investigation on wide-band-radiation-absorbing media comprising mixtures of powdered graphite, ferroxcube, etc. in paraffin. A waveguide method of measurement was used. The results are in good agreement with measurements reported by Lewin (2139 of 1947). Magnetic losses pass through a maximum value at a particular value of particle size. Best results are obtained with iron and graphite; using 2 per cent by volume of these materials in paraffin, the attenuation at 3 cm λ is 75 and 45 db/m respectively.

537.226:546.212:621.317.335.3.029.64 147

A New Method for Measurement of Complex Dielectric Constant at Centimetre Wavelengths. Application to the Study of the Adsorption of Water—Le Bot. (See 226.)

537.311.31:537.311.1 148

Determination of the Free Path of Conduction Electrons from Galvanomagnetic Effects—H. Scheffers. (*Ann. Phys., Lpz.,* vol. 18, pp. 29–34; May 15, 1956.) A formula is derived for the variation of resistance of a pure univalent metal in a transverse magnetic field, and is used in conjunction with results of measurements on Au, Ag, and Cu to determine the free path of the conduction electrons; the values found agree reasonably well with those determined *e.g.* from the anomalous skin effect. The results demonstrate the validity of Fermi statistics.

537.311.33 149

A Few Notes on the Statistics of the Recombination and Trapping in Semiconductor—J. Nishizawa and Y. Watanabe. (*Sci. Rep. Res. Inst. Tohoku Univ., Ser. B,* vol. 7, pp. 149–170; December, 1955.) Improvements on the statistical theory of Shockley and Read (420 of 1953) are discussed.

537.311.33 150

Nature of an Ohmic Metal-Semiconductor Contact—F. A. Kröger, G. Diemer, and H. A. Klasens. (*Phys. Rev.,* vol. 103, p. 279; July 15, 1956.)

537.311.33:535.215:538.63 151

A Simplified Model for the Study of the Photomagnetoelectric effect [in semiconductors] in a Strong Magnetic Field—J. Lagrenaudie and A. A. Pires de Carvalho. (*Ann. Télécommun.,* vol. 11, pp. 127–130; June, 1956.) The model proposed explains many of the observed phenomena, in particular the Kikoin current/magnetic-field characteristics for the cases when surface recombination is 1) important or 2) negligible; the theory is confirmed by the experimental work of Kurnick and Zitter (2429 of 1956).

537.311.33:535.215:538.63 152

The Photomagnetoelectric Effect [in semiconductors]—J. Lagrenaudie. (*Ann. Télécommun.,* vol. 11, pp. 131–138; June, 1956.) An approximate theory is presented for the case of a weak magnetic field for constant and variable illumination. Use of the effect in determining many of the characteristics of semiconductors is discussed.

537.311.33:537.32 153

Single-Crystal Bismuth Telluride—L. Ainsworth. (*Proc. Phys. Soc.,* vol. 69, pp. 606–612; June 1, 1956.) Crystals about 1 cm × 2 cm × 3 cm, all of *p*-type, were produced by pulling from the melt, in a hydrogen atmosphere; values of the thermoelectric power, energy gap and electrical and thermal conductivities have been determined.

537.311.33:[546.28+546.289+546.682.86 154

Range-Energy Relation for Low-Energy Alpha Particles in Si, Ge, and InSb—G. W. Gobeli. (*Phys. Rev.,* vol. 103, pp. 275–278; July 15, 1956.) Measurements of the stopping power of thin foils of Si, Ge, and InSb for $\alpha$ particles of energies between 0.7 and 4.45 mev are reported. Measurements on Al, Cu, Ag, and Au were also made, so that the results could be compared with those obtained by other methods: good agreement was obtained. The ranges for the semiconductors interpolate smoothly between those of the metals.

537.311.33:[546.28+546.289 155

Impurity-Band Conduction in Germanium and Silicon—E. M. Conwell. (*Phys. Rev.,* vol. 103, pp. 51–61; July 1, 1956.) Theory adapted from that presented by Baltensperger (1451 of 1954) is used to estimate the concentrations in Ge and Si above which the impurity band merges into the conduction band; the results agree reasonably well with those obtained experimentally. An estimate is also made of the range of concentrations for which the usual band theory could be applied to electrons in the impurity band. An approximate theory of the conduction process is developed which can account for the sharp increase in impurity-band resistivity with decreasing impurity concentration, the importance of compensation, and the order of magnitude of the resistivity.

537.311.33:[546.28+546.289 156

Galvanomagnetic Theory for Electrons in Germanium and Silicon: Magnetoresistance in the High-Field Saturation Limit—L. Gold and L. M. Roth. (*Phys. Rev.,* vol. 103, pp. 61–66; July, 1956.) Simple theory based on a constant scattering time and ellipsoidal energy surfaces is used to derive a conductivity tensor. Calculated values of the saturation magnetoresistance are compared with experimental results discussed by Abeles and Meiboom (147 of 1955) and Shibuya (739 of 1955).

537.311.33:[546.28+546.289 157

Electron Multiplication in Germanium and Silicon—Y. Watanabe. (*Sci. Rep. Res. Inst. Tohoku Univ., Ser. B,* vol. 7, pp. 45–66; September, 1955.) Published experimental results

relating to carrier multiplication in Ge and Si are compared with ionization rates in gases; it is shown that the processes may be described by similar functions.

**537.311.33:[546.28+546.289]:536.21    158**
Thermal Conductivity of Germanium and Silicon at Low Temperatures—G. K. White and S. B. Woods. (*Phys. Rev.*, vol. 103, pp. 569–571; August 1, 1956.) Report of measurements made on high-purity *n*- and *p*-type Ge and on a single crystal of *n*-type Si. The results are discussed in the light of theories regarding the mechanism involved.

**537.311.33:[546.28+546.817.221 +546.722.221    159**
Effect of the Surface Treatment on the Crystal Rectifier: Part 1—J. Nishizawa and Y. Watanabe. (*Sci. Rep. Res. Inst. Tohoku Univ., Ser. B*, vol. 7, pp. 75–105; September, 1955.) Research work done from 1949 onwards on the control of the impurity concentration just below the surface of a semiconductor is discussed in the light of recent developments. The materials investigated are Si, FeS, and PbS.

**537.311.33:546.28    160**
Calculations on the Band Structure of Silicon—D. P. Jenkins. (*Proc. Phys. Soc.*, vol. 69, pp. 548–555; July 1, 1956.) "The electronic band structure of silicon has been calculated along two axes of momentum space. The results agree qualitatively with experiment, but quantitative values are distorted by uncertainties in the potential used. The variational cellular method used is not too laborious in application, but appears to suffer from unexpected disadvantages at low-symmetry points in momentum space."

**537.311.33:546.28    161**
Quenched-In Recombination Centers in Silicon—G. Bemski. (Phys. Rev., vol. 103, pp. 567–569; August 1, 1956.) "Measurements of lifetimes of minority carriers in *p*- and *n*-type silicon indicate that quenching from temperatures above 400° C. introduces recombination centers. The energy of formation of these centers is about 0.6 ev. These centers anneal at temperatures in the neighborhood of the quenching temperatures with an activation energy for annealing of about 0.8 ev."

**537.311.33:546.28    162**
Polarization of Phosphorus Nuclei in Silicon—G. Feher and E. A. Gere. (*Phys. Rev.*, vol. 103, pp. 501–503; July 15, 1956.) Report of an experimental verification of a scheme for polarizing nuclei proposed previously [1760 of 1956 (Feher *et al.*)]. The polarization of the nuclei is deduced from electron-spin resonance lines which indicate the population of the energy levels.

**537.311.33:546.28    163**
Concentration Effects on the Line Spectra of Bound Holes in Silicon—R. Newman. (*Phys. Rev.*, vol. 103, pp. 103–106; July 1, 1956.) "The effect of impurity concentration on the line spectra of holes bound to B, Al, and Ga acceptors in silicon has been studied at 21°K. Details of the spectra (*i.e.*, relative positions, line shapes, intensities) differ for the different acceptors. Below an acceptor concentration of about $10^{16}/cm^3$ the spectra are concentration-independent under the conditions of measurement. Above $10^{16}/cm^3$ the spectral lines begin to broaden and by $\sim10^{18}$ /cm³ the line structure has been almost completely destroyed. A qualitative discussion of the concentration effects is given."

**537.311.33:546.28    164**
Avalanche Breakdown Voltage in Silicon Diffused p-n Junctions as a Function of Impurity Gradient—H. S. Veloric, M. B. Prince, and M. J. Eder. (*J. Appl. Phys.*, vol. 27, pp. 895–899; August, 1956.) "A method is presented for controlling the reverse breakdown voltage ($V_B$) in a silicon graded junction. The significant process parameters are shown

to be resistivity, time of diffusion, and temperature of diffusion. For a constant resistivity, $V_B$ increases with the fourth root of the time of diffusion and with the square root of the depth of diffusion as predicted by theory. Statistical analysis shows that the mean breakdown voltage for a large group of units can be predicted within 2 per cent. The method fails for very low or high resistivity material."

**537.311.33:546.28    165**
Spin-Spin Paramagnetic Relaxation in a Semiconductor—A. Abragam and J. Combrisson. (*C.R. Acad. Sci., Paris*, vol. 243, pp. 650–652; August 13, 1956.) Measurements are reported on As-doped Si subjected to a magnetic field whose intensity is reduced rapidly from a high to a low value.

**537.311.33:546.289    166**
Purification of GeCl₄ by Extraction with HCl and Chlorine—H. C. Theuerer. (*J. Metals, N.Y.*, vol. 8, pp. 688–690; May, 1956.)

**537.311.33:546.289    167**
Separation of Germanium and Cadmium from Zinc Concentrates by Fuming—H. Kenworthy, A. G. Starliper, and A. Ollar. (*J. Metals, N.Y.*, vol. 8, pp. 682–685; May, 1956.)

**537.311.33:546.289    168**
On the Preparation and Regeneration of Clean Germanium Surfaces—P. H. Robinson, A. J. Rosenberg, and H. C. Gatos. (*J. Appl. Phys.*, vol. 27, p. 962; August, 1956.) Brief report of experiments on the thermal regeneration of oxygenated Ge surfaces originally created by crystal cleavage in vacuum. Effects due to CO picked up from graphite crucibles are mentioned.

**537.311.33:546.289    169**
Structure Sensitivity of Cu Diffusion in Ge—A. G. Tweet and C. J. Gallagher. (*Phys. Rev.*, vol. 103, p. 828; August 1, 1956.) Experiments show that the diffusion depends markedly on the degree of perfection of the Ge crystals.

**537.311.33:546.289    170**
Light Emission due to Recombination via Traps in Germanium—C. Benoit à la Guillaume. (*C. R. Acad. Sci., Paris*, vol. 243, pp. 704–707; August 20, 1956.) Formulas are derived for the infrared emission to be expected when recombination takes place 1) directly and 2) indirectly, via traps. Low-temperature measurements on *n*-type specimens with In or Al junctions indicate a maximum in the emission spectrum for process 2) at a wavelength of about $2.45\,\mu$ giving an energy level about 0.53 ev below the conduction band for the traps. No emission due to process 2) is observed at temperatures above 200° K.

**537.311.33:546.289:537.52    171**
Effects of Low-Energy Gas Discharges on Evaporated Metal/Semiconductor Contacts—P. A. Hartig and R. N. Noyce. (*J. Appl. Phys.*, vol. 27, pp. 843–847; August, 1956.) An investigation is reported of the effects on the rectification properties of changes in the states at the oxide/semiconductor interface. Specimens of *n*-type Ge were exposed to discharges in various gases prior to deposition of platinum contacts. Measurements on the diodes thus formed indicated that oxygen and nitrogen lower the surface barrier, giving rise to ohmic contacts, while hydrogen and argon tend to keep the surface region highly *p*-type. The mechanisms responsible for these different types of behavior are discussed.

**537.311.33:546.289:543.23-8    172**
Effects of Copper on Ultrasonic Attenuation in Germanium—L. J. Teutonico, A. Granato, and R. Truell. (*Phys. Rev.*, vol. 103, pp. 832–833; August 1, 1956.) Experimental evidence indicates that observed variations in the magnitude of the attenuation are caused by precipitation of Cu on dislocations in the Ge.

**537.311.33:546.289:621.396.822    173**
Relaxation Time of Surface States on Germanium—R. H. Kingston and A. L. McWhorter. (*Phys. Rev.*, vol. 103, pp. 534–540; August 1, 1956.) Surface states associated with adsorbed ions or imperfections in the oxide layer are believed to control the position of the energy bands at the surface with respect to the Fermi level. A study of these states by the "field-effect" method is reported; the field is applied perpendicular to the surface and the change in conductance is observed. The results indicate that the relaxation or capture time of these states is much longer than that of the interface states, and depends on the surface treatment and ambient gas. Some surface treatments produce a very wide range of time constants. The significance of the results in relation to the inverse frequency variation of excess noise is discussed.

**537.311.33:546.46.814    174**
Electrical Conduction in Magnesium Stannide at Low Temperatures—H. P. R. Frederikse, W. R. Hosler, and D. E. Roberts. (*Phys. Rev.*, vol. 103, pp. 67–72; July 1, 1956.) Measurements of conductivity, Hall effect and magnetoresistance of Mg₂Sn were made at temperatures between 2° and 80°K; anomalous variations are observed similar to those of Ge and other semiconductors. The results obtained at the lowest temperatures suggest that conduction takes place in a surface layer rather than at impurity levels in the bulk of the material. See also 1104 of 1956. (Blunt *et al.*).

**537.311.33:546.482.21:537.525.8    175**
The Influence of a Glow Discharge on the Conductivity of Cadmium-Sulphide Single Crystals and the Production of Ohmic Contacts—J. Fassbender. (*Z. Phys.*, vol. 145, pp. 301–318; May 11, 1956.) Both the dark conductivity and the photoconductivity are increased by several orders of magnitude as a result of exposure of the crystal to the glow discharge. The effect is attributed to reduction of the layers close to the surface, whereby sulphur vacancies occur in the lattice. The production of nonrectifying contacts results from the production of an excess boundary layer due to a high concentration of impurity centers.

**537.311.33:546.482.31    176**
Crystallization of Hexagonal Cadmium Selenide with ZnSe, InAs, and In₂Se₃—N. A. Goryunova, V. A. Kotovich, and V. A. Frank-Kamenetski. (*Zh. Tekh. Fiz.*, vol. 25, pp. 2419–1421; December, 1955.) The synthesis was carried out by alloying stoichiometric quantities of the constituents in electric ovens. The synthesized specimens were studied by X-ray diffraction. The investigation shows that different types of forced crystallization of CdSe are obtained in the presence of crystalline phases of other substances.

**537.311.33:546.561-31    177**
Electrical Conductivity of Cuprous Oxide—A. I. Andrievski, V. I. Voloshchenko, and M. T. Mishchenko. (*Zh. Tekh. Fiz.*, vol. 25, pp. 2422–2427; December, 1955.) Measurements were made on specimens prepared by different methods and having different grain sizes and different numbers of grains per unit surface. The results are fully reported and discussed. The conductivity is proportional to the number of grains per unit surface of the specimen, independent of the method of preparation.

**537.311.33:546.621.86    178**
Electrical Properties of Aluminium Antimonide—F. Kover. (*C.R. Acad. Sci., Paris*, vol. 243, pp. 648–650; August 13, 1956.) Measurements have been made of the Hall and Seebeck constants and the resistivity of *p*-type specimens of AlSb. The results are consistent with degeneracy of the specimens, indicating strong curvature of the valence band.

**537.311.33:[546.623-31+546.47-31    179**
Electrical Conductivity of Aluminium Oxide

in the light of experimental results obtained on permalloys.

**538.221:621.318.12** 202

Structure and Magnetic Properties of Permanent-Magnet Alloys during Isothermal Precipitation Hardening: Part 1—Grain Growth, Critical Grain Size and Coercive Force—E. Biedermann and E. Kneller. (*Z. Metallkde*, vol. 47, pp. 289–301; May, 1956.) X-ray and electron-microscope observations of grain growth were made on Cu-Ni-Fe and Cu-Ni-Co alloys for different values of annealing time; coercive-force measurements were also made, and the grain size corresponding to maximum coercive force was determined.

**538.221:621.318.134:621.372** 203

Ferrites—(PROC. IRE vol. 44; October, 1956.) The main part of this issue is devoted to a group of papers together providing a comprehensive survey of the physical properties of ferrites and their applications in electronics. Abstracts of some of these papers are given individually; titles of the others are as follows:—

A Survey of the Properties and Applications of Ferrites Below Microwave Frequencies—C. D. Owens (pp. 1234–1248).

Fundamental Theory of Ferro- and Ferrimagnetism—J. H. Van Vleck (pp. 1248–1258).

Magnetic Resonance in Ferrites—N. Bloembergen (pp. 1259–1269).

The Nonlinear Behavior of Ferrites at High Microwave Signal Levels—H. Suhl (pp. 1270–1284).

Microwave Resonance Relations in Anisotropic Single-Crystal Ferrites—J. O. Artman (pp. 1284–1293).

Dielectric Properties of and Conductivity in Ferrites—L. G. Van Uitert (pp. 1294–1303).

Methods of Preparation and Crystal Chemistry of Ferrites—D. L. Fresh (pp. 1303–1311).

Intrinsic Tensor Permeabilities on Ferrite Rods, Spheres, and Disks—E. G. Spencer, L. A. Ault, and R. C. LeCraw (pp. 1311–1317).

Permeability Tensor Values from Waveguide Measurements—E. B. Mullen and E. R. Carlson (pp. 1318–1323).

Resonance Loss Properties of Ferrites in 9-kmc Region—S. Sensiper (pp. 1323–1342).

Anisotropy of Cobalt-Substituted Mn-Ferrite Single Crystals—P. E. Tannenwald and M. H. Seavey (pp. 1343–1344).

The Elements of Nonreciprocal Microwave Devices—C. L. Hogan (pp. 1345–1368).

Frequency and Loss Characteristics of Microwave Ferrite Devices—B. Lax (pp. 1368–1386).

Ferrites as Microwave Circuit Elements—G. S. Heller (pp. 1386–1393).

Topics in Guided-Wave Propagation in Magnetized Ferrites—M. L. Kales (pp. 1403–1409).

**538.221:621.385.833** 204

A Study of Bitter Figures using the Electron Microscope—D. J. Craik. (*Proc. Phys. Soc.*, vol. 69, pp. 647–650; June 1, 1956.) "The Bitter figure technique is modified to give patterns in the form of thin films which be can removed from the specimen and examined by high power phase-contrast microscopy and by electron microscopy. This makes possible the observation of extremely fine detail and direct measurement of the width of domain walls."

**538.221:621.385.833** 205

Observation of Magnetic Domains in Electron-Shadow Photographs—M. Blackman and E. Grünbaum. (*Nature, Lond.*, vol. 178, pp. 584–585; September 15, 1956.) A cusped-edge effect observed in the electron-shadow photograph of an unmagnetized single crystal of hexagonal cobalt is explained on the basis of the domain structure of the crystal.

**539.23:546.621-31:537.533.7** 206

Penetration of Electrons in Aluminum Oxide Films—J. R. Young. (*Phys. Rev.*, vol. 103, pp. 292–293; July 15, 1956.) Measurements on films of thickness 85–5000 Å indicate that the relation between penetration range $R$, in mg/cm$^2$, and electron energy $E$, in keV, is expressed by $R = 0.0115 E^{1.35}$.

**548.5** 207

The Distribution of Impurity in a Semi-infinite Solidified Melt—O. W. Memelink. (*Philips Res. Rep.*, vol. 11, pp. 183–189; June, 1956.) An expression is obtained for the transient impurity distribution in a crystal pulled from an impure melt which is identical with that derived by different methods by other workers [3675 of 1955 (Hulme) and 1758 of 1956 (Smith *et al.*)]; a table and curves of calculated results are given.

**621.315.61:537.226** 208

The "Representivity" of Values of Dielectric Properties of Electrical Insulating Materials in Relation to the Physical and Chemical Factors—W. M. H. Schulze. (*Schweiz. Arch. Angew. Wiss. Tech.*, vol. 22, pp. 137–149; May, 1956.) Apparent discrepancies between published values of dielectric properties are attributed to insufficiently close specification of the history of the material and the test conditions; an indication is given of the most important factors involved.

**621.315.61:537.533** 209

Experimental Investigation of Insulating Surfaces by means of Field Emission—H. Klumb and K. Neubeck. (*Naturwissenschaften*, vol. 43, p. 247; June, 1956.) A thin layer of the insulating material is coated on to the tungsten point of a field-emission microscope, a monatomic layer of metal is evaporated on to it from the side, and the distribution of field emission over the surface is then observed.

**621.315.612.6** 210

Effect of Pressure on Glass Structure—O. L. Anderson. (*J. Appl. Phys.*, vol. 27, pp. 943–949; August, 1956.)

**621.315.615:537.533** 211

Controlled Field Emission in Hexane—W. B. Green. (*J. Appl. Phys.*, vol. 27, pp. 921–925; August, 1956.) The investigation of the properties of hexane described previously (846 of 1956) has been extended to a study of the potentialities of this material as a field-emission source.

**621.315.616** 212

Studies on the Dielectric Polymers: Part 1—The Electrical Properties of the Dielectric Rubber—M. Matsudaira and K. Takei. (*Sci. Rep. Res. Inst. Tohoku Univ.*, Ser. B., vol. 6, pp. 11–17; 1954.) Mixtures of various plastics with powdered BaTiO₃ give materials having dielectric constants of 10–200; tan $\delta$ is $<0.05$ at room temperature.

**621.315.616.96:621.3.002.2** 213

Epoxide Resins in the Electronics Industry—A. G. Goodchild. (*Brit. Commun. Electronics*, vol. 3, pp. 293–299; June, 1956.) Practical hints are given on the handling of epoxide resins; the properties of some typical potting compositions are tabulated.

**676.4:621.395.623.7** 214

On a Model of Paper with the Dynamic Properties [suitable for loudspeaker cones]—T. Nimura and K. Kido. (*Sci. Rep. Res. Inst. Tohoku Univ.*, Ser. B, vol. 6, pp. 45–62; 1954.)

**537.311.33:621.314.7** 215

Progress in Semiconductors [Book Review]—A. F. Gibson, R. E. Burgess, and P. Aigrain (Eds). Publishers: Heywood, London, 220 pp., 50s, 1956. (*Nature, Lond.*, vol. 178, pp. 508–509; September, 1956.) Contains sections on 1) the preparation, properties and applications of Si; 2) the importance of Ge filaments in physical studies; 3) the Seebeck effect in semiconductors; 4) the electrical properties of phosphors; 5) the advantages of $p$-$n$-$i$-$p$ and unipolar transistors relative to conventional types; 6) photomagnetoelectric effects; 7) field effects.

## MATHEMATICS

**517.9** 216

On Stability Questions for Pendulum-Type Equations—G. Seifert. (*Z. Angew. Math. Phys.*, vol. 7, pp. 238–247; May, 1956. In English.)

**517.949.8** 217

Graphical Solution of certain Nonlinear Differential-Difference Equations—W. J. Cunningham. (*J. Franklin Inst.*, vol. 261, pp. 621–629; June, 1956.) Equations relevant to nonlinear oscillators (51 above) are discussed.

**517** 218

Handbuch der Laplace-Transformation; Band II: Anwendung der Laplace-Transformation (I. Abteilung) [Book Review]—G. Doetsch. Publishers: Birkhäuser, Basel and Stuttgart, 436 pp., D.M. 56; 1955. (*Arch. Elekt. Übertragung*, vol. 10, p. 223 May, 1956.) A clear and detailed exposition of the application of Laplace transforms in asymptotic series expansions, convergent series expansions, and ordinary differential equations.

## MEASUREMENTS AND TEST GEAR

**531.76:621.372.632** 219

High-Resolution Millimicrosecond Time-Interval Measurements based upon Frequency Conversion—C. Cottini, E. Gatti, and G. Giannelli. (*Nuovo Cim.*, vol. 4, pp. 156–157 July 1, 1956. In English.) Two time-separated pulses respectively excite two high-$Q$ circuits whose resonance frequencies differ by *e.g.* 1 per cent at about 30 mc. The resulting trains of damped oscillations are mixed in a Ge-diode ring demodulator; the time separation between the pulses is determined from the phase of the resultant oscillation.

**621.3.018.41(083.74)+529.786]:621.396.91** 220

Standard Frequencies and Time Signals WWV and WWVH—H. Amemiya. (PROC. IRE, vol. 44, pp. 1470–1473; October, 1956.) Details are given of routine services provided by broadcasting from the National Bureau of Standards stations, including standard radio and audio frequencies, standard time intervals, standard musical pitch, time signals, and radio propagation forecasts. Similar information is given in *Electronic Ind. and Tele-Tech.*, vol. 15, pp. 65–66, 111; September, 1956.

**521.317.3:621.317.733** 221

Errors in Bridge Measurements—C. G. Mayo and J. W. Head. (*Wireless Eng.*, vol. 33, pp. 265–268; November, 1956.) "Any linear network used as a bridge can be regarded as subject to three errors which are independent of the unknown and standard impedances (or admittances) being compared. These errors (which may be complex quantities having any argument) are: 1) a ratio error, 2) an error impedance in series with the standard, 3) an error admittance in parallel with the unknown. One of these errors can be reduced to zero by adjustment; the other two can be determined by measuring two known impedances (or admittances) of different orders of magnitude."

**621.317.3:621.396.822:537.311.33** 222

Fluctuation Noise—F. J. Hyde. (*Wireless Eng.*, vol. 33, pp. 271–276; November, 1956.) "A method of measuring the current noise generated in temperature-sensitive solid-state devices, in narrow bandwidths about frequencies chosen between $5 \times 10^{-3}$ and $10^{-1}$ c/s is described. The central feature of the equipment is a resistance-capacitance coupled amplifier of the type originally described by Schneider [1207 of 1946]. Spurious electrical fluctuations arising from the circuitry and thermal fluctuations of the ambient oil surrounding the specimen were shown not to cause any error. Specification of the noise necessitates determination of the impedance of the specimen at the frequencies in question: this may be derived from measurements of the limiting small-signal resistances of the speci-

taining three terms associated respectively with the ground wave, the space wave and the diffracted wave, for propagation across a discontinuity such as a sea/land boundary; the treatment is extended to cover propagation across a number of boundaries. Divergencies from Millington's results (1758 fo 1949) are discussed.

**621.396.11:551.510.535                    252**
Scattering of Radio Waves and the Horizontal Gradient of Ionization in the Ionosphere—S. S. Banerjee and P. G. Surange. (*Sci. and Cult.*, vol. 21, pp. 750–753; June, 1956.) It is shown that a linear $P'f$ curve for back-scattered echoes via the $F_2$ layer is obtained when there is no horizontal gradient of ionization; a nonlinear $P'f$ curve is obtained when such a gradient exists. Experimentally derived curves are presented illustrating both cases.

**621.396.11:621.396.674.3                    253**
Low-Frequency Radiation from a Horizontal Antenna over a Spherical Earth—J. R. Wait. (*Canad. J. Phys.*, vol. 34, pp. 586–595; June, 1956.) The problem is treated by introducing a boundary impedance characterizing the earth's surface; scalar wave functions are used in the analysis. The results indicate that, in general both vertically and horizontally polarized ground waves are propagated, but the horizontally polarized component has negligible magnitude at low radio frequencies, by Norton (PROC. IRE, vol. 25, pp. 1203–1236; September, 1937.)

**621.396.11.029.5                    254**
Statistics of Long-Distance Night-Time Propagation of Long and Medium Waves—W. Ebert. (*Tech. Mitt. Schweiz. Telegr.-TelephVerw.*, vol. 34, pp. 198–209; May 1, 1956.) The determination of the time distribution of field-strength values from a record of sky-wave measurements is described. Analysis of measurements made under the auspices of the Union Européene de Radiodiffusion over the period 1952–1955 indicates that the mean distribution obtained from records for periods of one or two hours is fairly close to a Rayleigh distribution. The distribution of the diurnal median field-strength values and the construction of night-time propagation curves on the basis of these data are discussed.

**621.396.11.029.6                    255**
Propagation Properties of Metre Waves at Not Very Great Distances from the Transmitter—J. Houtsmuller. (*Tijdschr. Ned. Radiogenoot.*, vol. 21, pp. 103–144; May, 1956.) Propagation theory for a spherical earth is summarized, and vhf service-area field-strength measurements made by the Netherlands Post Office are reported. Variations due to irregularities of terrain are noted, and the influence of the receiving-antenna height is studied. Field distortion at 67.5 mc due to a water tower, and at 97.5 mc due to a house, are illustrated. Marked seasonal effects due to vegetation were observed. The reflection coefficient of a water surface for waves of 8.8 cm λ and the influence of tides are also discussed.

**621.396.11.029.6                    256**
On the Method of Estimating Electric Field Intensity in V.H.F. Propagation in Mountainous Region—Y. Nomura. (*Sci. Rep. Res. Inst. Tohoku Univ.*, *Ser. B*, vol. 6, pp. 157–166; 1955.) The theory of diffraction by parallel screens established previously (2494 of 1954) is applied; practical examples are given.

**621.396.11.029.62:523.5                    257**
The Duration of Forward-Scattered Signals from Meteor Trails—P. A. Forsyth and E. L. Vogan. (*Canad. J. Phys.*, vol. 34, pp. 535–545; June, 1956.) Measurements were made over an east–west and a north-south path in Canada, both about 1000 km long, using frequencies of about 38 and 50 mc simultaneously. The duration of signals observed over the east-west path appeared to be greater than the duration

of those over the north-south path. The variation with wavelength and equipment sensitivity is similar to that for back-scattered signals studied by McKinley (*ibid.*, vol. 31, pp. 758–767; July, 1953, and 1399 of 1954). While the variation with equipment sensitivity is consistent with accepted theories, the variation with wavelength has not yet been explained.

**621.396.945:550.372                    258**
Determination of the Electrical Properties of Rock Strata from the Attenuation of Radio Waves—A. G. Tarkhov. (*Bull. Acad. Sci. U.R.S.S., Sér. Géophys.*, pp. 599–608; May, 1956. In Russian.) Formulas are given relating the dielectric constant and the conductivity of the medium with the attenuation of waves at two frequencies in the range between about $10^4$ and $10^7$ cps. Results of experiments involving the transmission of such waves through the earth are reported.

## RECEPTION

**621.396.62:621.396.822                    259**
The Improvement of Signal/Noise Ratio by Signal Repetition—G. Günther and G. Kraus. (*Frequenz.*, vol. 10, pp. 169–177; June, 1956.) The factors governing the improvement in signal/noise ratio attainable by signal repetition are examined theoretically. Experimental results are basically confirmed, although a general solution is not obtained. A more detailed analysis, including statistical considerations, shows that the improvement factor due to *n*-fold addition lies between on and *n*.

**621.396.621.54:621.376.33                    260**
U.S.W. F.M. Reception with Low Distortion—U. Köhler. (*NachrTech.*, vol. 6, pp. 219–225; May, 1956.) The design of a receiver to have an over-all distortion factor not exceeding 0.4 per cent is discussed, the distortion being distributed equally between rf, IF, demodulator and af stages. Detailed attention is given mainly to the IF and demodulator stages. The design developed includes four IF stages and a Foster-Seeley discriminator; the adjacent-channel selectivity is 28 db, and tubes can be changed without necessitating re-alignment.

**621.396.621.54.029.55:621.314.7                    261**
An Experimental All-Transistor Communications Receiver—C. J. Heinen. (*QST*, vol. 40, pp. 11–16; May, 1956.) A superheterodyne receiver for wavelengths between about 15 and 80 m is described. The transistor types used include a surface-barrier SB-100 in the hf oscillator and mixer stages and 2N76 or similar in the two 455-kc IF stages and the beat frequency oscillator and af stages; a Type-2N34 crystal diode is used in the second detector stage.

**621.396.823:621.315                    262**
Design Features of Insulators and Transformers which Contribute to Radio Interference from Power Lines—J. W. Orner. (*Trans. S. Afr. Inst. Elect. Eng.*, vol. 47, pp. 153–171. Discussion, pp. 171–178; May, 1956.) Investigations indicate that interference from power lines is due mainly to the ionization of small pockets of air associated with insulating material. Methods of testing for the existence of such overstressed air gaps are suggested. Freedom from interference can theoretically be achieved on any power line, but economic considerations may constitute a limiting factor.

## STATIONS AND COMMUNICATION SYSTEMS

**621.39:534.78:621.396.61                    263**
A New System for Speech Transmission—P. Marcou and J. Daguet. (*Ann. Télécommun.*, vol. 11, pp. 118–126; June, 1956.) A speech signal of the form $a(t) \cos \phi(t)$ may be transformed by simple apparatus into a constant-level signal of the form $\cos \phi(t)$, without loss of information value. A.s.s.b. radiotelephony transmitter using class-C amplification has

been constructed based on this principle; the performance is markedly superior to that of conventional transmitters of comparable power.

**621.39.001.11                    264**
Improvement of Communication Efficiency during a Finite Interval of Time—N. Honda. (*Sci. Rep. Res. Inst. Tohoku Univ., Ser. B.*, vol. 6, pp. 167–182; 1955.) A method is described for calculating the efficiency of coding of messages of finite length, for channels with and without noise; the results are compared with those of Shannon for messages of infinite length.

**621.395.44.078                    265**
The Design of Automatic Pilot Regulators—G. Kraus. (*Arch. Elekt. Übertragung*, vol. 10, pp. 175–187; May, 1956.) Automatic level control and equalization for long-distance carrier-current systems is discussed.

**621.396.1                    266**
Danger! The Radio Spectrum is bursting at the Seams—D. G. Fink. (*J. Franklin Inst.*, vol. 261, pp. 477–493; May, 1956.) Discussion of frequency-allocation problems in the U.S.A. Suggestions for ameliorating the position are offered. See also 2446 of 1953.

**621.396.41:621.376.3:621.396.82                    267**
Interference due to Echoes in F.M. Multichannel Radio Links—G. Bosse and M. Wagner. (*Frequenz*, vol. 10, pp. 137–142; May, 1956.) If the same frequency deviation is used in all channels, interference caused by echoes due to multipath propagation or mismatch between antenna and feeder is greater in the higher-frequency channels, where path noise is also greatest. Calculations are made of the improvement which can be effected in the uniformity of the distribution of interference over the channels by introducing pre-emphasis. The results are shown in families of signal-noise-ratio frequency characteristics for different values of echo time delay, with amount of pre-emphasis as parameter. Methods described by Bennett *et al.* (3089 of 1955) are used.

**621.396.65/.66                    268**
Speech-Controlled Blocking Circuits in Radiotelephony—K. Fischer. (*Elektronische Rundschau*, vol. 10, pp. 158–162; June, 1956.) Circuits are discussed which enable the incoming and outgoing radio paths to be isolated while ensuring good transmission between both of these paths and the subscriber's line. The suitability of various available types of blocking circuit for use in different communication systems is assessed from the level available at the two-wire line input in relation to the gain in the transmission path and the signal/noise ratio at the receiver output.

**621.396.93                    269**
Equipment for Mobile and Fixed Radio Services—(*Elektronische Rundschau*, vol. 10, pp. 173–174, 178; June, 1956.) Communication equipment shown at the 1956 German Industries Fair at Hanover is described.

## SUBSIDIARY APPARATUS

**621.311.6:621.316.722.2/.3                    270**
Generator of Linearly Varying Voltage—J. Levinson. (*Rev. Sci. Instrum.*, vol. 27, pp. 536–539; July, 1956.) A square-wave generator provides the input to a parallel feedback integrating circuit consisting of a series resistance and a capacitance in parallel with two dc amplifiers. Linear variation is obtained in both rising and falling half cycles; maximum current in either direction is 3 a, and the cycling period is up to 1 min.

## TELEVISION AND PHOTOTELEGRAPHY

**621.397.3                    271**
On the Theory of Mertz and Gray—H. Schönfelder. (*Frequenz*, vol. 10, pp. 142–147; May, 1956.) The theory of the television-signal spectrum developed by Mertz and Gray (*Bell Syst. Tech. J.*, vol. 13, pp. 464–515; July, 1934).

is discussed on the basis of pulse analysis. Spectrum lines at odd multiples of half-line-frequency are absent when an oldd number of scanning lines is used, in consequence of phase reversals at frame frequency. The analysis is used to examine conditions when a color subcarrier is used, both still and moving picture subjects being considered. The correctness of the theory is confirmed for the case of a diagonal-strip picture also.

**621.397.5** 272
Simultaneous Television—G. Muller. (*Television*, pp. 119–120; May, 1956.) An outline is presented of a suggested system whereby all the points of the television picture could be transmitted simultaneously. A photoemissive mosaic is combined with a transparent electrode to form a capacitor comprising a multiplicity of elementary capacitors; by arranging the two plates to be nonparallel, the capacitance variations associated with different distributions of illumination can be given characteristic values. These capacitance variations can be used to vary the frequency or other parameter of a carrier wave. Quite a narrow band should suffice for the transmission.

**621.397.5** 273
Spectrum of Television Signals—D. A. Bell and G. E. D. Swann. (*Wireless Eng.*, vol. 33, pp. 253–256; November, 1956.) Systematic measurements were made at Birmingham on the signals broadcast from the B.B.C. station at Sutton Coldfield. The video-frequency band from zero to 11 kc was studied using an analyzer with a bandwidth of 4 cps, and the general trend of the spectrum over the ranges 90–550 kc and 1.38–1.97 mc was explored using a narrow-band communication receiver. The results indicate that the amplitudes of the line-frequency harmonics decrease faster than $1/n$, where $n$ is the order of the harmonic. The relation between peak-white power and sideband power is discussed as a basis for estimating the risk of interference with other services.

**621.397.5:535.623** 274
Amateur Television—M. Barlow. (*Wireless World*, vol. 62, pp. 548–549; November, ber, 1956.) An account is given of progress made since the previous report (541 of 1954). Interesting results have been obtained with color systems. Photoconductive camera tubes are used most widely.

**621.397.5+681.142]:621.395.625** 275
Recording and Reproduction of Frequencies above 100 kc/s on Magnetic Media. Application to Storage Elements and to the Recording of Television Images—Charlet. (See 33.)

**621.397.6:535.623** 276
Directions of Improvement in N.T.S.C. Color-Television Systems—D. Richman. (Proc. IRE, vol. 44, pp. 1125–1139; September, 1956.) The N.T.S.C. television signal is analyzed and proposals are made for modifying it so as to improve the resolution and facilitate compatibility while permitting use of a simplified receiver with equal, wide, flat pass bands.

**621.397.6.001.4:621.317.3** 277
A Self-Synchronizing Line Selector for Television—O. Macek. (*Frequenz*, vol. 10, pp. 193–197; June, 1956.) The principles of operation of a line selector are described, and details are given of an instrument suitable for use at any point in a television transmission system where the complete video waveform is available. Narrow variable-width columns as well as lines can then be displayed on a monitoring cro.

**621.397.61:535.623** 278
Color TV Commercials by Use of Direct Artwork with Opaques—P. B. Laeser. (*J. Soc. Mot. Pict. Telev. Eng.*, vol. 65, pp. 284–286; May, 1956.) Short description of flying-spot scanning equipment for dealing with material prepared on opaque cards about 3/16 in. thick,

the useful area being up to 9 in. ×12 in. The pickup unit uses three photocells.

**621.397.61.029.62** 279
Band-I Television-Transmitter Design, with Particular Reference to the Transmitters at Crystal Palace—V. J. Cooper and W. J. Morcom. (*Proc. IEE*, Part B, vol. 103, pp. 651–663. Discussion, pp. 663–666; September, 1956.) Reliability is increased by designing the equipment to operate in two parallel chains. Noise-rejection and clamp circuits are provided, thus ensuring satisfactory operation in the presence of distortion and noise in the input signal. Modulation is at high level; use of tetrodes in the output stage permits use of receiver-type valves throughout the modulator. Steps taken to make the transmitters adaptable for color transmissions with an N.T.S.C.-type signal are indicated. Power output is 17 kw for the vision transmitter, and 4'/4 kw for the sound transmitter.

**621.397.611.2:621.383.4** 280
Transmitting Television Tube with Photoresistor—N. L. Artem'ev, V. K. Sokolov, and S. K. Temiryazeva. (*Radiotekhnika i Electronika*, vol. 1, pp. 245–252; February, 1956.) A vidicon-type camera tube using a stibnite photoconductive layer is described its characteristics are presented graphically. With 10–50 v on the signal plate, an illumination of 5–50 lux, and a velocity of the projected image of a moving object on the screen of 3 mm/sec, the resulting signal current is 0.2–0.8 $\mu$a and the resolution up to 300 lines.

**621.397.62** 281
Four-Standards Television—H. d'L. Banting. (*Wireless World*, vol. 62, pp. 559–562; November, 1956.) The problems of designing television receivers for use in Belgium are discussed. The four standards used are: the French 819–line system, the Belgian 625- and 819-line systems and the C.C.I.R. 625-line system with negative modulation and fm sound.

**621.397.62:621.385.832** 282
The Television Picture Tube—Rothe and Gundert. (See 336.)

**621.397.8** 283
Contribution to the Study of Echo Phenomena in Television—J. Polonsky, L. Amster, and G. Melchior. (*Ann. Radioélect.*, vol. 11, pp. 57–69; January, 1956.) Methods of reducing echoes by modifying the transmitting-antenna installation are discussed. Large reductions can be obtained by using wide-band antennas with polyphase feed. Alternatively, echoes can be compensated by coupling two transmitters to a common antenna or by inserting an echo of opposite sense at rf or video frequency.

**621.397.8** 284
The Evaluation of Geometrical Distortion in Television Pictures—H. Stier. (*NachrTech.*, vol. 6, pp. 248–251; June, 1956.) Geometrical distortion may be determined in terms of either the variation of the spot tracing velocity or the distance of the spot from its assigned position at a particular instant. The estimates obtained by the two methods differ widely and at different rates with respect to the cause of the distortion, which may lie in the optical, electron-optical, or circuit parts of the system. The permissible limits of distortion are discussed.

**621.397.8:621.396.11** 285
Practical Tests of Picture Reception in Mountainous Territory—R. A. Raffin. (*Télév. Franç.*, no. 131, pp. 6–8; June, 1956.) Tests were made of the reception of signals from the 200-w television station Lyon-Fourvière at Roanne, about 70 km away. The receiver used was a modified form of one described previously (*ibid.*, no. 118, p. 20; May, 1955.) The IF bandwidth was reduced to 6 mc, thus improving sensitivity and reducing noise, and the frame-synchronizing signal was derived by integrating

rather than differentiating the rear flank of the synchronizing pulse, thus improving picture stability. Reception was possible every day, though the quality varied with time of day and meteorological conditons. The best propagation conditions appeared to correspond to stormy or cloudy weather with a fairly high ceiling, say 1000–1500 m. Local variations of signal strength and the usefulness of antenna pre-amplifiers are discussed.

### TRANSMISSION

**621.396.61:621.39:534.78** 286
A New System for Speech Transmission—Marcou and Daguet. (See 263.)

**621.396.61:621.376.3:621.396.969.3 :621.317.361** 287
The Calibration Problem in Frequency-Modulation Systems [for measurement of distance]—H. Familier and B. Ginger. (*Ann. Radioélect.*, vol. 11, pp. 118–124; April, 1856.) A system for determining the instantaneous frequency deviation of the transmitter is described. An auxiliary reference oscillator is amplitude-modulated by an oscillation whose frequency exceeds the deviation by an amount greater than the relative drift of the transmitter oscillator. The modulation side-bands constitute a pair of reference frequencies of known separation; these are mixed with the frequency-modulated signal. The timebase used to effect the fm also gates a counter circuit into which the mixture is fed, and a control voltage derived from the counter output is used to stabilize the deviation. The method is illustrated by examples. An absolute accuracy better than 1 in $10^3$ is attainable.

### TUBES AND THERMIONICS

**621.314.63+621.314.7** 288
Maximum Power of Semiconductor Junction Elements—J. P. Vasseur. (*Ann. Radioélect.*, vol. 11, pp. 3–28; January, 1956.) Universal curves are presented from which the maximum permissible power dissipation can be determined for Ge junctions cooled by 1) conduction or 2) convection. The influence of mounting, circuit arrangement, and signal waveform is discussed. The time taken to reach temperature equilibrium is investigated. Both diodes and transistors are considered.

**621.314.63** 289
High-Frequency Silicon-Aluminum Alloy Junction Diodes—M. B. Prince. IRE TRANS., vol. ED-2, pp. 8–9; October, 1955.) A preliminary account of experiments on diodes made by alloying the point of a thin Al wire to a wafer of $n$-type Si; the diameter of the alloyed region is probably <0.000 5 in. Rectification at frequencies up to 500 mc has been obtained. The reverse saturation current is about $10^{-9}$ A.

**621.314.63** 290
On the Transient Behaviour of Semiconductor Rectifiers—B. R. Gossick. (*J. Appl. Phys.*, vol. 27, pp. 905–911; August, 1956.) The investigation reported previously (1236 of 1956) is extended to cover large-amplitude transients in surface-barrier and point-contact diodes.

**621.314.63:546.28** 291
Reverse Current and Carrier Lifetime as a Function of Temperature in Silicon Junction Diodes—E. M. Pell and G. M. Roe. (*J. Appl. Phys.*, vol. 27, pp. 768–772; July, 1956.) Previous work on Ge diodes [3424 of 1955 (Pell)] is extended to Si grown-junction diodes; measurements over the temperature range $-190°$ to $+200°C$. are reported. The lifetime becomes constant over a low-temperature range; this is consistent with published recombination theory. The results indicate that charge generation from centres about 0.5 ev deep is responsible for most of the reverse current at temperatures up to well above room temperature.

**621.314.63:546.28                          292**

Silicon Wide *p-n* Junction—J. Nishizawa. (*Sci. Rep. Res. Inst. Tohoku Univ., Ser. B*, vol. 6, pp. 183–215; 1955.) The effect of the depletion layer in Si *p-n* junctions is investigated by determining the $V/I$ characteristics for various positions of the contact whisker, using a specially prepared junction having a very wide high-resistance layer.

**621.314.7                          293**

The Effect of Surface Treatments on Point-Contact Transistor Characteristics—J. H. Forster and L. E. Miller. (*Bell Syst. Tech. J.* vol. 35, pp. 767–811; July, 1956.) The properties of formed point contacts on Ge are discussed; technique for observing the equipotentials surrounding such contacts is described. Differences between donor-free and donor-doped contacts are emphasized. Differences are also observed between unformed point contacts subjected to different surface treatments. These considerations are applied to a study of transistor point-contact forming in contemporary practice. High yields from the forming process can be expected with oxidized surfaces; the use of chemical washes to remove soluble Ge-oxide surface films greatly reduces the forming yields.

**621.314.7                          294**

The Variation of Junction Transistor Current Amplification Factor with Emitter Current—N. H. Fletcher. (PROC. IRE, vol. 44, pp. 1475–1476; October, 1956.) A modification is suggested to the analysis presented by Webster (2798 of 1954) which brings the results into line with those of Rittner (3390 of 1954).

**621.314.7                          295**

Some Aspects of Transistor Progress—H. W. Loeb. (*J. Brit. IRE*, vol. 16, pp. 515–528; September, 1956.) Developments during the past seven years are surveyed. Transistor technology is outlined and advances in theory are illustrated by reference to the development of the equivalent circuit.

**621.314.7                          296**

An Analysis of Transistor Performance as a Function of Frequency and for Realistic Geometries—J. H. O. Harries. (*Quart. J. Mech. Appl. Math.*, vol. 9, pp. 212–223; June, 1956.) "The boundary shapes of actual fused impurity transistors are not such that analytical solutions of the governing differential equations can be found for the flow of the carriers. A relaxation method is described for solving these differential equations for realistic boundary shapes and as a function of frequency. The linear small signal theory of transistor operation is used."

**621.314.7                          297**

Inductive A.C. Admittance of Junction Transistor—M. Onoe and A. Ushirokawa. (PROC. IRE, vol. 44, p. 1475; October, 1956.) Experiments are briefly reported indicating that both the collector and the emitter admittances become inductive under certain conditions.

**621.314.7:(47)                          298**

Technical Data of New [Russian] Transistors—(*Radio, Moscow*, p. 55; June, 1956.) Data are given of junction and point-contact transistors for af and hf (up to 10 mc). For other types, see 3777 of 1955.

**621.314.7:621.3.015.3                          299**

Theory of Transient Processes in Semiconductor Triodes—E. I. Adirovich and V. G. Kolotilova. (*C.R. Acad. Sci. U.R.S.S.*, vol. 108, pp. 629–632; June, 1956. In Russian.) Expressions are derived for the transfer characteristics of a transistor operating in grounded-emitter and grounded-collector circuits. The

corresponding expression for the grounded-base circuit was given previously (1892 of 1956).

**621.314.7:621.375.4                          300**

Design of Circuits using Junction Transistors for High Frequencies—Vasseur. (See 65.)

**621.314.7:621.385                          301**

The Analogy between Vacuum Valves and Transistors—H. Beneking. (*Arch. Elekt. Übertragung*, vol. 10, pp. 214–221; May, 1956.) Continuation of previous work (2542 of 1954) on equivalent networks for junction transistors.

**621.314.7.012                          302**

Some Remarks on the Method of Presenting the Transistor Characteristics—Y. Watanabe and N. Honda. (*Sci. Rep. Res. Inst. Tohoku Univ., Ser. B*, vol. 6, pp. 63–98; 1954.) Equivalent circuits and various circuit parameters are derived from the $I_c/I_e$ characteristics for fixed values of $V_c$ and the $I_c/I_e$ characteristics for fixed $V_e$.

**621.314.7.012.8                          303**

A Contribution to the Mathematical Treatment and Measurement Technique of the Transistor as a Linear Quadripole—T. Scheler and H. W. Becke. (*Frequenz.*, vol. 10, pp. 107–116; April, 1956.) Matrix analysis is presented and an equivalent circuit is developed comprising a cascaded arrangement of a symmetrical T network, an ideal transformer and an "ideal" transistor which is presented as an ideal power amplifier. On the basis of this circuit, all the important transistor data can be calculated from five measurements, one of which serves as a check. Results obtained with a particular junction transistor are reported.

**621.314.7.012.8                          304**

A Transistor Simulator—J. Lüscher and P. Choquard. (*Tech. Mitt. schweiz. Telegr.-Teleph Verw.*, vol. 34, pp. 193–197; May 1, 1956.) Discussion of a complex equivalent network useful for investigations of the parameters and response of junction transistors.

**621.314.7.012.8                          305**

Equivalent Quadripole Networks of the Transistor—W. Taeger. (*Frequenz*, vol. 10, pp. 186–189; June, 1956.) Two equivalent networks are examined and equations, derived by matrix methods, are given for the transistor characteristics in grounded-base and grounded-emitter circuits. The characteristics of the Type-OC601 are evaluated on the basis of the manufacturer's provisional data.

**621.383.2                          306**

Electron-Optical Investigations of Composite Photocathodes—L. N. Bykhovskaya and Yu. M. Kushnir. (*Zh. Tekh. Fiz.*, vol. 25, pp. 2477–2485;) The investigation was carried out using a special tube incorporating the photocathode under test, together with a cylindrical anode and a fluorescent screen. The distribution integral and spectral sensitivities over the surface of SB-CS, Bi-Cs and Cs-O photocathodes were studied. The results are fully reported and discussed, and a number of electron images of the photocathodes are shown. Photoelectric fatigue was observed in caesium-oxygen cathodes.

**621.383.2                          307**

The Photoeffect of Antimony-Caesium Cathodes sensitized by Oxygen.—B. I. Dyatlovitskaya. (*Zh. Tekh. Fiz.*, vol. 25, pp. 2264–2276; November, 1955.) Experimental investigations indicate that the effect of the sensitization by oxygen is to lower the potential barrier without affecting the composition of the intermediate layer. Associated factors increasing the photosensitivity are discussed in detail, including the decrease in the external work function and the change in the emission depth of photoelectrons.

**621.383.2                          308**

The Effect of the Adsorption of Dipole Molecules of Barium Oxide on the Photoelec-Emission of an Antimony-Caesium Cathode—V. M. Gavrilyuk. (*Zh. Tekh. Fiz.*, vol. 25, pp. 2469–2476; December, 1955.) A method is proposed for regulating the work function of an Sb-Cs photocathode by adsorption of BaO on its surface. At a concentration of $2 \times 10^{14}$ BaO molecules/cm3 the work function decreased by about 0.1 ev, with a corresponding increase in the photocurrent by a factor of 2.5. A theoretical interpretation of the phenomena is given; the changes occuring in the photocathode during the adsorption of BaO are in many respects similar to those due to oxygen sensitization, as investigated by Dyatlovitskaya (307 above).

**621.383.27                          309**

The Development of Unfocused Secondary-Electron Multipliers—F. Eckart. (*Z. Angew. Phys.*, vol. 8, pp. 303–312; June, 1956.) Photomultipliers with grid- and louvre-type dynodes are considered. Five-fold multiplication per stage can be attained with a voltage of 150 v using AlMg louvre-type dynodes; this high factor is partly due to the simplicity of the construction, which facilitates surveillance during preparation. Resolving powers of the same order as those of focused multipliers are attainable. 53 references.

**621.383.27                          310**

Secondary-Electron Multipliers—G. H. Hille. (*Elektronik*, vol. 5, pp. 117–123; May, 1956.) The properties of a number of commercial photomultiplier cells are described and tabulated.

**621.383.4                          311**

The Lead Sulphide Photoconductive Cell M. Smollett and J. A. Jenkins. (*Electronic Eng.* vol. 28, pp. 373–375; September, 1956.) The principles and performace of PbS photocells are described, and some applications are indicated.

**621.383.5                          312**

The Hamaker-Geezhold Effect—G. Blet and A. Ritti. (*Rev. d'Optique*, vol. 35, pp. 193–214; April, 1956.) A report is presented of a detailed study of current build-up and decay in Se barrier-layer photocells. The results support theory based on trapping of photoelectrons. The trap density is estimated at $10^{10}$ per mm² of surface. Various methods of eliminating the time-lag effects have been tried; modulating the light beam is the only simple means found to be effective.

**621.385.832:621.397.62                          313**

The Television Picture Tube—H. Rothe and E. Gundert. (*Arch. Elekt. Übertragung*, vol. 10, pp. 188–194; May, 1956.) Modern cathode-ray pictures tubes are reviewed with particular attention to factors affecting spot modulation and resolution.

**621.314.7:537.311.33                          314**

Progress in Semiconductors. [Book Review] Gibson, Burgess, and Aigrain (Eds). (See 215.)

## MISCELLANEOUS

**621.37/.38].004.5/.6                          315**

Reliability of Military Electronic Equipment—L. M. Clement. (*J. Brit. IRE* vol. 16, pp. 488–495; September, 1956.) "Reliability is defined and the history of U.S. reliability programs discussed from the point of view of liaison between Government (user) and industry. Examples are given of reliable equipments and their success is shown to be due to the attitude of the management responsible for their production. Steps for ensuring the design and production of reliable equipment are described."