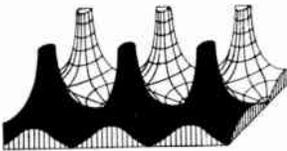


# Proceedings of the IRE



## Poles and Zeros



**DX and Progress.** On March 6, Pioneer IV, our first artificial asteroid established what appeared as an all-time dx record

by transmitting its last usable signal at a distance of 410,000 miles toward the sun. Soon thereafter we learned of the earlier praiseworthy work done in February by the Lincoln Laboratory group, in aiming a radar signal at Venus and receiving the return at a round trip distance of 56,000,000 miles.

We recall a day when even terrestrial DX was an event, and well remember sitting in while our local DX achiever contacted Brazil using a 199 tube and 0.2 watt plate input for a phenomenal record of 30,000 miles per watt. Being in the 199 days (a type now so obsolete it is not even on the Not Recommended list) it was also much before we had uncovered Hartley's law, although we knew well the Hartley oscillator, when we knew of bands but not bandwidth, and utilized antennas designed and placed according to the limits of the real estate instead of the requirements of impedance, gain, and directivity. And then the receiving equipment used therewith was a TRF, blooper detector, and one step with Baldwin cans (if you really were in the bucks).

With such a background, and many of us have it or can exceed it in historical interest, it helps our egos to note that with 265 kilowatts announced power at 28,000,000 miles one way to Venus, the Venus experiment succeeded in reaching the figure of 105 miles per watt! The little old 199, pulsed with good honest ham CW, can apparently still claim the title, if not surpassed by modern transistors. We admit, of course, that it is harder to hit a rabbit with a rifle than with a shotgun, and the high-power radar at Millstone Hill, Massachusetts, is a rifle and Venus is no rabbit, being more of a sitting duck!

It probably should be overlooked in passing, but we like figures well enough to bring out the fact that the system loss approximated 150 db. The system transmitted a pulsed signal for 4.5 minutes, and it took several months to statistically determine that a return signal had been received. Such a signal represents essentially one bit of information (did or did not), so that we suspect some sort of record may have been set in terms of channel capacity utilized per bit.

All the above is but prologue to our main thesis upon which we set out in writing this column—that in the field of electronics engineering, application no longer trails basic research by many years. In fact it would appear that the large and diversified basic research, plus the close relations and inseparable nature of much of our research and engineering development has had available the methods, systems, or equip-

ment when the need arose, or that invention must no longer wait upon necessity. Cannot we even say today that when necessity arrives, invention is here to meet it?

Take our recent timid ventures into space—could these have even passed the stage of science fiction without the computer for solving navigational problems, without the transistor and printed circuits for miniaturization, without the radar and radio astronomy antennas for communications? Was the computer developed after its need in a space program became apparent? Was the transistor a result of a program to reduce the size and power input of communication devices for satellites? Were our massive controllable antennas designed and built to track a midget aluminum sphere at an insignificant distance in space?

The answers are obviously negative; the computer was a result of a desire of thinking men to solve more problems faster, the transistor a stop along the way in a major attempt to understand better the nature of electrical materials, and our antennas were a product of basic and pure research into the far reaches of interstellar space. Yet when needed they were ready.

Our conventions, exhibits, and symposia give some evidence that we may be stockpiling other developments for which present application is unclear, toward a day when another new challenge arises—perhaps direct electrical conversion of fusion energy? Can we hope through intensive and close coordination of engineering with research, to provide the tools before the need and thus more quickly master nature for mankind's benefit?

**We Grow Again.** Once more we take this opportunity to welcome new IRE Sections: the Benelux Section of Belgium, the Netherlands, and Luxembourg; the Milan, Italy, Section; and that in Orlando, Florida, U.S.A.

This action brings to 103 our total of Sections, of which eight are located outside the United States and Canada. The old timer in this group is Buenos Aires, established in 1939; the others include Israel, Egypt, Tokyo, Rio de Janeiro, Colombia, Benelux, and Milan which have been founded since 1954.

We greeted our most recent foreign section, that in Quebec, in French in these columns, but because of the diversity of languages represented in our new Sections we are going to ask their indulgence for a greeting in English.

We are indeed glad to see this increased interest in international electronics. Good luck, Benelux, Milan, and Orlando.

—J.D.R.



## *Ferdinand Hamburger, Jr.*

*Director at Large, 1959–1961*

Ferdinand Hamburger, Jr. (A'32-M'39-SM'43-F'53) was born in Baltimore, Maryland, on July 5, 1904. He was educated in the public school system of that city, and received the degrees of Bachelor of Engineering in 1924 and Doctor of Engineering in 1931 from The Johns Hopkins University, where he was a Charles A. Coffin Fellow in 1930–1931.

During the period between the undergraduate and graduate degrees he participated in a program of dielectric research at The Johns Hopkins University. In 1931 he was appointed instructor in electrical engineering; in 1947, professor of electrical engineering; in 1954, chairman of the department of electrical engineering; and in addition, in 1958, director of the Radiation Laboratory of The Johns Hopkins University. He served as chief test engineer for Bendix Radio Division from 1942 to 1945 while on partial leave of absence from the University. He has acted as consultant for the Research and Standards Section, Bureau of Ships, Navy Department; United States District Court;

and The Radio Corporation of America; and others. He has served as research contract director of a number of research investigations supported by the Department of Defense at The Johns Hopkins University.

Dr. Hamburger joined the IRE in 1932 and was made a Fellow in 1953. He was Regional Director of the Central Atlantic Region in 1950–1951 and has served on the Nominations, Appointments, Policy Advisory, and Education Committees. He has been a member of the Editorial Board since 1956 and was vice-chairman in 1958. He was IRE representative at The Johns Hopkins University from 1941 to 1955. He was largely responsible for the formation of the Baltimore Section in 1939 and its reorganization in 1944; he served as its chairman in 1940–1941.

He is a Fellow of the American Institute of Electrical Engineers and is presently serving as vice-chairman of its Instrumentation Division. He is a member of the Society of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu and is a registered Professional Engineer in the State of Maryland.

## Scanning the Issue

**IRE Enters Space** (Berkner, p. 1048)—On March 25 at the annual IRE banquet, over 1000 members and guests sat spellbound while a distinguished radio physicist described the enormous opportunities which the space age has placed in the hands of the electronics profession. His speech ranged over many topics—astronomy, physics, relativity, meteorology, communications, international planning—but always with electronics in the central role. He discussed in particular the possibility of using satellites, hovering some 22,000 miles above the equator, as microwave relays for global communications, providing a timely sequel to the description of passive satellites which appeared in these pages just two months ago. To recapture for its readers the full impact of this memorable speed, PROCEEDINGS has discarded its traditional formality in favor of publishing a tape recorded transcript of the talk in its original, eloquently informal style.

**Relativity and Space Travel** (Pierce, p. 1053)—Lunik and Pioneer IV have brought man to the threshold of interplanetary travel at speeds that are a significant fraction of the speed of light. In escaping from his planet, man will also be leaving his familiar Newtonian world and entering a relativistic one in which time, distance and mass are no longer absolute and invariable nor everywhere the same, but instead depend on the relative velocities of the observer and the observed. Since the electronics engineer will be intimately concerned with man's exploits in space, it is highly desirable that he have some understanding of the subject of relativity as it applies to problems of space flight and that he be aware of the pitfalls of reasoning, such as the famous clock paradox, into which others have repeatedly fallen. This understanding is now clearly provided by a unique and highly readable paper in which readers will find the answers to the following interesting questions: 1) Since a clock runs slower at velocities near the speed of light, does this mean that if one twin took a long space voyage at very high speed he would be several years younger than his earthbound twin when he returned? 2) Within the next year or two a very precise clock will be placed in a satellite and compared with one on earth. What do we expect this comparison will show? Why should a 2000-mile orbit be avoided? 3) Photon rockets have been suggested for attaining a speed very close to that of light. Will they really go that fast? 4) What about using space ships propelled by the energy of interstellar matter? How fast would they travel?

**A Maser Amplifier for Radio Astronomy at X-Band** (Giordmaine, *et al.*, p. 1062)—This paper reports one of the first major applications of the maser. Using a three-level ruby maser as a preamplifier, the U. S. Naval Research Laboratory has been studying microwave radiation from Venus, Jupiter and other radio sources since April of last year, with a resulting twelvefold improvement in sensitivity. The effective input noise temperature, including background radiation into the antenna, was found to be 85°K. This work demonstrates not only that the maser has substantially improved the performance of radio telescopes, but that it should be possible to reduce noise temperatures to nearly 10°K, resulting in even more dramatic improvements in the future.

**Tantalum Printed Capacitors** (Berry and Sloan, p. 1070)—The authors have found a way of applying a metal electrode, instead of an electrolyte, directly to a tantalum oxide film without creating short circuits or low breakdown strengths. This represents an important step forward because it makes possible the manufacture of the highly regarded tantalum capacitor by printed circuit techniques—a development of substantial significance at a time when microminiaturization is of such widespread interest.

**Linear Diversity Combining Techniques** (Brennan, p.

1075)—It is unlikely that two signals, separated in space, frequency, time, or polarization, would fade together. This fact has provided communications engineers with a powerful tool for overcoming the effects of fading, by using two or more separated transmission channels. This paper is concerned with current techniques for selecting or combining the two or more signals after they are received so as to improve the quality or reliability of reception. The author explains the fundamental concepts involved, clarifies earlier work on the subject, presents important new results, and provides extremely useful comparisons of the three most promising signal-combining techniques now in use. This much-needed work is the most comprehensive, the most unified, and the most clearly written treatment of the subject that has yet been published. It appears with the blessings of the PGCS.

**Physical Principles of Avalanche Transistor Pulse Circuits** (Hamilton *et al.*, p. 1102)—This paper presents for the first time a simple and straightforward theory which explains, to a good degree of accuracy, the operation of a transistor in the avalanche region. Since avalanche transistors have recently been found to provide a new and simple way of generating millimicrosecond pulses, this development will be of importance to those who are working in high-frequency pulse circuitry as well as those in the field of semiconductor devices.

**Pulse Amplification Using Impact Ionization in Germanium** (Steele, *et al.*, p. 1109)—Millimicrosecond pulses and avalanche phenomena, described theoretically in the preceding paper, are the subject of important new experimental work in this paper. Moreover, the avalanche process in this instance is produced by means of a new physical phenomenon found to occur when impurity-type semiconductors are subjected to extremely low temperatures (about 4°K), involving the rapid ionization of neutral impurity atoms under the influence of small electric fields. The pulse amplifier which the authors describe is believed to be the first solid-state device which operates on this principle.

**The Transpolarizer: An Electrostatically Controlled Circuit Impedance with Stored Setting** (Pulvari, p. 1117)—Recent improvements in ferroelectric materials have made it feasible for the author to devise a novel storage, switching and control element, consisting of two or more ferroelectric dielectrics connected in series. In addition to having many of the properties and functions of magnetic devices such as the transfluxor, the transpolarizer shows promise of finding use in solid-state display devices.

**Geometric-Analytic Theory of Transition in Electrical Engineering** (Bolinder, p. 1124)—The author shows how some well-known concepts in modern geometry may be applied to plotting the trajectories of circuit parameters in the complex plane, affording an interesting approach to frequently occurring circuit theory problems.

**A Phenomenological Theory of the Reggia-Spencer Phase Shifter** (Weiss, p. 1130)—This paper looks into the interesting question of why the behavior of a novel ferrite phase shifter, described in these pages a year and a half ago, differs so radically from the familiar Faraday rotation devices. The simple terms in which the analysis is carried forward not only aids in understanding this device, but may eventually lead to producing other novel effects in ferrites.

**A New Look at the Phase-Locked Oscillator** (McAlear, p. 1137)—This paper is aptly titled. It deals with a versatile device that has been looked at many times in the past, but never with such a freshness of viewpoint nor in such understandable terms. Little more could be added here that is not already clearly explained by the author.

Scanning the Transactions appears on page 1164.

# IRE Enters Space\*

L. V. BERKNER†, FELLOW, IRE

**Summary**—Man's escape from the confines of his planet offers him revolutionary opportunities for performing whole new ranges of scientific experiments, notably in such fields as astronomy, physics and geophysics. Electronics, because it provides the vital nerve system for such experiments, will be at the very center of these new exploits in space. Moreover, earth satellites, possibly in a 24-hour equatorial orbit, promise to open a new era in global communications in which almost limitless bandwidths may become available at relatively low cost.

Space will become a major part of the activities of the IRE and its members in the future. Already there is developing a need for international planning of space-age communications standards and of regulations for utilizing satellites for man's greatest benefit.

## INTRODUCTION

IT WAS just forty-seven years ago, in 1912—within the span of most of our lifetimes—that Jack Hogan, Robert Marriott, and Dr. Alfred N. Goldsmith came together here in New York to form the Institute of Radio Engineers. In this short interval have come the great strides in technology that electronics has wrought. Indeed, we in the IRE have the opportunity to put our nose in almost everyone's business, because we supply the nerve system for almost every technology of which man can dream.

In the development of this vast technology, our Institute has played a very mighty part. The skill and the vision that has been required to organize this Institute so that the full power and strength of technology in electronics could bear on the problems of our country, has arisen from the vision and the dedication of the men that have led the IRE.

Among these, of course, we could cite the thousands of men who have worked on the committees of standards, who have worked on interengineering and industry committees such as JTAC, and who have organized and led our Professional Groups. But we must mention particularly our past Presidents who have carried the real responsibility for the organization and policy of this now greatest technical, or should I say quasi-scientific, society in the world. Among these Presidents I would particularly mention one man who can't be here tonight, and I am very sorry that he can't. His great leadership has led to the growth of the Professional Groups—those Groups which really carry much of the vitality of the IRE, which continue to make it grow in a solid and healthy technical fashion. I refer to Dr. W. R. G. Baker, who has long been a friend to all of us here in the IRE and who is also our Treasurer.

\* Original manuscript received by the IRE, April 15, 1959. This address is transcribed from tape in its original informal style, as given ad lib before the Annual IRE Banquet, New York, N. Y., March 25, 1959.

† President, Associated Universities, Inc., New York 19, N. Y.

Nor, of course, can we forget the people who have to do the day-by-day work, such as supervising publication of our wonderful journal, and the Editors of the PROCEEDINGS, and finally the men who run the basic organization underlying all of this—George Bailey, our great Executive Secretary, and Larry Cumming, our Technical Secretary.

This is the background that has prepared the Institute of Radio Engineers to enter into the space age that started scarcely more than a year ago. It was only on March 17th in 1958 that the first Vanguard satellite went up, the third of all of the satellites to be erected. It is still transmitting on one hundred-eight megacycles, using solar batteries to power it, and one suspects that our grandchildren and perhaps even our great-grandchildren will be listening on that frequency to wonder a little about the exploits of our present generation. Tonight I will try to demonstrate to you that the future operations in space will become an important and even a major part of the interests of this great Institute in the future.

Up to the year 1957 we were poor two-dimensional creatures who had to live on the surface of the earth. We couldn't escape above it, we couldn't get down very deep below its surface; we were confined like the ham in a sandwich and could only dream about what might happen if we could get to the outside. Now, we have gotten to the outside. We have seen for the first time the whole of the universe in the full range of its color. We can observe from the longest to the shortest wavelengths, from the highest to the lowest particle energies.

This escape is a bit reminiscent of the old story written by Mr. Abbott, the English schoolmaster, (I believe it was in the year 1848) when he described a two-dimensional world that he called "Flatland." In this two-dimensional world the people were two-dimensional geometric figures. Men could be triangles or squares or pentagons or hexagons, and if they were indeed very bright, they acquired so many sides that they finally became circles; whereas the women, as I recall it, were designated by straight lines (or really very narrow parallelograms). The story goes that one of the citizens of Flatland, a Hexagon, finally succeeded in escaping into three-dimensional space, and the story describes his wonder at the phenomena that he finds in three dimensions. Of course, Flatland was really a satire on the inability of men of the time to understand four dimensions. But it has unusual application in this day when we ourselves have just escaped into this three-dimensional world. Now, what is space going to mean to man in the future?

## SCIENTIFIC ASPECTS OF SPACE

Last night at the Symposium on Space you heard the discussions in which we mentioned that space science would be an important aspect of space explorations. But there will also be commercial and military applications. It would repay us to look back just for a moment over the potentialities of science that can be done in space. Indeed, the important point is that we can now do scientific experiments that would be quite beyond our range of capability without this opportunity to escape beyond our atmosphere and on towards the planets, and perhaps eventually even toward some of the stars.

The first science that comes to mind is the field of astronomy. Here one can imagine looking at the sun, our nearest star, not only in the ordinary spectrum, but down into the ultraviolet, the X radiation, to really understand the physics of this variable body that gives us our energy, our light and heat. Already this is being done, and some of our distinguished members have published extraordinary papers in the PROCEEDINGS describing our first views of the sun as we see it from beyond our atmosphere—a truly revolutionary opportunity.

But, going on to the stars, one can anticipate that we will learn much more of matter and its origin as we study astrophysics with telescopes that can extend over the whole range of color. Already in radio astronomy we are planning to look at the very long wavelengths that the ionosphere obscures from us, and at the much shorter wavelengths extending from a centimeter or so right on down to the infrared. There will be a real challenge to the members of the Institute to produce methods of detection in this intermediate range of wavelengths that will permit observation in the range one centimeter to, say, fifteen  $\mu$ . Then, finally, astronomers will need new telescopes that can provide good definition in the ultraviolet and X-ray regions of radiation.

Aside from the fun of these observations, one can pretty safely predict that our new knowledge of the physics of the universe, of its origin, its creation, its destruction, will be extended immensely by this new panorama that we have of space.

Then one can turn to the field of physics. "Well," you say, "what problems are there in physics that you can solve?" Let us take the case of general relativity. You say, "But hasn't general relativity been proven from the studies of the regression of the perihelion of Mercury?" The answer is that physicists would have said so until a year or two ago, but ever since parity as a general concept has tumbled down the drain, everyone is prepared to question again very seriously the full generality of some of the most fundamental concepts of our physics. Our operations in space permit us to make a very clear and unambiguous test of the general theory of relativity by comparing a very precise atomic

clock in a satellite with an identical clock on the earth. This experiment will be done in the next year or two. But beyond this, space science provides the opportunity to study fields, both gravitational and magnetic, on a scale involving new orders of distance in space. So, this escape into space really gives us a new dimension to carry on research in some of the most fundamental phenomena that occur in the field of physics.

In the field of geophysics, the advantages of space science are obvious. Last night you heard two distinguished physicists tell you of the studies that we are starting in the outer atmosphere, and beyond it in the region that lies between the outer atmosphere and the sun. I suspect that one of the most immediate rewards from studies in space will come to the field of meteorology. Remember that now our meteorologists can observe only about twenty per cent of the earth's surface—the rest is covered by oceans, the poles, and other inaccessible regions. However, with the satellite that we launched only last month, for the first time we can see and measure the cloud cover of the earth. It is hard to believe that up to the present time we have not known what proportion of the earth's surface is covered with cloud—we are just on the threshold of acquiring this information. We all know that the atmosphere of the earth behaves like the gas in a great heat engine that receives its heat generally in the vicinity of the equator and loses its heat near the poles. But changing cloud cover can change the heat balance—the input and output of energy of the heat engine—through reflection of solar radiation from cloud tops and absorption of terrestrial radiations in atmospheric gas dipoles. Soon we will know the distribution of the heat balance everywhere over the earth, so that we will have some opportunity to compute and to understand the performance of this heat engine which drives the gas of our atmosphere that gives us our weather.

Then, looking farther on, perhaps ten years or more, we begin to think of exploration of the moon and the planets. Among these I would mention Mars, Venus, Mercury, and perhaps the moons of Jupiter. It may be that ten, fifteen, twenty, or perhaps as long as fifty years hence, we will be sitting down to an IRE banquet with the spice on our meat from Mars, and I won't suggest what might come from Venus.

## COMMUNICATION BY SATELLITES

In this space activity certainly electronics plays the key and central part. From the very experiment itself, in which electronics must always be the vital nerve center, to the storage where the information acquired must be stored for subsequent transmission back, through the telemetry system, through the beacons that locate the vehicle, through the systems of control—in every element of space activity our electronics physicists and engineers are involved in the game.

But in spite of the great importance of these general

aspects of space activity, we must turn to the basic question of space communications that will affect the interests of IRE most of all. Here we find a variety of applications that open Pandora's box—the basic opportunity for communication channels over long distances, perhaps a thousand, ten thousand, or even a million times beyond the opportunities that we have had heretofore. First of all, remember that when we transmit to and receive from satellites, we finally achieve the engineer's dream of having  $1/r$  in the equation without having any exponent—for the exponent  $e^0 = 1$ . And I'll tell you that getting rid of this exponent does wonders for the engineer's good nature.

One can think of a variety of ways in which transmission and reception from satellites can be carried out to give us immense bandwidths for long-distance communications. Already you have read in the PROCEEDINGS, by John Pierce and others, the descriptions of how simple reflecting bodies sent into space—perhaps a hundred-foot sphere, perhaps a great corner reflector—can permit the relay of information by reflection from the satellite back to earth. Certainly these kinds of communications will be accomplished in the very near future—probably before this year is out. But, if you let your mind run just a bit further, you will realize that the real potentialities of communications from satellites lie in the receivers and amplifiers and transmitters that can be put in the satellites themselves. In this way, they provide you the opportunity to control very precisely as you choose, the kinds of communications you want to relay and the quality of those communications.

Someone said the other day, "Well, I don't suppose it will be worthwhile sending up a satellite for communications unless you have at least a two hundred megacycle bandwidth, or perhaps a five hundred megacycle bandwidth," and this is probably true. When you start to calculate power, you will say, "Oh, my goodness, we will never be able to do this with the power required to handle these wide bandwidths." But, with a little thought, you will realize that indeed you don't want very much power in the satellite; indeed a few milliwatts per kilocycle will do not only a very adequate job, it will also permit a much wider application of satellite communications than would an overpowering amount of output power. Indeed, one comes to the conclusion, with a little further study, that unless you keep the power down rather carefully, you can easily destroy the situation that you most want to acquire, and that is to realize cheaply the maximum number of transmission bands.

"Well," you will say, "what orbits must these satellites have in order to provide all of this communication? Should they be going around the equator, or around the poles?" The answer is, of course, one can select orbits to meet the requirements of a variety of situations. But, in thinking about the orbit, you pretty soon come to the conclusion that perhaps the most valuable orbit,

the one likely to be most important, would be the so-called twenty-four hour orbit. Now, this is a very nifty satellite orbit—it is one in which the satellite is fired off to the eastward from the equator, or in fact from any other point and then is stabilized on an equatorial orbit. It is stabilized on the eastward orbit at just the right velocity, so that the satellite rotates around the earth just as fast as the earth itself turns. Then, of course, the satellite just hovers above a selected point above the equator or, if you haven't stabilized it on the equator, just oscillates slowly back and forth each twenty-four hours along a portion of the meridian, but always in your own plain view and control. Indeed, we should be able to launch a twenty-four hour satellite with suitable control in the very near future.

Now, when you have an object up there twenty-two thousand seven hundred miles (I believe that is what the orbital height comes out to be), and you control it pretty much over your head all the time, you can see that this can be a very useful device indeed. But, you say, "Well, after all, you put one of these things up there and then you have your two-hundred megacycles or your five-hundred megacycles, isn't that it?" But the answer, of course, is not as simple as that, because these electronics engineers are always clever fellows. They say, "Oh, but we would like to put some antennas around which are highly directive," and then instead of having one of these hovering satellites, we will have a hundred, perhaps a thousand, each one with a nice antenna looking at it. You can move your antennas a little bit, and because you know just where these satellites are, you can pick out the satellite you want, to bring in your connections with London or Moscow or Paris or whatever city you like. Then there will be enough megacycles bandwidth so each of us can have some.

I have described this in just a little detail to challenge your imagination with respect to the immense potentialities involved. Of course, the matter looks quite easy until you start to dig into the—you know—the nasty details. Then you discover that there are problems of position control and reliability, and above all, problems of terminal equipment on the ground. As far as power is concerned, it is beginning to look pretty good. Our solar cells seem to be doing all right, but more recently we hear a good deal of talk about the gas thermocouples (plasma thermocouples if you want to sound scientific), operated from small amounts of nuclear energy. It looks as though you would be able to supply all the satellite power you would want on a vehicle for a long time. So, one concludes that we are on the threshold of having made available to us at a relatively small cost, perhaps at a lower cost than any other form of long-distance communication, almost limitless bandwidths for point-to-point, or even in some cases mobile, communications.

One is only a bit frightened about the terminal equip-

ment and about man's capability to use this potentially vast quantity of communications intelligently. However, at this Convention of the Institute we have learned that small solid-state circuits have come out that are barely larger than a pin-head. I believe I heard it said that five million such circuits could be put in a cubic foot. So you say, "After all, computers of any kind can be reasonably designed." Indeed, one is a bit startled to learn that computers, aside from ordinary programming, are now permitted to have judgment, and you can bring in whatever quantity or quality of judgment you want into the computer operation.

The problems of cost, of course, inevitably come into the picture, and the cost of launching large rockets is now very great indeed. But we must remember that in doing our initial rocket launching, we are a bit in the position of an aircraft producer that wanted to get a new aircraft. Every time he flew this aircraft it would crash, and from the crash he would have to determine the defects and then develop another aircraft that would do a little better. This is precisely the situation in the rocket business. In spite of all the static testing, one has to go through flight after flight during which your rocket crashes before you finally begin to get one that is useful and effective.

I do not think one can be very critical of the people who have done a magnificent job of this rocket development. They are now perfecting our rockets to the point where they are becoming relatively reliable through relatively few test flights. I am told that within a very short time a rocket that will put some four-hundred pounds into a suitable orbit can be launched for less than a half-million dollars. When you begin to calculate the kinds of communications that this will provide over a long period of time, you realize that the real cost per message unit lies not in the rocket and satellite itself, but in the terminal equipment on the ground at both ends.

#### INTERNATIONAL REGULATION

When we think about the problems of control, multiplying our communications by a factor of a thousand or ten thousand or more over the present, one cannot help but be impressed with the dreadful problems that face the International Telecommunications Union, the International Frequency Board, the CCIR and our FCC in the future. I imagine that the present meeting of CCIR out at Los Angeles is beginning to worry about this pretty seriously, because rumors of the capabilities of these devices are just beginning to be discussed rather widely.

The question before us, with these potentialities immediately ahead, is how soon should we proceed with the standards—with the planning—that underlies an advance in communications of this kind? To describe the thinking in some of the other countries, I am going to read to you from the minutes of the Technical Coun-

cil of the Ministry of Communications, Moscow, held in January, 1959, and I quote a part of these minutes:

S. I. Katayev pointed out that in the solution of a number of problems for the subsequent development of television technology, including the selection of the standards, it is necessary, even now, to take into account the possibility of utilizing the artificial Earth Satellite. It is also necessary to start to develop a series of concrete answers to problems, such as the transmitting and receiving radio equipment designed for several years' operation without servicing; the selection of an optimum wavelength for the cosmic retransmitter; automatic scanning of large antennas in cosmic space; sufficiently powerful, independent, long-duration power sources; investigation of the economic variants of color television designed to utilize the cosmic retransmitters; etc., etc.

It says further:

Professor Katayev proposed to include the basic measures for realization of a cosmic television program retransmitter in the Seven-Year Plan for the development of the economy of the U.S.S.R.<sup>1</sup>

So, you see, that very serious thinking is going on abroad concerning the development of the means and standards required for the use of these immense potentialities for space communication in the future.

In the field of international planning, of course, we have the International Telecommunications Union and under it, all of its very excellent administrative and advisory bodies to deal with problems of frequency allocation. During the last session of the United Nations there was proposed a resolution which was adopted, but not unanimously, to have an international committee set up to study the general control, regulation and utilization of satellites for man's greatest benefit. Unfortunately, because of some disagreement over this, this committee has not yet been called together. But, I think all of you who have followed this problem would agree that it is rather urgent that man undertake early planning for this powerful new potentiality if it is to be used for his benefit and not for his detriment.

In the field of science, fortunately, all nations have joined in the scientific utilization of satellites under a committee called COSPAR, the International Committee on Space Research, organized and operating under the International Council of Scientific Unions. I am proud to tell you that at the last meeting of that Committee at the Hague, at which all important nations, including the U.S.S.R., were represented, the U.S. delegation was able to make a proposal which had been approved by our Government. This proposal invited scientists of nations outside of the United States, acting under COSPAR, to prepare selected experiments to be carried in U.S. satellites in the future. In my opinion this is a very wise decision on the part of our Government, since it means that scientists abroad whose science *must* be done in space in the future, will have access to that space through the good auspices of the United States. I believe that it is an offer of the greatest significance to our future.

<sup>1</sup> *Radiotekhnika*, Moscow, no. 1, p. 67; January, 1959.

## CONCLUSIONS

In closing, I think we cannot avoid considering in this space age the implications of the race between ideologies in the world at the present time. True, we are in a military race, but beyond this the U.S. is really in a race for intellectual superiority. For, assuming that through wisdom we can avoid in one way or another a disastrous military conflict, then the future of the Western system depends upon the ability of the West to demonstrate the superiority of that system. If you try to avoid the implications of this race, you can not avoid the knowledge that if the U.S. is not racing with the U.S.S.R., at least the U.S.S.R. is racing with the U.S. I would read to you the closing remarks of Academician A. N. Nesmeyanov, the President of the Soviet Academy of Sciences, on the occasion of the first anniversary of the launching of the Sputnik. Nesmeyanov says:

In marking the anniversary of the launching of the first artificial earth satellite, we may be confident that in the name of the great ideals of humanity our people, under the leadership of the Communist Party, will accomplish new and ever more notable feats.

Our guarantee of this is the socialistic structure of our country, which gives wide rein to the development of science and ensures the bringing-out of the notable talents of our people.<sup>2</sup>

In the face of that position, can you doubt the Russian desire to achieve through their system intellectual superiority?

We have one great advantage, and that is the immense freedom that is enjoyed by each one of our citi-

zens. This freedom challenges the individual, without being pressed by his government, to do his part in bringing the free society of men in which we live to a position of unquestioned leadership. For in a free system, it is the individual, not the government, that determines the competence of the system. In meeting this challenge, we in the Institute of Radio Engineers, who manage and control the nerve center of man's technology, are at the center of the affair. We cannot avoid it. It is something that is placed on our shoulders, and I do believe that this Institute, the men in it, the men who have created it, the men who will nurture its growth in the future, have the sense of responsibility, the spirit and the courage to meet this challenge—a duty that the free world delegates to them today.

I would finally say, with respect to the Institute itself, that rarely does a well-developed technology find an entirely new opportunity. But just such an opportunity, one that can multiply our communications a thousand times over, has been brought to us by space. Space promises cheap communications, international television, the opportunity for spreading good will if it is managed correctly. It promises a new order of acquisition and transmission of weather data, which may eventually lead to better forecasting and perhaps control or modification of our weather. Over the horizon are benefits which we can only imagine at the present time, for we are in the situation of Hexagon who has just escaped from two-dimensional space, and is still looking with wonder at the three-dimensional space around him.

<sup>2</sup> "A great victory of Soviet science," trans. by E. R. Hope from *Vestnik Akad. Nauk SSSR*, vol. 28, no. 11, pp. 3-9; 1959.

# Relativity and Space Travel\*

J. R. PIERCE†, FELLOW, IRE

**Summary**—This paper treats in terms of the special theory of relativity: a “clock paradox” involving the fact that the frequency of an atomic oscillator on a moving body is lowered but the mass which is converted into radiation is increased; the case of the twin who goes on a space trip at near-light speed and returns younger than his brother on earth; the shift in frequency in the presence of a gravitational field; the clock rate on a satellite; the speed attainable by a photon rocket; and a space ship propelled by the energy of interstellar matter.

## I. INTRODUCTION

TODAY engineers are building rockets and satellites with velocities unprecedented among gross, man-made objects. Further, they are seriously speculating concerning even higher velocities in connection with interplanetary travel. There have been repeated suggestions that a “photon rocket” could be used to attain a speed very close to that of light.

It is well known that the theory of Newtonian mechanics is inaccurate in dealing with velocities appreciable compared with the velocity of light. The solution to the problem of applying relativity consistently to the practical and speculative problems of space flight is less well known, and various questions and errors recur repeatedly.

Fortunately, it is possible to treat many interesting problems of space travel by means of special relativity, together with a few reasonable *ad hoc* assumptions. This is the purpose of the present paper. Portions of the material will be found in a number of sources.<sup>1-5</sup> A reasonably complete and unified account is attempted here.

Before we proceed to particular problems, a few general remarks may be helpful. In making use of relativity, it is important to avoid certain easy pitfalls. One of these is the notion of simultaneity. We make matters more difficult for ourselves, for instance, if in reckoning time on a space ship we talk about the time *on earth* at which the space ship turns around and heads toward earth instead of away from it. This can lead to a seeming “clock paradox” which will be discussed later.

Feynman has expounded the matter of simultaneity in a wonderfully simple manner. In pursuing his explanation we should note that, according to special rela-

tivity, an object going past us appears to be foreshortened in the direction of motion.

Imagine two very long starships, ships which to us are equally long, passing each other while traveling in opposite directions, as shown in Fig. 1. Suppose that just as they are opposite one another, two bolts of lightning pass instantly between them, one from the nose of the upper, leftward-traveling ship, which we will call ship *L*, to the tail of the lower, rightward-traveling ship, ship *R*, and the other from the tail of ship *R* to the nose of ship *L*. An observer in the center of the upper ship, ship *L*, will see the left-hand bolt first, because he tends to catch up with the light emitted by it, while he travels away from the light emitted by the bolt to the right. On the other hand, an observer in the center of ship *R* will see the light from the right-hand bolt first, for he rushes to meet the light from the right and flees the light from the left.

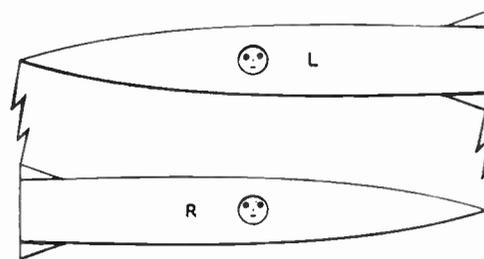


Fig. 1.

Special relativity tells us that the man in ship *L* has a perfect right to say, “I am standing still, and if I see the left-hand bolt of lightning, the one at my nose, before the right-hand bolt, the one at my tail, it is because the bolt at my nose occurred first. Indeed, it had to, for I see that that other ship, ship *R*, is much shorter than my ship, ship *L*, and hence his tail was opposite my nose before his nose reached the position of my tail.”

On the other hand, the man in ship *R* says, “I am standing still. I saw the right-hand bolt, the one at my nose, first, and so it occurred first. Indeed, it had to, for that ship *L* is much shorter than my ship *R*, and his tail was opposite my nose before his nose reached the position of my tail.”

Clearly, no argument can resolve this contradiction as to which bolt of lightning struck first, if, indeed, either did. An assumption of some instantaneous form of communication (telepathy, for instance) which enables us to determine a “true” time order of events would involve us in a hopeless paradox. We simply can’t have such a thing and relativity too: one or the other must go.

\* Original manuscript received by the IRE, April 1, 1959.

† Bell Telephone Labs., Inc., Murray Hill, N. J.

<sup>1</sup> J. H. Jeans and E. Whittaker, “Relativity,” in “The Encyclopedia Britannica,” vol. 19, pp. 89-98, 1955.

<sup>2</sup> J. J. Coupling, “On atomic jets,” *Astounding Science Fiction*, vol. 54, pp. 115-127; January, 1955.

<sup>3</sup> S. F. Singer, “Application of an artificial satellite to the measurement of the general relativity ‘red shift,’” *Phys. Rev.*, vol. 104, pp. 11-14; October 1, 1956.

<sup>4</sup> R. Schlegel, “New clock problems in special relativity,” *Bull. Am. Phys. Soc.*, vol. 2, p. 239; April 25, 1957.

<sup>5</sup> C. G. Darwin, “The clock paradox in relativity,” *Nature*, vol. 180, pp. 976-977; November 9, 1957.

As the relative velocities of the two ships approach the speed of light, one captain will insist that event *A* occurs a time  $L/c$  before event *B*, while the other insists that event *A* occurs at time  $L/c$  after event *B*, there will be a possible disagreement of  $2L/c$  about what is simultaneous with what, where  $L$  is the length that each assigns to his ship. In general, for events occurring a distance  $L$  apart, there is a range of time  $2L/c$  over which we cannot say whether one event occurred before or after another.

We can assert that event *B* at location *b* occurred after event *A* at location *a*, which is a distance  $L$  away only when a light signal generated by event *A* at *a* reaches *b* before event *B* occurs; this takes a time  $L/c$ . The same statement holds if we interchange *A* and *a* with *B* and *b*.

## 11. SOME EQUATIONS OF SPECIAL RELATIVITY

It is not the purpose of this paper to derive the equations of special relativity. The purpose is to explain them sufficiently so that they can be used correctly and to apply them to certain problems in space travel.

The equations we need deal with an observer who considers himself to be standing still and a space ship, planet, solar system or galaxy which is moving past him, from left to right, with a velocity  $v$ . As we noted earlier, to the observer, all objects in the moving system appear to be shrunk or contracted by a factor which we will call a shrinking factor  $S$ . This factor is

$$S = \sqrt{1 - (v/c)^2}. \quad (1)$$

Here  $c$  is the velocity of light.

$$c = 3 \times 10^8 \text{ meters/sec.} \quad (2)$$

The moving object appears shrunk only in the direction of motion, not crosswise to it.

The clocks on the moving object also appear to be going slower by this same shrinking factor  $S$ , so that when our clock indicates the passage of one hour, the clocks on the moving system indicate the passage of only a fraction  $S$  of an hour.

Let us now consider the people in the moving system. One of them shoots a projectile in the direction of motion and says, "I have shot this forward with a velocity  $u_m$ ." But, when we measure the velocity by our "fixed" standards, we obtain a velocity  $u$

$$u = \frac{u_m + v}{1 + vu_m/c^2}. \quad (3)$$

We note that the velocity  $u$  which we observe can never be greater than  $c$ , the velocity of light. Suppose, for instance, that a man in the moving system shines a beam of light to the right, a beam which he says has the speed of light, so that  $u_m = c$ . Both he and the beam are rushing to the right, he with a velocity  $v$ . Yet, if we put  $u_m = c$  into (3), we get for  $u$  simply  $c$ , the velocity of light. Eq. (3) will not be used in the subsequent work, but it is an important tool.

When we observe matter in motion with a velocity  $v$ , it appears to have more mass. Thus, something which to a man in the moving system appears to have a mass  $m_0$  (called the rest mass) appears to us, past whom it is moving with a velocity  $v$ , to have a mass  $m$  given by

$$m = \frac{m_0}{\sqrt{1 - (v/c)^2}}. \quad (4)$$

The rest mass is the quantity of matter; it is what remains constant when we accelerate a body, while the relativistic mass  $m$  increases.

We are all familiar with the relativistic expression for energy  $E$ ,

$$E = mc^2. \quad (5)$$

This is universally applicable. As a moving object has more mass than it would if it were standing still [according to (4)], (5) tells us that it will also have more energy. The additional energy is of course the kinetic energy of the body, that is, the energy associated with its motion.

In order to make calculations concerning space travel, we must associate with these laws of relativity two universal and basic laws of physics: the conservation of energy and the conservation of momentum. These laws say that the total energy and momentum must be the same before and after a physical event, such as the acceleration of a starship.

We can express the energy of the system in terms of its masses, at rest or moving, by means of relation (5). As the energy of each mass is proportional to the mass by the same constant,  $c^2$ , we see that the conservation of energy means the conservation of mass. *Rest mass* may be diminished, but if we use the energy produced to set matter in motion, the relativistic masses [as given by (4)] total the same as do the original masses. Or, electromagnetic radiation such as light may be produced, which, according to relativity, also has mass.

The momentum  $p$  of a material body of mass  $m$  and velocity  $v$  is

$$p = mv = \frac{m_0 v}{\sqrt{1 - (v/c)^2}}. \quad (6)$$

Momentum of course has a direction; this is the same direction as the velocity of the body.

According to relativity, electromagnetic radiation, including gamma rays, X rays, light and radio waves, must have mass and momentum as well as energy. Let us call the energy of a certain amount of radiation  $E_r$ , its mass  $m_r$  and its momentum  $p_r$ . The radiation travels with the velocity of light,  $c$ . The energy, mass and momentum are related by the following equations:

$$E_r = m_r c^2 \quad (7)$$

$$p_r = E_r/c = m_r c. \quad (8)$$

We see that (7) and (8) are really just the same as (5) and (6). However, radiation has no rest mass  $m_0$ .

We are now equipped with all the physical laws we need for our calculations, except a quantum law governing radiation, which we will encounter in the next section.

### III. A "CLOCK PARADOX"

In accord with the conservation of energy and the fact that energy is  $mc^2$ , when radiant energy such as light appears, matter must disappear. Thus, when an atom or molecule emits a quantum of light or other electromagnetic radiation, such as radio waves or gamma rays, it loses energy. Quantum mechanics tells us something more about this process. The energy of a quantum of light or other radiation is Planck's constant  $h$  times the frequency of the light,  $f$ . Hence, from the conservation of energy,

$$hf = m_r c^2 = E_r \tag{9}$$

or

$$f = m_r c^2 / h = E_r / h \tag{10}$$

$$h = 6.55 \times 10^{-34} \text{ joule/sec.} \tag{11}$$

Here  $m_r$  is the mass of the quantum of radiation and it is also equal to the mass lost by the atom or molecule in producing the radiation.  $E_r$  is the energy of the radiation.

Relativity tells us that the mass of each atom or molecule *increases* when the atom or molecule is traveling fast. Further, our best clocks are regulated by the frequencies of radiation of ammonia molecules or cesium atoms. If we took such a clock on a swiftly-moving space ship, the mass of each molecule or atom would increase, and hence the small fraction of matter which it would lose in emitting radiation would also have more mass. Thus, we might expect that the frequency of the electromagnetic wave produced by the atom or molecule would be higher than if the ship were standing still. Since this frequency governs the speed of our clocks, we might expect the clocks on the ship to go faster by a factor  $1/\sqrt{1-(v/c)^2}$ .

Relativity tells us, quite to the contrary, that the rate of the moving clocks is decreased by a factor

$$S = \sqrt{1 - (v/c)^2}. \tag{1}$$

How are we to explain this seeming paradox?

Let us explore this matter warily. As a first step, let us assume that we send waves of light or other radiation from earth to a moving space ship. What frequency will these waves seem to have to an observer on the space ship?

We will consider two cases, shown in Fig. 2. In (a), the ship is headed toward the waves of the radiation with a velocity  $v$ , and of course the waves of radiation go toward the ship with the velocity  $c$  and of light. If  $\lambda$  is the wavelength of the waves, that is, the distance between wave crests, then the frequency  $f$  of the waves, that is, the number of crests  $\lambda$  apart which will strike a fixed object in a second, is clearly

$$f = c/\lambda. \tag{12}$$

However, we see the ship and the waves as having a relative velocity  $c+v$ , so we see the wave crests as striking the ship with a frequency  $f_1$  given by

$$f_1 = \frac{c+v}{\lambda} = (1+v/c)f. \tag{13}$$

Is this the frequency at which the man on the ship observes the wave crests to arrive? No, for his clock runs slow by a factor  $\sqrt{1-(v/c)^2}$ . Thus, to him the crests seem to arrive more rapidly, with a frequency  $f_d$  given by

$$f_d = \frac{f_1}{\sqrt{1 - (v/c)^2}}. \tag{14}$$

By using the value of  $f_1$  given by (13) in (14) we find

$$f_d = \frac{\sqrt{1+v/c}}{\sqrt{1-v/c}} f. \tag{15}$$

This is known as the *relativistic Doppler frequency*.

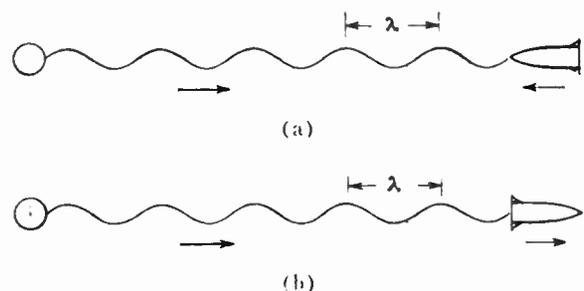


Fig. 2.

In the case shown in Fig. 2(b), in which the ship is going away from the source of radiation, we see the waves of radiation overtaking the ship with a relative velocity

$$c-v.$$

Accordingly, the frequency  $f$  with which we see crests strike the ship is

$$f_1 = (1 - v/c)f. \tag{16}$$

However, because the ship's clock runs slow, those on the ship observe a frequency

$$f_d = \frac{\sqrt{1 - v/c}}{\sqrt{1 + v/c}} f. \tag{17}$$

So far, we have stood still on earth and watched an observer on the moving ship measure the frequency of the radiation that we send him. But who is to say which is moving, the ship or the earth? Suppose that the captain of the ship watches an atom on the (to him) rapidly-moving earth emit radiation. The energy produced must be the change in relativistic mass, which is greater than the rest mass, times the square of the velocity of

light. Since the speed of the earth results in an increase in the mass, does this increase in mass directly account for the frequency of radiation emitted? No, it does not!

What we do know is that both energy and momentum must be conserved in the emission of radiation from the sources on earth. Let us assume that the radiation is emitted downward, at an angle  $\theta$  with respect to the normal to the path of the earth, which travels to the right with a velocity  $v$ , as shown in Fig. 3.

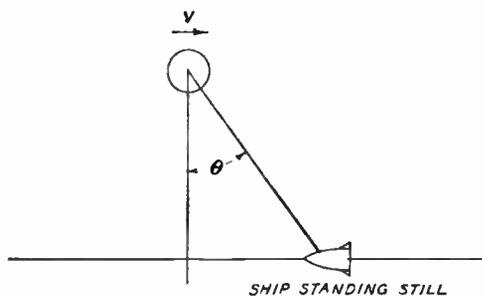


Fig. 3.

Let us consider the energy and momentum of an object traveling to the right with a velocity  $v$  which has  $x$  and  $y$  components  $v_x$  and  $v_y$ . The magnitude  $v$  of the velocity is

$$v = \sqrt{v_x^2 + v_y^2}. \tag{18}$$

The energy  $E$  will be

$$E = \frac{m_0 c^2}{\sqrt{1 - (v_x/c)^2 - (v_y/c)^2}}. \tag{19}$$

The  $x$  and  $y$  components of momentum,  $p_x$  and  $p_y$ , will be

$$p_x = \frac{m_0 v_x}{\sqrt{1 - (v_x/c)^2 - (v_y/c)^2}} \tag{20}$$

$$p_y = \frac{m_0 v_y}{\sqrt{1 - (v_x/c)^2 - (v_y/c)^2}}. \tag{21}$$

Let us suppose that the rest mass  $m_0$  is changed by a small amount  $dm_0$  and that the  $x$  and  $y$  components of velocity are changed by small amounts  $dv_x$  and  $dv_y$ . By differentiation, we find the changes in energy and momenta to be

$$dE = \frac{m_0(v_x dv_x + v_y dv_y)}{(1 - (v_x/c)^2 - (v_y/c)^2)^{3/2}} + \frac{c^2 dm_0}{\sqrt{1 - (v_x/c)^2 - (v_y/c)^2}} \tag{22}$$

$$dp_x = \frac{m_0(1 - (v_y/c)^2)dv_x}{(1 - (v_x/c)^2 - (v_y/c)^2)^{3/2}} + \frac{v_x dm_0}{\sqrt{1 - (v_x/c)^2 - (v_y/c)^2}} \tag{23}$$

$$dp_y = \frac{m_0(1 - (v_x/c)^2)dv_y}{(1 - (v_x/c)^2 - (v_y/c)^2)^{3/2}} + \frac{v_y dm_0}{\sqrt{1 - (v_x/c)^2 - (v_y/c)^2}}. \tag{24}$$

Let us now apply (22)–(24) to our case, in which the earth is whizzing by in the  $x$  direction with a velocity  $v$ . In this case

$$v_y = 0 \\ v_x = v. \tag{25}$$

Accordingly,

$$dE = \frac{m_0 v dv_x}{(1 - (v/c)^2)^{3/2}} + \frac{c^2 dm_0}{\sqrt{1 - (v/c)^2}} \tag{26}$$

$$dp_x = \frac{m_0 v dv_x}{(1 - (v/c)^2)^{3/2}} + \frac{v dm_0}{\sqrt{1 - (v/c)^2}} \tag{27}$$

$$dp_y = \frac{m_0 dv_y}{\sqrt{1 - (v/c)^2}}. \tag{28}$$

Now let  $E_r$  be the energy of the emitted radiation. The earth, in emitting the radiation, must lose an energy equal to the energy of the radiation, so that we must have

$$dE = -E_r. \tag{29}$$

Further, in order for momentum to be conserved, we must have

$$dp_x = -(E_r/c) \sin \theta \tag{30}$$

$$dp_y = (E_r/c) \cos \theta. \tag{31}$$

Let us further call the rest mass which a molecule or atom on earth loses in emitting radiation  $m_r$ . Thus

$$dm_0 = -m_r. \tag{32}$$

By substituting (29)–(32) into (26)–(28), we obtain

$$E_r = \frac{-m_0 v dv_x}{(1 - (v/c)^2)^{3/2}} + \frac{m_r c^2}{\sqrt{1 - (v/c)^2}} \tag{33}$$

$$(E_r/c) \sin \theta = \frac{-m_0 dv_x}{(1 - (v/c)^2)^{3/2}} + \frac{m_r v}{\sqrt{1 - (v/c)^2}} \tag{34}$$

$$(E_r/c) \cos \theta = \frac{m_0 dv_y}{\sqrt{1 - (v/c)^2}}. \tag{35}$$

We note that  $dv_y$  appears in (35) only. This equation can be regarded as giving  $dv_y$  directly in terms of  $E_r$ . We can eliminate  $dv_x$  by using (33) and (34); we obtain

$$E_r = \frac{\sqrt{1 - (v/c)^2}}{1 - (v/c) \sin \theta} m_r c^2. \tag{36}$$

Here  $m_r c^2$  is the energy of a photon as seen from earth, while  $E_r$  is the energy of the photon according to an observer who says that the earth is moving past him with a velocity  $v$ , as shown in Fig. 3. As frequency is

proportional to energy, then if  $f$  is the frequency of a photon as observed on earth and  $f_d$  as observed by our "fixed" observer, (36) tells us that

$$f_d = \frac{\sqrt{1 - (v/c)^2}}{1 - (v/c) \sin \theta} f. \quad (37)$$

When  $\theta = \pi/2$ , so that  $\sin \theta = 1$ , we have the case of earth and ship approaching, to which (15) applies, and we are gratified to see that (37) indeed reduces to (15). When  $\theta = -\pi/2$ , so that  $\sin \theta = -1$ , we have the case of earth and ship separating, and indeed, (37) reduces to (17), which applies to this case.

When  $\theta = 0$ , so that  $\sin \theta = 0$ , the earth is neither approaching nor receding from the ship. Here we observe the oscillator frequency of the ship, unaffected by relative motion; we see from (37) that it is

$$\sqrt{1 - (v/c)^2} f. \quad (38)$$

This is exactly in accord with the shrinking factor  $S$  of (1).

At first thought, it may seem puzzling that while the relativistic mass of the matter lost during radiation of the quantum is  $m_r/\sqrt{1 - (v/c)^2}$ , the mass of the radiation produced is only  $m_r\sqrt{1 - (v/c)^2}$ . We note, however, that for momentum to be conserved in the  $x$  direction after the earth has lost rest mass, the earth must travel faster, and this means that the rest of the earth must gain energy during the process of radiation.

We may also note that radiation which to the fixed observer appears to be traveling normal to the  $x$  axis must appear to an observer on the "moving" earth to have a component of motion in the  $-x$  direction. In illustration, we may note that a swiftly flying airplane would have to eject a bomb violently backward in order for it to fall straight down to a man on earth.

#### IV. THE CLOCK PARADOX

We have just considered a clock paradox. This is not, however, the clock paradox that most people mean when they use the words. In the clock paradox, one of two identical twins goes off in a starship at 99.5 per cent of the speed of light. As observed from earth, his clock runs only 1/10 normal speed. After five years pass on earth, the ship turns around, and 10 years from takeoff he lands on earth. The twin steps out, but he has aged only 1 year, while his brother who stayed on earth is 10 years older than when the ship started.

However, it has been argued that to the twin on the ship, the clock on earth should have appeared to be running at 1/10 normal speed. Should we say that the twin on earth ages only one year while the twin on the ship ages 10 years?

Such arguments are confusing because they involve one in considerations of the time on distant earth when the ship turns around, or the time on the ship as reck-

oned from earth when the ship turns around. As we have seen, what events (perhaps clock readings) we take to be simultaneous depend on velocity.

We can avoid a lot of trouble if we make only such time comparisons as involve objects at the same point. Among such objects we may include electromagnetic waves emitted by the ship or by the earth as well as material objects.

How do we measure the clock rate on some object in motion with respect to ourselves, anyway? We can infer this rate from the relativistic Doppler frequency. Let us, however, work with this directly and see what sort of answers the approach leads us to.

First, let us consider the starship from the point of view of earth people. As indicated in Fig. 4, the ship goes out at a velocity  $v$  toward some point \* (perhaps a star). It then turns around and returns at the same velocity  $v$ . The point \* is at a distance  $L$  from earth as measured by earthmen.

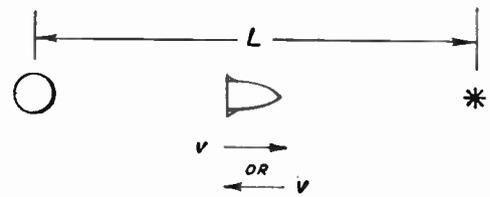


Fig. 4.

Suppose that the ship carries a standard oscillator just like one left on earth. Each oscillator would oscillate with the same frequency  $f$  if it were standing still. Now, let us measure elapsed time in terms of cycles of oscillation of these oscillators. On earth, the time  $t_e$  in seconds between the departure and the arrival of the starship will be

$$t_e = \frac{2L}{v}. \quad (39)$$

The number of cycles of oscillation  $N_e$  of the earth clock is thus

$$N_e = ft_e = \frac{2Lf}{v}. \quad (40)$$

What about radiation received from the ship? On the outward journey, when the ship is going away from earth, the Doppler frequency  $f_{d1}$  received at earth will be less than  $f$ ; in fact, (17) says that it will be

$$f_{d1} = \frac{\sqrt{1 - v/c}}{\sqrt{1 + v/c}} f. \quad (41)$$

But, radiation of this frequency is emitted from the ship right up to the time it turns around, when it is a distance  $L$  from the earth. The last of this radiation, which travels toward earth with a velocity  $c$ , will reach earth at a time  $L/c$  later than the time at which the ship turns

around. Hence, the radiation received on earth has the frequency  $f_{d1}$  for a time

$$L/v + L/c \tag{42}$$

out of the total time of the journey,  $2L/v$ . For the rest of the time, which must be

$$2L/v - (L/v + L/c) = L/v - L/c, \tag{43}$$

the Doppler frequency received will be that for a source approaching the earth, as given by (15); we will call this frequency  $f_{d2}$ .

$$f_{d2} = \frac{\sqrt{1 + v/c}}{\sqrt{1 - v/c}} f. \tag{44}$$

To get the total number of cycles of radiation  $N_s$  which have been produced on the ship and have struck the earth between the time the ship left and the time it returned, we should multiply the time (42) by the frequency (41) and add the product of the time (43) and the frequency (44). Thus

$$N_s = \frac{L f}{v} \left[ (1 + v/c) \frac{\sqrt{1 - v/c}}{\sqrt{1 + v/c}} + (1 - v/c) \frac{\sqrt{1 + v/c}}{\sqrt{1 - v/c}} \right]$$

$$N_s = \frac{2L f}{v} \sqrt{1 - (v/c)^2}. \tag{45}$$

We see by comparing (40) and (45) that according to this calculation, the oscillator on the ship must have run slower than the oscillator on earth by just the proper relativistic time shrinking factor,  $S$  of (1).

But, suppose that we are on the ship? Here the picture is as shown in Fig. 5. The universe is whizzing by in one direction or the other with a velocity  $v$ . Because of this, the distance between earth and \* has shrunk relativistically to a distance

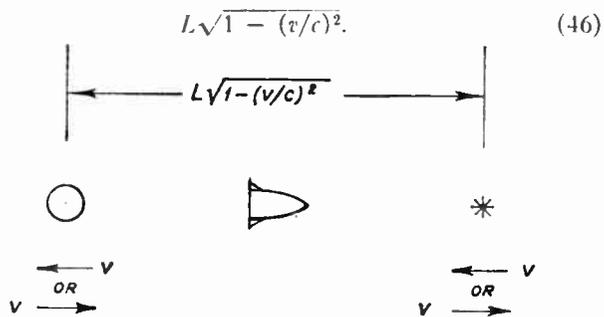


Fig. 5.

Accordingly, on the ship the "trip" from earth to \* and back takes a time  $t_s$

$$t_s = \frac{2L\sqrt{1 - (v/c)^2}}{v}. \tag{47}$$

The number of cycles  $N_s$  emitted by the ship's oscillator in this time is

$$N_s = f t_s = \frac{2L f \sqrt{1 - (v/c)^2}}{v}. \tag{48}$$

What about the radiation from earth? On the "journey out" the earth will appear to be moving away from us with a velocity  $v$  for a time half that of (47), which is the time for the two-way trip. But, when we reach \* we change our motion with respect not only to the earth and the star, but with respect to the radiation from the earth as well. The waves of radiation which we immediately encounter are already in space; we rush to meet them, and the observed frequency immediately changes from that of (17) to that of (15). Hence, we have the lower and the higher Doppler frequencies each for the same time. Thus, the total number of cycles  $N_e$  which leave earth and fall on the ship during both parts of the trip is

$$N_e = \frac{L f \sqrt{1 - (v/c)^2}}{v} \left[ \frac{\sqrt{1 - v/c}}{\sqrt{1 + v/c}} + \frac{\sqrt{1 + v/c}}{\sqrt{1 - v/c}} \right]$$

$$N_e = \frac{2L f}{v}. \tag{49}$$

Let us now compare the calculations made by the earthbound observer with those made by the observer on the ship. The results are exactly the same.  $N_s$  from (45) is exactly  $N_s$  from (48), and  $N_e$  from (40) is exactly  $N_e$  from (49). If we work matters out carefully, there is no clock paradox involved. The crew on the starship age less rapidly than those left behind on earth.

It may of course be objected that the acceleration of turning does dire things to the ship's clock. Here we can at least say that the effect should be the same for a long as for a short trip, and could only be additive.

The reader can easily arrive at the above result in another simple way. Assume two ships, the first going from the earth to \* with a velocity  $v$  and the second going from \* to the earth with a velocity  $v$ . Assume that the ships pass one another at \*. Assume that ship 1 and earth set their clocks to zero as ship 1 passes earth. Imagine that ship 2 sets its clock by the clock of ship 1 as ship 2 passes ship 1. Let earth and ship 2 compare clocks as ship 2 passes earth. It will be found that the reading of the clock of ship 2 (which represents elapsed time of ship 1 from earth to star plus elapsed time of ship 2 from star to earth) is smaller than the reading of the earth clock by the factor  $\sqrt{1 - (v/c)^2}$ , no matter what frame of reference we use.

### V. FREQUENCIES ON A SHIP WITH A CHANGING VELOCITY

Consider the moving space ship shown in Fig. 6. Here the acceleration  $a$  is upward. In  $a$ , at time  $t$ , the ship of length  $L$  is moving downward with a velocity  $v$ . In  $b$ , a time  $L/c$  later, the ship is stationary and the observer at the tail-end of the ship is right next to our stationary observer O. Let us assume that the relativistic Doppler shift accounts for the frequencies observed, and explore the consequences.

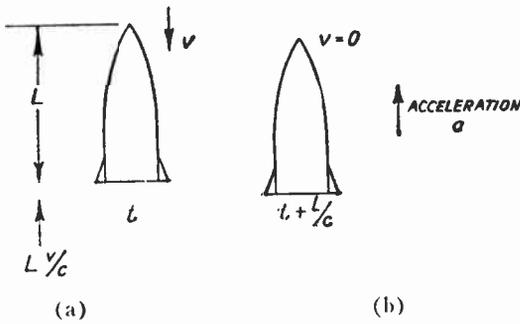


Fig. 6.

As we are dealing with moderate accelerations and velocities, we will assume Newtonian laws of motion. Further, we can at an appropriate point neglect  $Lv/c$  in comparison with  $L$ .

The radiation that the stationary observer, and the other stationary observer on the tail of the ship, see at time  $t + L/c$  left the nose a time very nearly  $L/c$  earlier, when the ship was traveling downward with a velocity  $v$ . Hence, the observed relativistic Doppler frequency for this radiation will be

$$f_a = f \frac{\sqrt{1 + v/c}}{\sqrt{1 - v/c}} \doteq f(1 + v/c). \quad (50)$$

Thus, the stationary observer and the observer in the tail of the ship find radiation as if from a clock running faster by a fraction  $\Delta f/f$

$$\Delta f/f = v/c. \quad (51)$$

Now, assuming Newtonian acceleration, very nearly

$$v = a \quad t = aL/c. \quad (52)$$

When we substitute  $v$  from (52) in (51) we obtain

$$\Delta f/f = aL/c^2. \quad (53)$$

Einstein's principle of equivalence asserts that a field due to acceleration is indistinguishable from a gravitational field. A man cooped up in the space ship thus has no way of knowing that the ship is changing velocity. He may think of it as sitting on the ground and experiencing a gravitational field of acceleration  $a$ . Hence, we conclude that a clock at height  $L$  in a gravitational field of acceleration  $a$  runs faster than a clock on the ground by a fraction given by (53).

This is a well-known relativistic clock effect which predicts a red shift in radiation from a massive body such as the sun.

We can deduce this effect in another way. If a mass  $m$  is lifted to a height  $L$  against a field of acceleration  $a$ , it gains an energy  $maL$  and this is a gain of a fraction  $aL/c^2$  of the total energy  $mc^2$ . Thus if a fraction of the mass is converted to photons which return to the ground a distance  $L$  below, these photons must have an energy greater by this same fraction  $aL/c^2$  than if they had been emitted before the mass was raised. As frequency is proportional to energy, the frequency of the photons

from the source at a height  $L$  will be raised by just the fraction given by (53).

In connection with the "clock paradox" of Section IV, it has been objected by some that "relativity" requires that the same time elapse for the traveler as for stay-at-homes, but this is not so.

Special relativity can deal only with relative time rates ascribable to velocity differences, not with those caused by gravitational fields or accelerations. In Section IV we scrupulously took relative velocities into account. We insisted, however, that the ship change its velocity with respect to earth, star and the radiation as well. We did this because the rockets applied a force to the ship, but not to the earth, the stars and the radiation.

The ship captain might insist, however, that during acceleration his ship is being supported in a gravitational field, while all the heavens fall past it. In this case he would have to take into account the effect of this gravitational field in increasing the frequency of the radiation traveling toward him between his ship and earth. No doubt the complete and general relativist could do not only this, he could also treat a case of protracted acceleration and deceleration. It seems simplest, however, to proceed as in Section IV, and other results must be consistent with the results obtained there.

### VI. CLOCK RATE ON A SATELLITE

Let us imagine a very accurate oscillator on a satellite in a circular orbit at radius  $r$ . According to Newton's laws, the orbital velocity is given by

$$v^2 = gr(r_0/r)^2. \quad (54)$$

Here  $r_0$  is the radius of the earth and  $g$  is the acceleration of gravity at the surface. Accordingly, there is a fractional *reduction* in clock rate as observed from the earth of

$$1 - \sqrt{1 - (v/c)^2} \doteq (1/2)gr_0(r_0/r)/c^2. \quad (55)$$

The potential energy of a unit mass at radius  $r$  with respect to the surface (at  $r_0$ ) is

$$g(r_0/r)(r - r_0). \quad (56)$$

In accordance with Section V, there is a fractional *increase* in clock rate because of the gravitation field of

$$g(r_0/r)(r - r_0)/c^2. \quad (57)$$

Thus, the net fractional change of clock rate,  $\Delta f/f$ , is

$$\frac{\Delta f}{f} = \frac{g}{c^2} \frac{r_0}{r} \left( r - \frac{3}{2} r_0 \right). \quad (58)$$

We must note that this will be the frequency when the satellite is directly overhead; a Doppler shift will be present when the satellite is approaching or receding, in accordance with (37) (overhead corresponds to  $\theta = 0$ ).

If the gravitational "red shift" were not taken into account, the frequency decrease at a radius of 6000

miles (a height of 2000 miles) would be about 4 parts in  $10^{10}$ . When the red shift is taken into account the net frequency change at this altitude is, according to (58), zero.

We should note that in making calculations of this sort, in accord with Section V, we take into account velocities and gravitational fields in the observer's frame of reference. Fields experienced by someone else, as, for instance, a force felt in the satellite because of an acceleration of the satellite, are irrelevant.

## VII. THE PHOTON ROCKET

Can we ever hope to travel with anything like the speed of light? For this purpose, some invoke the photon rocket. Of the total rest mass  $m_0$  of the rocket we regard all but a fraction  $a$  (or a rest mass  $am_0$ ) as fuel. We will not burn this fuel, nor fission it, nor fuse it. We will turn it all completely into radiation and use the pressure of this radiation to thrust the ship forward.

Of course we don't know how to turn matter completely into radiation. If we could do this, and if the minutest fraction of the radiation released heated the ship rather than pushing it forward, it would of course fry the crew. We will, however, assume that all the energy goes into pushing the ship, and none into heating it.

Before the ship sets off on its journey, it has zero velocity with respect to earth, a mass  $m_0$  and a total energy  $m_0c^2$ . The ship shoots out radiation of total energy  $E_r$  to the left and itself shoots off to the right, attaining a velocity  $v$ . Its remaining rest mass is now  $am_0$  and its relativistic mass is

$$\frac{am_0}{\sqrt{1 - (v/c)^2}}. \quad (59)$$

According to the conservation of energy, the energy of the radiation produced,  $E_r$ , plus the final energy of the ship must be equal to the initial energy of the ship,  $m_0c^2$

$$E_r + \frac{am_0c^2}{\sqrt{1 - (v/c)^2}} = m_0c^2. \quad (60)$$

According to the conservation of momentum, the momentum of the radiation (directed to the left) must equal the momentum attained by the ship (directed to the right), or

$$E_r/c = \frac{am_0v}{\sqrt{1 - (v/c)^2}}. \quad (61)$$

From (60) and (61) we obtain

$$v/c = \frac{1 - a^2}{1 + a^2}. \quad (62)$$

Now, we remember that to make time pass only a tenth as fast on the ship as on the earth, we must attain 99.5 per cent of the speed of light; that is,  $v/c$  must be 0.995. If we use a photon rocket to attain this speed, what fraction of the initial mass or matter will we have left?

The answer turns out to be a fraction

$$a = 0.05. \quad (63)$$

This sounds extreme, but think of the trip out and back! We start out with a mass  $m_0$ . After getting up to 99.5 per cent of the speed of light we have left of the ship only a rest mass of 0.05  $m_0$ . After stopping at the far end, we have only a mass (0.05) (0.05)  $m_0$ , or 0.0025  $m_0$ . After starting back and then stopping at earth, we are left with only a fraction 0.00000625 of our original mass, which sounds rather impractical.

Did we do the wrong thing? Instead of using up rest mass to shoot radiation off to the left, we might have used it to shoot a part of the ship off to the left with a velocity  $v_1$  and a relativistic mass  $m_1$ , while the ship proper, which retains a fraction  $a$  of the total original rest mass  $m_0$ , goes off to the right with a velocity  $v$ . Remember,  $m_1$  is *relativistic* mass and  $am_0$  is *rest mass*. With this in mind, the conservation of energy becomes

$$m_1c^2 + \frac{am_0c^2}{\sqrt{1 - (v/c)^2}} = m_0c^2. \quad (64)$$

The conservation of momentum is

$$m_1v_1 = \frac{am_0v}{\sqrt{1 - (v/c)^2}}. \quad (65)$$

From these equations we obtain

$$v/c = \sqrt{1 - a^2(1 + v/v_1)^2}. \quad (66)$$

We are of course interested in the case in which the fraction  $a$  is very small and  $v/c$  is very near to unity. In this case, unless the mass shot out to the left is very much greater than the remaining mass of the ship which moves to the right,  $v_1/c$  will be very nearly unity too. In other words,  $v/v_1$  will be nearly equal to one. If this is true and  $a$  is small, then (66) is approximately

$$v/c = 1 - 2a^2. \quad (67)$$

Under the same circumstances, according to (62) the photon rocket gives us a speed of approximately

$$v/c = 1 - 2a^2. \quad (68)$$

We see that (67) and (68) are the same. The remaining fraction  $a$  of the initial rest mass attains the same speed whether we turn all of the rest of the mass into photons and use these to accelerate the rest mass  $am_0$ , or whether we use some of the mass to shoot a part of the mass of the ship rearwards with a speed close to the speed of light.

We could do a little better if we could use the energy released to accelerate the ship by pushing against some fixed or very massive body—the universe, for instance. In this case all the energy will appear in the ship, and the conservation of energy will lead directly to the velocity, giving

$$\frac{am_0c^2}{\sqrt{1 - (v/c)^2}} = m_0c^2$$

$$v/c = \sqrt{1 - a^2} \tag{69}$$

when  $a$  is very small, very nearly

$$v/c = 1 - a^2/2. \tag{70}$$

We see that this is somewhat better than the other two cases [see (67) and (68)]. It is still not good enough to allow us to make a trip with nearly the speed of light. Moreover, I can't think of a plausible way to do it, except perhaps in taking off from a planet.

How is it that we find it possible on earth to accelerate electrons and protons far closer to the speed of light? The answer is, that we do have a massive body to push against, and besides this we do not rely on the energy of the particles, but instead have a huge fixed power plant; it is with these that we accelerate measly little atomic particles to speeds near that of light. With a sort of fixed space cannon and a huge mass-conversion power source we might shoot a space ship from earth with 99.5 per cent of the speed of light, but it seems unlikely that the crew would survive the acceleration.

### VIII. FUEL FROM SPACE

Space is full, loosely speaking, of matter, which is mostly hydrogen. It is said that on the average there is about one atom of hydrogen per cubic centimeter, but to be on the safe side let us assume some larger number  $M$  of hydrogen atoms per cubic centimeter, or  $M \times 10^6$  atoms per cubic meter. As the mass of a hydrogen atom is

$$1.66 \times 10^{-27} \text{ kilograms} \tag{71}$$

the mass of hydrogen per cubic meter is

$$1.66 \times 10^{-21} M \text{ kilograms.} \tag{72}$$

Let us assume a space ship of mass  $m_0$  kilograms which, in literally tunneling through space, sweeps up all the matter over an area  $A$ , turns it into radiation, and uses the radiation to push the ship forward. In a distance  $L$  the ship has acquired and used a mass which we shall call  $m_i$  of interstellar matter;  $m_i$  will be given by

$$m_i = 1.66 \times 10^{-21} maL \text{ kilograms.} \tag{73}$$

If  $v$  is the final velocity attained by the ship of rest mass  $m_0$  and  $E_r$  is the energy of the radiation produced,

the conservation of energy tells us that

$$E_r + \frac{m_0c^2}{\sqrt{1 - (v/c)^2}} = (m_i + m_0)c^2. \tag{74}$$

Initially, there is no radiation and neither the interstellar matter nor the ship has any momentum. Finally, in the tunnel behind the ship there is no interstellar matter (we used it up), there is radiation with momentum to the left, and the ship has a momentum to the right. The conservation of momentum tells us

$$E_r/c = \frac{m_0v}{\sqrt{1 - (v/c)^2}}. \tag{75}$$

From (74) and (75), we obtain

$$v/c = \frac{(1 + m_i/m_0)^2 - 1}{(1 + m_i/m_0)^2 + 1}. \tag{76}$$

From (73) we easily see that if the distance of travel,  $L$ , is zero,  $m_i$  is zero, and thus (76) tells us that we are at a standstill.

How fast are we going after a ten-light-year trip? A light year is

$$\text{one light year} = 9.48 \times 10^{15} \text{ meters.} \tag{77}$$

Let us assume a generous scoop (made of force fields, no doubt) which is 100 meters square, so that

$$A = 10^4 \text{ square meters.} \tag{78}$$

We thus see that the mass of interstellar matter collected will be

$$m_i = (1.66 \times 10^{-21})(10^4)(10)(9.48 \times 10^{15})M$$

$$m_i = 1.57 M \text{ kilograms.} \tag{79}$$

Surely, a habitable starship will weigh several thousand kilograms. Let us assume the mass of the ship to be 15,700 kilograms (about 17.5 tons). Let us generously assume not one but 1000 hydrogen atoms per cubic centimeter. This makes

$$m_i/m_0 = \frac{(1.57)(1000)}{(15,700)} = 0.1. \tag{80}$$

From (80) and (76), we obtain

$$v/c = 0.093. \tag{81}$$

Clearly, it is impossible to attain a velocity close to that of light by using interstellar matter as fuel.

# A Maser Amplifier for Radio Astronomy at X-Band\*

J. A. GIORDMAINE†, L. E. ALSOP†, C. H. MAYER‡, MEMBER, IRE, AND  
C. H. TOWNES†, SENIOR MEMBER, IRE

**Summary**—The design and operating characteristics of a maser radiometer for use in radio astronomy at 3-cm wavelength are discussed. The operating system which is described has a bandwidth of 5.5 mc and an input noise temperature, including background radiation into the antenna, of about 85°K. An rms fluctuation level of about 0.04°K is attained using an averaging time of 5 seconds. A discussion of the factors determining the sensitivity of such devices is presented.

## INTRODUCTION

THE low noise anticipated for maser amplifiers at the inception of maser techniques has by now received considerable experimental verification.<sup>1</sup> On the theoretical side, a lower limit to the effective input noise temperature<sup>2</sup> arising from spontaneous emission has been set at  $h\nu/k = 0.048\nu^\circ\text{K}$  for a two-level maser (or a three-level maser in the absence of noise associated with the pumping source), amplifying at a frequency  $\nu$  expressed in kmc. Experimental measurements of the noise figure of two- and three-level masers have been hampered by the presence of thermal noise from circuit components considerably in excess of the limiting noise. However, a value of close to 2°K has been set experimentally for the contribution of the amplification process in some maser devices.<sup>3</sup>

Realization of an amplification technique with such low inherent noise and the availability of a suitable three-level maser medium in ruby<sup>4</sup> has made possible the development of an operational radiometer whose sensitivity is largely limited by the purely thermal radiation from components preceding the amplifier and from antenna spillover.

The radiometer to be described has been in use in the study of continuum radiation from the planets and

other discrete radio sources with the U. S. Naval Research Laboratory 50-foot reflector since April, 1958. An effective system noise temperature of the order of 85°K has been attained, of which about 20°K is due to noise entering the antenna from external sources. The remaining noise of 65°K corresponds to a noise figure of  $1 + 65/290 = 1.2$ . An improvement in sensitivity for detection of thermal radiation of a factor of about 12 was obtained over a superheterodyne receiver with a noise figure of 9.8 db<sup>5</sup> and an IF bandwidth of 5.5 mc. The latter receiver, of course, accepted signal energy in both image bands. This system should be regarded only as a first approximation to what may be achieved by use of masers. Nevertheless its noise figure is notably better than that of previous radiometers, and a description of its characteristics will perhaps be of value for future and more finished system design.

The 5.5-mc bandwidth was obtained using a maser with voltage gain bandwidth product of  $\sim 50$  mc. A gain bandwidth product of  $\sim 100$  mc was available from the same maser under laboratory conditions of optimum magnetic field homogeneity.

## IMPROVEMENT IN SENSITIVITY WITH A MASER PREAMPLIFIER

The radiometer was a modified Dicke-type system<sup>6</sup> in which the receiver consisted of a ruby maser pre-amplifier followed by a superheterodyne receiver. We wish to estimate the sensitivity of such a radiometer in terms of the maser and receiver parameters.

Let

$T_1^\circ\text{K}$  = operating noise temperature at the input of the maser, including contributions from the following noise sources:

- Radiation from lossy ferrite switch and circulator preceding the maser.
- Waveguide losses between the input horn and the maser cavity.
- Radiation from cavity walls at the temperature of the cooling bath.
- Spontaneous emission in ruby.
- Ground pickup through antenna spillover..
- Thermal radiation from atmospheric absorption.

$T_2^\circ\text{K}$  = equivalent temperature of the receiver input, equal to  $(F - 2)290^\circ\text{K}$  including mixer noise, IF amplifier noise, and mixer conversion loss.

<sup>5</sup> The noise figure of 9.8 db corresponds to an effective input noise temperature of  $(F - 2)290 = 2190^\circ\text{K}$  at the input to the 5.5-mc band-pass. This is just twice the noise figure measured by use of a broadband noise source substituted for the antenna horn. It is expressed in this form for easy comparison with a single-sideband receiver.

<sup>6</sup> R. H. Dicke, "Measurement of thermal radiation at microwave frequencies," *Rev. Sci. Instr.*, vol. 17, pp. 268-275; July, 1946.

\* Original manuscript received by the IRE, January 6, 1959; revised manuscript received, March 27, 1959. Work at Columbia University, New York, N. Y., supported jointly by the U. S. Army Signal Corps, the Office of Naval Res., and the Air Force Office of Scientific Res.

† Dept of Phys., Columbia University.

‡ Radio Astronomy Branch, U. S. Naval Res. Lab., Washington, D. C.

<sup>1</sup> J. P. Gordon and L. D. White, "Noise in maser amplifiers—theory and experiment," *Proc. IRE*, vol. 46, pp. 1588-1594; September, 1958. References to the published theoretical and experimental work on maser noise figures are given in this paper.

<sup>2</sup> The term "effective input noise temperature" is used according to the definition in the reference above and is equal to  $(F - 1)290^\circ\text{K}$ . In a paper by L. E. Alsop, J. A. Giordmaine, C. H. Townes, and T. C. Wang, "Measurement of noise in a maser amplifier," *Phys. Rev.*, vol. 107, pp. 1450-1451, September 1, 1957, the term "noise figure" was used for the quantity  $(F - 1)$ , leading to the quotation of maser noise figures less than unity. The reference temperature was taken as  $300^\circ\text{K}$ . In this paper the term noise figure and the symbol  $F$  are used in accordance with the accepted definition. See H. T. Friis, "Noise figures of radio receivers," *Proc. IRE*, vol. 32, pp. 419-422; July, 1944.

<sup>3</sup> A. C. McWhorter and F. R. Arams, "System-noise measurement of a solid state maser," *Proc. IRE*, vol. 46, pp. 913-914; May, 1958.

<sup>4</sup> G. Makhov, C. Kikuchi, J. Lambé, and R. W. Terhune, "Maser action in ruby," *Phys. Rev.*, vol. 109, pp. 1399-1400; February 15, 1958.

$b$  = maser voltage gain-bandwidth product, expressed in mc.  
 $\Delta\nu_1$  = maser bandwidth at half power, expressed in mc.  
 $\Delta\nu_2$  = IF bandwidth at half power, expressed in mc.  
 $\Delta\nu_0$  = bandwidth of filter following lock-in detector, expressed in mc.

In the absence of the maser preamplifier, the superheterodyne receiver accepts continuum signal power from both the signal band and the image band, *i.e.*,  $2\Delta\nu_2$ , the only input preselection occurring in the mixer crystals. The 30 cps half-wave lock-in detector is preceded by a narrow band audio filter. For square-switching and demodulation waveshapes, the rms fluctuation level at the receiver output, expressed in terms of antenna temperature, is given<sup>6,7</sup> approximately by

$$\Delta T = \sqrt{2} T_2 \left( \frac{\Delta\nu_0}{\Delta\nu_2} \right)^{1/2} \tag{1}$$

It is assumed that the IF response is constant over  $\Delta\nu_2$  and otherwise zero, and that  $\Delta\nu_0 \ll \nu_0 \ll \Delta\nu_2$ , where  $\nu_0$  is the switching frequency. The output filter frequency response is somewhat complicated but may be approximated by a bandwidth  $\Delta\nu_0 = 1/4\tau$ , where  $\tau$  is a response time of 5 seconds. Eq. (1) becomes

$$\Delta T = \frac{T_2}{(2\tau\nu_2)^{1/2}} \tag{2}$$

For a noise figure of 9.8 db,  $T_2 = 2190^\circ\text{K}$ . Taking  $\tau = 5$  seconds and  $\Delta\nu_2 = 5.5$  mc, we obtain from (2)  $\Delta T = 0.3^\circ\text{K}$ . The disagreement between this temperature fluctuation and the observed value of  $0.5^\circ\text{K}$  is attributed to appreciable distortion of the assumed square wave modulation waveshape and to use of half-wave lock-in detection with an audio filter which did not completely reject harmonics of the switching frequency.

In the presence of the maser preamplifier, the noise temperature  $T_2$  referred to the input of the system is

$$T_2^* = \frac{T_2}{2 + \frac{(G-1)\Delta\nu_1}{\Delta\nu_2}} \tag{3}$$

where it is assumed for simplicity that the maser has uniform power gain  $G$  over the bandwidth  $\Delta\nu_1$  within the receiver signal band, and unity gain elsewhere including the image band. In this notation,  $b = \sqrt{G}\Delta\nu_1$  and as may be seen from (23)  $b$  is approximately independent of  $G$  for a given maser design. It is noted that in the extreme case of  $G = 1$ , *i.e.*, of the maser removed from the system,  $T_2^* = T_2/2$  since both the signal and image bands are receptive. Where the maser has useful gain and bandwidth, the effective receiver bandwidth approaches  $\Delta\nu_1$  and from (2) we obtain

$$\Delta T = \left( \frac{2}{\tau} \right)^{1/2} \left[ \frac{T_1}{(\Delta\nu_1)^{1/2}} + \frac{T_2^*}{(\Delta\nu_2)^{1/2}} \right] \tag{4}$$

For the useful case of  $G \gg 1$  and  $G\Delta\nu_1 \gg 2\Delta\nu_2$

$$\Delta T = \left( \frac{2}{\tau\Delta\nu_1} \right)^{1/2} \left[ T_1 + \frac{T_2}{G_1} \left( \frac{\Delta\nu_2}{\Delta\nu_1} \right)^{1/2} \right] \tag{5}$$

Where the available gain bandwidth  $b$  of the maser is small, *i.e.*,

$$b < \sqrt{\frac{2T_2}{T_1}} \Delta\nu_2,$$

the fluctuation level (5) can be minimized by adjusting the gain of the maser to an optimum value

$$G = \left[ 2 \frac{T_2}{T_1} \left( \frac{\Delta\nu_2}{b} \right)^{1/2} \right]^{4/3} \tag{6}$$

With optimum gain,

$$\Delta T = \frac{3}{(2\tau)^{1/2}} \left[ \frac{2T_2T_1^2(\Delta\nu_2)^{1/2}}{b^2} \right]^{1/3} \tag{7}$$

For a larger gain-bandwidth product,

$$b > \sqrt{\frac{2T_2}{T_1}} \Delta\nu_2,$$

the optimum gain is just the gain for which the maser bandwidth is equal to the receiver bandwidth. In this case

$$\begin{aligned} \Delta T &= \left( \frac{2}{\tau\Delta\nu_2} \right)^{1/2} \left( T_1 + \frac{T_2}{G} \right) \\ &= \left( \frac{2}{\tau\Delta\nu_2} \right)^{1/2} \left[ T_1 + \frac{T_2(\Delta\nu_2)^2}{b^2} \right] \end{aligned} \tag{8}$$

The limiting sensitivity for

$$b \gg \sqrt{\frac{T_2}{T_1}} \Delta\nu_2$$

is

$$\Delta T = \left( \frac{2}{\tau\Delta\nu_2} \right)^{1/2} T_1 \tag{9}$$

In this case the effective system input noise temperature is just the input temperature to the maser.

Now the bandwidth of a microwave superheterodyne receiver can be varied considerably without a large modification in noise figure. As a corollary of (8) and (9), the optimum receiver bandwidth for use with a maser preamplifier of gain bandwidth product  $b$  is

$$(\Delta\nu_2)_{\text{opt}} = b \sqrt{\frac{T_1}{3T_2}} \tag{10}$$

Assuming that the optimum receiver bandwidth can be used, and that the gain of the maser has been opti-

<sup>7</sup> W. Selove, "A dc comparison radiometer," *Rev. Sci. Instr.*, vol. 25, pp. 120-122; February, 1954.

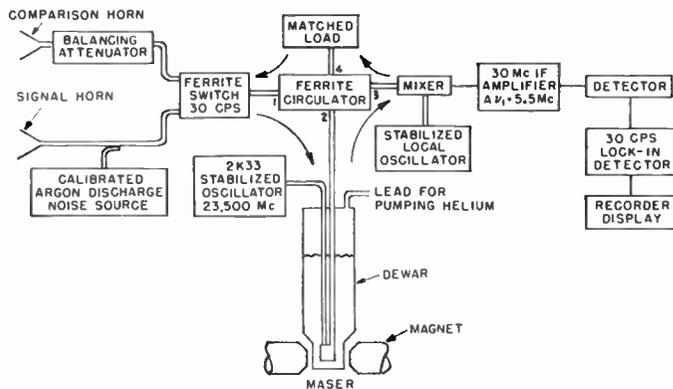


Fig. 1—Schematic diagram of the maser radiometer. The components, with the exceptions of the lock-in amplifier and the recorder display, were mounted at the focus of the reflector.

mized, the minimum fluctuation level obtainable with a maser of gain bandwidth product  $b$  is given by

$$\Delta T = \frac{3.6}{\sqrt{2b\tau}} T_1 \left( \frac{T_2}{T_1} \right)^{1/4} \quad (11)$$

In the maser radiometer to be described, pertinent parameters for the superheterodyne receiver are given by Table I. The minimum observed fluctuation level was about  $0.04^\circ\text{K}$ . This result represents an improvement of a factor of approximately 12 in sensitivity over the radiometer without the maser preamplifier.

TABLE I  
CHARACTERISTICS OF SUPERHETERODYNE RECEIVER

$T_2 = 2190^\circ\text{K}$ (noise figure 9.8 db)
$\Delta\nu_2 = 5.5$ mc
$\tau = 5$ seconds

The contributions to the input noise temperature  $T_1$  are listed below. The contribution from antenna spillover is an estimate based on absolute gain measurements in the forward direction, and an assumed pattern in the backward direction.

- Ferrite switch (antenna-comparison horn), 0.2-db forward loss,  $13^\circ$ .
- External waveguide losses (2 feet of RG 52/U and associated flanges),  $10^\circ$ .
- Ferrite circulator, 0.1-db forward loss,  $7^\circ$ .
- Atmospheric absorption (antenna elevation  $45^\circ$ ),  $8^\circ$ .
- Spontaneous emission,  $3^\circ$ .
- Cavity walls (bath temperature  $1.4^\circ\text{K}$ ),  $< 0.5^\circ$ .
- Antenna spillover,  $20^\circ$ .

The contribution of receiver noise to the input temperature is  $T_2^* = T_2/G = 26^\circ\text{K}$ , giving a total input noise temperature of approximately  $86^\circ\text{K}$ . The predicted fluctuation level (8) is  $0.02^\circ\text{K}$ , in reasonable agreement with the observed value of  $0.04^\circ\text{K}$ . The difference between predicted and observed fluctuations may come primarily, as in the case of the radiometer

\* D. C. Hogg, *J. Appl. Phys.*, to be published.

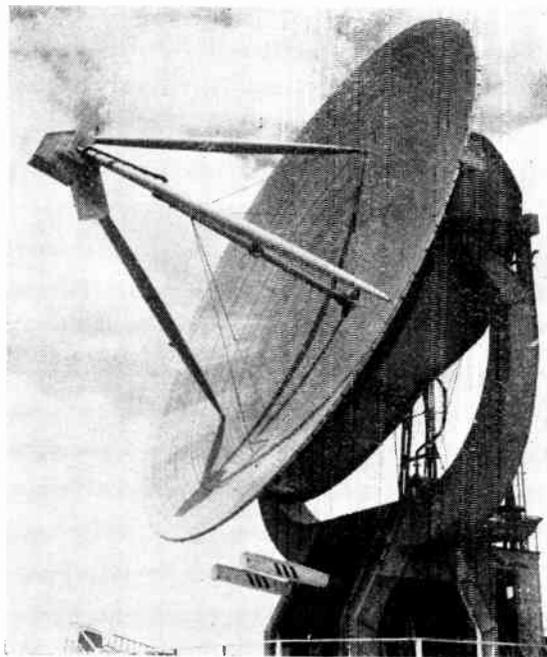


Fig. 2—The radiometer mounted near the focus of the NRL 50-foot reflector.

without a maser, from inaccurate assumptions about the modulation shape and characteristics of the coherent detection.

With an available voltage-gain bandwidth product of 50 mc, the optimum receiver bandwidth in the presence of  $2190^\circ\text{K}$  receiver input temperature is 4.7 mc [cf. (13)]. The actual receiver bandwidth of 5.5 mc was thus almost the optimum bandwidth.

#### APPARATUS: GENERAL DESCRIPTION

The radiometer is shown schematically in Fig. 1. The maser, the local oscillator, and the IF amplifier were mounted adjoining the focus of the 50-foot reflector.<sup>9</sup> The over-all mounting is indicated in Fig. 2, in which the radiometer is shown encased in a protective covering. The vacuum pump for maintaining helium pressure at 1 to 2 mm was mounted behind the reflector. In Fig. 3 the equipment at the focus is shown.

The apparatus was modified Dicke system<sup>10</sup> in which a ferrite circulator was used to switch alternately between the 50-foot antenna and a comparison horn pointed at the sky. A four-port circulator was used to isolate the maser from receiver radiation. The second detector output was synchronously detected in a half-wave lock-in detector, with the reference voltage the same square-wave voltage used to drive the ferrite switch. The output of the lock-in detector was displayed on a recorder.

Preparation of the system for observing was carried

<sup>9</sup> The attenuation in the length of RG52/U waveguide necessary to connect the feed horn to a maser amplifier at the base of the antenna would have introduced an additional  $150^\circ\text{K}$  to the input noise temperature.

<sup>10</sup> C. H. Mayer, T. P. McCullough, and R. M. Sloanaker, "Measurements of planetary radiation at centimeter wavelengths," *Proc. IRE*, vol. 46, pp. 260-266; January, 1958.

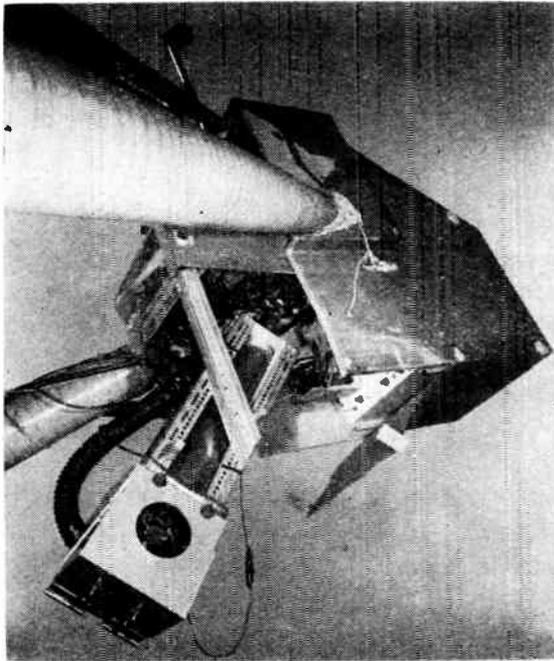


Fig. 3—View of the assembly mounted near the focus.

out on the ground by precooling the dewar containing the maser cavity and input waveguide by filling the outer portion of the dewar with liquid nitrogen and the inner dewar with liquid helium. The dewar is mounted at the focus, and the necessary waveguide connections are made. The helium capacity of the dewar allowed for 12 to 15 hours of continuous operation at 1.4°K.

Although not continuously variable, the operating frequency could be adjusted by moving the ruby about slightly in the cavity or by variation in a small amount of dielectric material also inserted in the cavity. The system was operated in the frequency range 8700–10,100 mc.

THE MASER

Theoretical Considerations

The maser was a three-level<sup>11</sup> solid state device using ruby<sup>4</sup> (Al<sub>2</sub>O<sub>3</sub>·0.01Cr<sub>2</sub>O<sub>3</sub>) as the paramagnetic medium. The energy levels of the Cr<sup>+3</sup> ion in ruby<sup>12,13</sup> are shown as a function of magnetic field *H* in Fig. 4 for the case of an angle of 60° between the direction of the static mag-

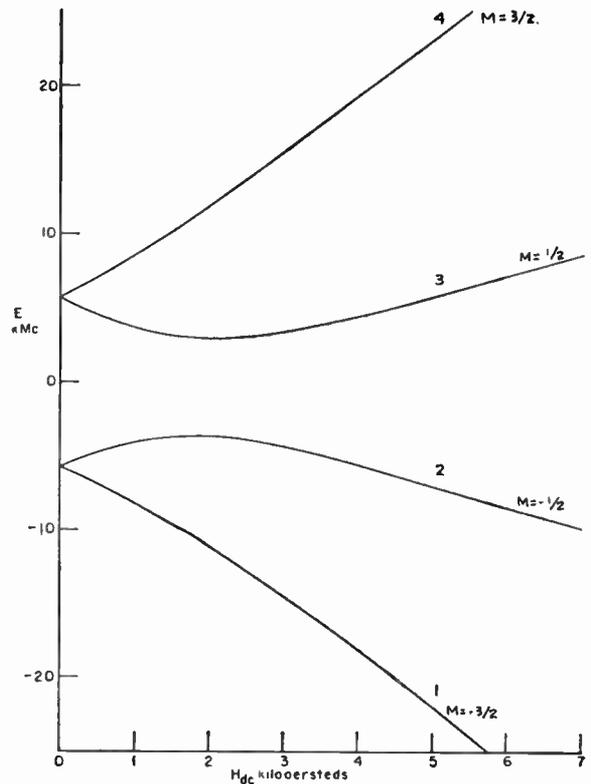


Fig. 4—Energy levels of Cr<sup>+3</sup> in a ruby as a function of magnetic field for an angle of 60° between the crystal axis of symmetry and the dc magnetic field.

netic field and the crystal axis of symmetry. The levels are numbered in order of increasing energy and by the quantum number *M* applicable at high magnetic field. At thermal equilibrium at a bath temperature *T*, the ratio of the population of level 3 to that of level 2 is the Boltzmann ratio  $\exp(-h\nu_{23}/kT)$  and in the high temperature approximation  $T > h\nu_{14}/k$

$$n_2 - n_3 \approx \frac{Nh}{4kT} \nu_{23} \tag{12}$$

where *N* is the total number of spins present, *h* is Planck's constant, *k* is Boltzmann's constant, and *n*<sub>2</sub>, *n*<sub>3</sub> are the populations of levels 2 and 3, respectively.

In the presence of pumping radiation at the frequency  $\nu_{13}$  sufficient to saturate the transition 1→3 the populations assume new steady state values and

$$n_2 - n_3 = \frac{Nh}{4kT} \left[ \frac{-\nu_{12} \left( w_{12} + \frac{w_{14}w_{42}}{w_{14} + w_{24} + w_{34}} \right) + \nu_{23} \left( w_{23} + \frac{w_{24}w_{43}}{w_{14} + w_{24} + w_{34}} \right)}{\left( w_{12} + \frac{w_{14}w_{42}}{w_{14} + w_{24} + w_{34}} \right) + \left( w_{23} + \frac{w_{24}w_{43}}{w_{14} + w_{24} + w_{34}} \right)} \right] \tag{13}$$

<sup>11</sup> N. Bloembergen, "Proposal for a new type solid state maser," *Phys. Rev.*, vol. 104, pp. 324–327; October 15, 1956.

<sup>12</sup> Energy levels can be derived from published tables. See P. M. Parker, "Nuclear quadrupole levels in single crystals," *J. Chem. Phys.*, vol. 24, pp. 1096–1102; October, 1956.

<sup>13</sup> W. S. Chang and A. E. Siegman, "Characteristics of Ruby for Maser Applications," Stanford Electronics Lab. Tech. Rep. No. 156-2; September 30, 1958 (unpublished). The labelling of the levels in this report is opposite that of the present paper.

where *w*<sub>*ij*</sub> sec<sup>-1</sup> is the transition probability for an ion to relax from level *i* to level *j*. The *w*<sub>*ij*</sub> are related to the relaxation times *T*<sub>*ij*</sub> by the relations

$$w_{ij} + w_{ji} = \frac{1}{T_{ij}}$$

and

$$w_{ij}/w_{ji} = \exp(-hv_{ij}/kT) \quad i < j. \quad (14)$$

Amplification due to stimulated emission at frequency  $\nu_{23}$  is obtained when the population of level 3 exceeds that of level 2. Thus we obtain from (13) as a condition for amplification at  $\nu_{23}$ ,

$$\nu_{12} \left( w_{12} + \frac{w_{14}w_{42}}{w_{14} + w_{24} + w_{34}} \right) > \nu_{23} \left( W_{23} + \frac{w_{21}w_{13}}{w_{14} + w_{24} + w_{34}} \right). \quad (15)$$

A negative quality factor  $Q_M$  can be assigned to the amplifying material as indicated

$$Q_M = \frac{2\pi\nu_{23} \frac{1}{8\pi} V_c (\overline{H_{rf}^2})_c}{h\nu_{23}(n_2 - n_3) \overline{W}_{32}} = \frac{V_c (H_{rf}^2)_c}{4h(n_2 - n_3) \overline{W}_{32}} \quad (16)$$

where

- $V_c$  is the cavity volume,
- $(\overline{H_{rf}^2})_c$  is the mean square of the RF magnetic field amplitude over the cavity, and
- $\overline{W}_{32}$  is the average probability for an RF induced  $3 \rightarrow 2$  transition.

For the ideal case where  $M$  is a good quantum number, the value of  $W_{M,M-1}$  is given by

$$W_{M,M-1} = r \frac{\pi^2}{4} \frac{g^2 \beta^2}{h^2} H_{rf}^2 (S + M)(S - M + 1) g(\nu) \quad (17)$$

where

- $\beta$  is the Bohr magneton  $0.93 \times 10^{-20}$  ergs/gauss,
- $g$  is the splitting factor, 2.0 in the case of  $\text{Cr}^{+3}$  in ruby,
- $g(\nu)$  is the normalized<sup>14</sup> line shape of the magnetic resonance,
- $M$  is the magnetic quantum number describing the state at high magnetic field, and
- $S$  is the effective spin quantum number describing the ion.

Eq. (17) applies when, as is normal, the width of the magnetic resonance is much larger than the amplifier bandwidth. The factor  $r$  is a measure of the effectiveness of  $H_{rf}$ . For  $g\beta H_{DC} \gg |D|$ ,<sup>15</sup>  $r=1$  for  $H_{rf}$  perpendicular to  $H_{DC}$  and  $r=0$  for  $H_{rf}$  parallel to  $H_{DC}$ . For  $g\beta H_{DC} \lesssim |D|$   $r$  becomes a function of  $H_{DC}$  and the relative orientations of  $H_{rf}$ ,  $H_{DC}$ , and the crystalline axis of symmetry. Tables have been prepared permitting the calculation of  $r$ .<sup>13</sup>

In the case of a homogeneously broadened line of Lorentz shape with total width  $\Delta\nu_M$  at half power, the peak value of  $g(\nu)$  is

$$g(\nu) = \frac{2}{\pi \Delta\nu_M}. \quad (18)$$

<sup>14</sup>  $\int_0^\infty g(\nu) d\nu = 1$ .

<sup>15</sup> The zero field splitting of  $\text{Cr}^{+3}$  in ruby is  $2D = -11.46$  mc.

Then the maximum value of  $W_{32}$  is given by

$$W_{32} = \frac{8\pi\beta^2 r H_{rf}^2}{h^2 \Delta\nu_M} \quad (19)$$

and

$$\overline{W}_{32} = \frac{8\pi\beta^2}{h^2 \Delta\nu_M} (\overline{r H_{rf}^2})_{x\text{tal}} \equiv \frac{8\pi\beta^2 \overline{r} (\overline{H_{rf}^2})_{x\text{tal}}}{h^2 \Delta\nu_M} \quad (20)$$

where  $(\overline{\quad})_c$  and  $(\overline{\quad})_{x\text{tal}}$  refer to averages over the volumes of the cavity and the ruby, respectively. From (16) and (20)

$$Q_M = \frac{hV_c \Delta\nu_M}{32\pi\beta^2 k(n_2 - n_3)} \frac{(\overline{H_{rf}^2})_c}{(\overline{H_{rf}^2})_{x\text{tal}}} = \frac{hV_{x\text{tal}} \Delta\nu_M f}{32\pi\beta^2 k(n_2 - n_3)}, \quad (21)$$

where the cavity filling factor  $f$  is defined by

$$f = \frac{V_c (\overline{H_{rf}^2})_c}{V_{x\text{tal}} (\overline{H_{rf}^2})_{x\text{tal}}}$$

The filling factor  $f$  has a maximum value of unity when the cavity is filled with paramagnetic material. The  $r$  factor has a value of unity, for example, at high values of  $H_{DC}$  for the case of a  $\text{TE}_{0mn}$  mode in a rectangular cavity with the  $E$  field parallel to  $H_{DC}$ .

At cavity resonance, the ratio of reflected to incident power,  $G$ , is related to  $Q_e$ , the external cavity  $Q$ , to  $Q_0$  the unloaded  $Q$  of the cavity, and to  $Q_M$  by

$$\sqrt{G} = \frac{\frac{1}{Q_e} - \left( \frac{1}{Q_0} + \frac{1}{Q_M} \right)}{\frac{1}{Q_e} + \left( \frac{1}{Q_0} + \frac{1}{Q_M} \right)} = \frac{\frac{1}{Q_e} + \frac{1}{|Q_M|} - \frac{1}{Q_0}}{\frac{1}{Q_e} - \frac{1}{|Q_M|} + \frac{1}{Q_0}} \quad (22)$$

From (22) and the relation  $\nu_{23} = Q\Delta\nu_1$  we obtain

$$(\sqrt{G} - 1)\Delta\nu_1 = 2\nu_{23} \left( \frac{1}{|Q_M|} - \frac{1}{Q_0} \right) \quad (23a)$$

$$(\sqrt{G} + 1)\Delta\nu_1 = \frac{2\nu_{23}}{Q_e} \quad (23b)$$

Thus at high gain the root gain bandwidth product is approximately a constant.

### Performance Characteristics

Two masers were constructed and used in the work. The first, designated by I, had a low filling factor. The second, II, whose characteristics were outlined in the Introduction, had a filling factor of almost unity and a near-optimum configuration of  $H_{rf}$ . The characteristics of the two masers are summarized in Table II.

The value of  $N$  is calculated from the crystal volume and an assumed concentration (Cr/Al) of 0.05 per cent. The rubies were prepared by Linde Air Products Corporation from a mixture with an initial concentration of 0.1 per cent.

TABLE II  
MASER CHARACTERISTICS

	Maser I	Maser II
Maser gain-bandwidth, $\sqrt{G}\Delta\nu_1$ mc	10	50
Magnetic $Q$ , $Q_M$	1800	380
Amplification frequency, $\nu_{23}$ mc	9000	9500
Pump frequency, $\nu_{13}$ mc	21,000	23,500
Orientation of magnetic field relative to crystal symmetry axis, $\theta$	59°	55°
Static magnetic field, $H_{DC}$ oer.	3360	3740
Unloaded cavity $Q(\nu_{23})$ , $Q_0(\nu_{23})$	10,000	7000
Bath temperature, $T^\circ\text{K}$	1.4°	1.4°
Total cavity $Q(\nu_{13})$ , $Q(\nu_{13})$	5000	4000
Pumping power mw	30	30
Cavity mode ( $\nu_{23}$ )	TE <sub>102</sub>	TE <sub>101</sub>
Cavity volume, $V$ , cm <sup>3</sup>	4.8	0.48
Crystal volume, $V_{\text{crystal}}$ cm <sup>3</sup>	0.54	0.35
Number of Cr <sup>+3</sup> ions, $N$	$1.3 \times 10^{19}$	$8.0 \times 10^{18}$
Filling factor, $f$		0.8
Transition probability factor, $F$		0.72
Maser bandwidth at half power, $\Delta\nu_M$ mc	200	200

The masers were operated over a range of frequencies between 8700 and 10,100 mc using a variety of cavity modes at the pumping frequency, so that the operating characteristics quoted should be considered only as representative. The value of 50 mc is the value of  $\sqrt{G}\Delta\nu_1$  obtained using a small permanent magnet. Under laboratory conditions with higher field homogeneity, a value of  $\sqrt{G}\Delta\nu_1 \leq 100$  mc was observed.

An accurate calculation of the expected  $\sqrt{G}\Delta\nu_1$  is complicated by lack of detailed knowledge of the transition probabilities  $w_{ij}$  between the various levels. A crude estimate of  $\sqrt{G}\Delta\nu_1$  for II using the parameters of Table II in (13) and (21), assuming  $W_{12} = W_{23}$  and neglecting the effects of  $\Delta M = 2$  relaxation, gives a value of 48 mc for  $\sqrt{G}\Delta\nu_1$ . This value is in general agreement with that observed.

The gain stability of the system was usually limited by variations in the impedance at the output of the maser, although small changes in critical mechanical parts or in magnetic field strength sometimes also gave troublesome gain variations. In common with any highly regenerative reflection type amplifier the gain was quite sensitive to the matching of the output load. With a maser gain of 20 db a mismatch introducing a voltage reflection coefficient of 1 per cent could give rise to a gain variation of  $\pm 0.8$  db.

Most of the observations with the present system were carried out with an rms gain variation of the order of 0.1 db/min. Most of this variation was periodic with a highly reproducible period of about 44 minutes. This component of the variation was attributed to the impedance variations arising from the moving surface of the liquid helium within the signal waveguide. This source of gain variation was eliminated by filling the signal waveguide with polystyrene foam which displaced the liquid helium. Although the system has not yet been observed over long periods of time with this modification, it appears that the average gain variation with the antenna stationary has been reduced to less than 0.01 db/min. This gain variation is not serious for most astronomical purposes if the ferrite switch connects

two sources of noise with effective temperatures that differ by less than 1°.

The sensitivity of the gain to load variations prevented the replacement of the ferrite switch by a mechanical chopping arrangement. The mechanical chopper was installed in the hope of removing the appreciable noise contribution from the ferrite switch.<sup>16</sup> In this arrangement, the signal and comparison horns were connected to terminals 1 and 4, respectively, (Fig. 1) of the circulator. The temperatures seen by the horns were compared by insertion of a motor-driven shorting blade at terminal 1. It was observed that the maser gain with the blade inserted was as much as 3 db greater than the gain with the blade removed. As the blade was inserted and removed higher transient gains were sometimes seen. This type of behavior is consistent with the inadequate isolation of about 22 db provided by the circulator between terminals 2 and 1. To avoid gain variations of this kind the ferrite switch was reinstalled.

#### Cavity and Coupling Arrangements

In both masers I and II the cavity was a rectangular type supporting a TE<sub>10n</sub> frequency. In II the amplifying mode was the fundamental TE<sub>101</sub> mode with the RF electric field approximately parallel to  $H_{DC}$ . No identification of the K-band modes was attempted. The density of modes at K-band was large and the mode selection was not found to be critical. Most modes of similar  $Q$  appeared to have equivalent effect in driving the system.

The cavity and coupling scheme for II are shown in Fig. 5. The coupling from RG53/U waveguide to the K-band mode was via a fixed iris in the lead window closing the cavity. The lead window was pressed onto the cavity edge by a stainless steel clamp to provide satisfactory electrical contact. The X-band coupling was via a thin inductive post passing across the X-band waveguide and protruding into the cavity along the RF electric field. In the case of I the coupling could only be adjusted internally. As a result, slight and irregular coupling variations accompanying the violent thermal cycling often prevented the optimum coupling from being available after cooling. In II the coupling probe could be adjusted from outside the dewar using a rack and pinion linkage.

The ruby was supported in the cavity of I by a perforated teflon block. The frequency of the cavity could be varied over a range of 800 mc by adjusting the position of the ruby in the teflon block. Frequency adjustments in II were obtained by moving the ruby relative to a dielectric spacer plate in the cavity.

Initially, microphonics were a persistent problem in operating the system. It was found that in I, motion of the antenna caused appreciable gain fluctuations due to small motions of the crystal in the cavity. In II the microphonics were almost completely eliminated by applying pressure to the crystal through spring loaded teflon rods pressing through three of the cavity walls.

<sup>16</sup> This arrangement was suggested by F. Arams.

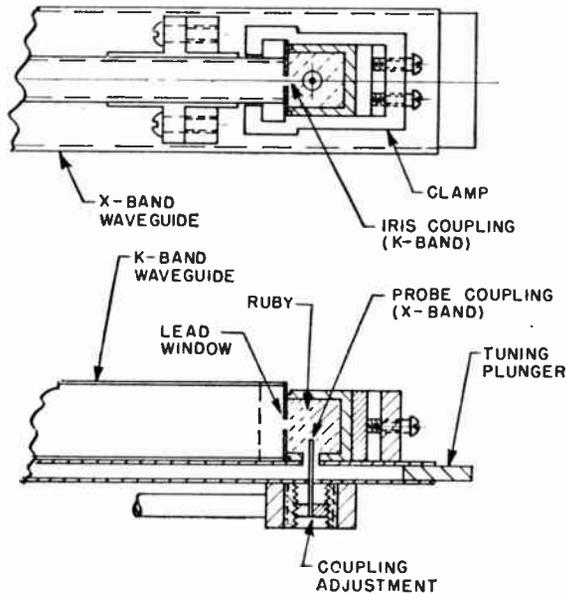


Fig. 5—Cavity and coupling scheme of maser II.

### Pumping

Pumping power was provided by a QK306 (21–22 kmc) or 2K33 (22–24 kmc) klystron mounted beside the dewar. At a pumping power of 30 mw no appreciable gain variation was found to be associated with variations in power output of the pump source. A power stabilization circuit using a current controlled ferrite isolator as a power regulator was installed but found unnecessary.

The klystron frequency was stabilized against long term drifts by being locked to an external wavemeter which could be tuned remotely. The stabilization scheme<sup>17</sup> (Fig. 6) involved sweeping the klystron frequency continuously over several hundred kc centered about the wavemeter frequency. By comparing the phase of the sweep frequency component of the crystal detector current with that of the sweep voltage in a lock-in detector, a correction signal was obtained. The correction signal was applied to the klystron repeller through a 2B23 magnetic diode. The klystron frequency was tuned to the cavity resonance by tuning the wavemeter for maximum maser gain. The maser gain was found to be relatively insensitive to driving frequency near the peak of the cavity resonance.

### Cryogenic System and Magnet

The elongated stainless steel dewar was of conventional design. The main helium storage chamber was 18 inches inside height and had a capacity of 3 liters. The "finger" extending between the magnet pole pieces was 1 inch ID and 1½ inches OD. The X-band waveguide was tapered in both dimensions to fit the finger. Both waveguides were of brass milled to a thickness of

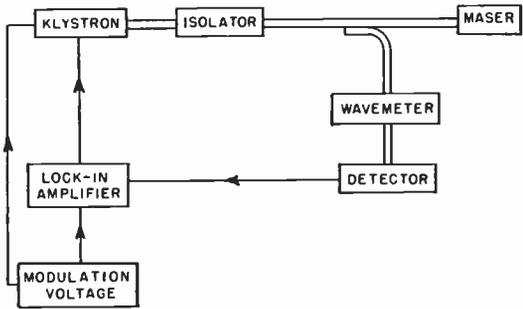


Fig. 6—Schematic diagram of pumping klystron stabilization system.

0.015 inch to minimize helium loss. The dewar was mounted at the focus so as to be vertical at an antenna elevation of 45°, limiting the maximum tilt of the dewar to 45°.

After the dewar was pumped down to a pressure of about 2 mm the remaining helium lasted 12 to 15 hours. The pump was of 15 cubic ft/min capacity and was mounted behind the reflector surface. The 30-foot connection to the focus was via 3-inch tubing.

The system was operated at 1.4°K. Good helium circulation inside the cavity and around the crystal was found to be important for effective cooling in the presence of microwave pumping power. In II the ruby faces were serrated to insure helium circulation around the crystal.

The magnetic field was provided by a modified commercial permanent magnet (No. 5725, Arnold Engineering Co., Marengo, Ill.). The commercial magnet produced a field of 1600 oersteds, with 3.5-inch diameter pole faces and a 3.6-inch diameter pole gap. The magnet was modified by addition of tapered Armco iron pole pieces providing a pole gap of 1¾ inches with pole diameter of 1¾ inches. Homogeneity of the field was improved by the use of annular shims ½ inch thick by 3/16 inch wide. With the shims and an available pole gap of 1½ inches the maximum available field at the center of the gap was about 3300 oersteds. Operating fields up to 4000 oersteds were obtained by using coils wound around the pole pieces. The coils were operated with a current-regulated power supply, and the field could be regulated within ±2 oersteds. The magnet was mounted in a cage which could be rotated to any desired angle relative to the ruby crystal symmetry axis.

### THE RECEIVER, MIXER, IF STRIP, AND X-BAND COMPONENTS

The receiver system has been described fully by Mayer, *et al.*<sup>10</sup> The characteristics are summarized in Table III.

As discussed in the Introduction, the noise figure of 9.8 db = 10 log F as defined for a receiver without reactive image rejection, indicates an equivalent temperature of (F–2)290°K at the input to the mixer on the basis of a 5.5-mc bandpass. Since the continuum signal information is detected over an equivalent 11-mc band-

<sup>17</sup> C. H. Townes and A. L. Schawlow, "Microwave Spectroscopy," McGraw-Hill Book Co., Inc., New York, N.Y., p. 484; 1955.

TABLE III  
RECEIVER CHARACTERISTICS

Intermediate frequency = 30 mc
IF bandwidth at 1/2 power = 5.5 mc
IF amplifier noise figure = 2.2 db
Receiver noise figure = 9.8 db

pass, the noise figure of the system is effectively  $F/2$  and the noise temperature

$$(F/2 - 1) 290^\circ\text{K}.$$

The reduction in gain fluctuation "noise" afforded by the ferrite switching technique<sup>10</sup> was found to be essential to utilize the improved sensitivity of the maser. The equivalent temperature at the input to the maser is estimated as about 47°K (excluding the ferrite switch). Maser gain variations of 10 per cent during an observation of 6 minutes, a not-uncommon occurrence, would have been disastrous in the absence of the switching, which reduced the signal from background temperature to the equivalent of a few degrees. To reduce further the effects of maser gain instability, a remotely controlled attenuator was inserted in the comparison horn waveguide permitting the antenna horn and the comparison horn to be balanced to within about 0.3°K. For this degree of balancing a ten per cent gain variation introduced a shift in the zero level of the order of the rms noise level.

RADIO ASTRONOMICAL OBSERVATIONS

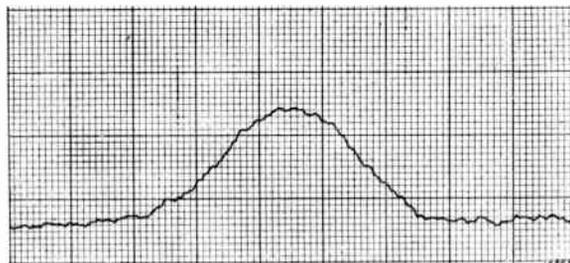
The maser radiometer has been applied to the study of discrete radio sources, in particular the planets Venus and Jupiter. A preliminary account of the early observations of Venus, Jupiter, and Virgo-A (NGC 4486) has been presented<sup>18,19</sup> and detailed results will be published elsewhere.

In Fig. 7 the performance of the maser radiometer (II) is compared with that of the radiometer without the maser. The drift scans of the intense source Cygnus-A (peak antenna temperature 4.6°K) were both obtained using an averaging time of  $\tau=5$  seconds. The rms fluctuation level for the drift scan using the maser radiometer is approximately 0.07°K in Fig. 7.

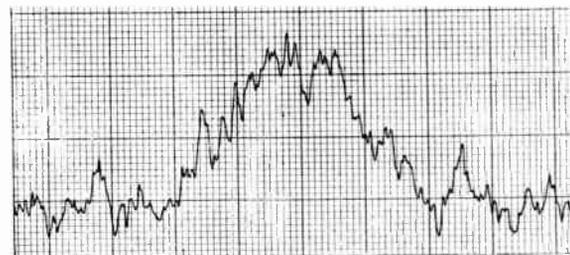
Fig. 8 is a drift curve of the radio source Virgo-A obtained using the maser radiometer (II) on August 21, 1958. The peak antenna temperature is 1.0° and the rms fluctuation level approximately 0.05°K, for an averaging time of 5 seconds.

CONCLUSION

This initial application of a maser amplifier indicates the substantial improvement in sensitivity available from maser techniques and the feasibility of using these



(a)



(b)

Fig. 7—Drift scans of the radio source Cygnus-A. The upper curve (a) was taken with the maser radiometer. The lower curve (b) was obtained by using the superheterodyne receiver without the maser preamplifier, and is reproduced through the courtesy of T. McCullough and R. Sloanaker. Peak antenna temperature 4.6°K.

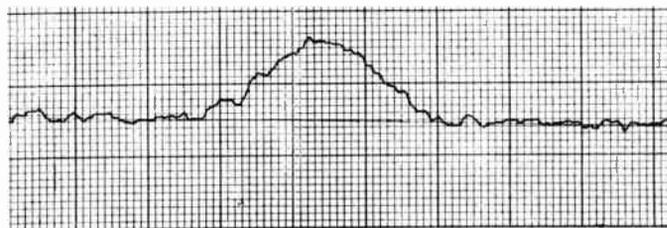


Fig. 8—Drift scan of the radio source Virgo-A obtained with the maser radiometer. Peak antenna temperature 1.0°K.

techniques in radio astronomy. The limitation on noise in such a system lies in the thermal radiation from components and environment preceding the amplifier, in particular from the ferrite switch, the circulator, and the hot ground seen by the antenna. It appears that the contribution from each of these sources can be eliminated or largely reduced, so that noise temperatures near 10°K should be obtainable in carefully designed systems. This, with masers of still wider bandwidth, should allow considerable further improvement of radiometer sensitivity in the microwave range.

ACKNOWLEDGMENT

The authors would like to thank Prof. C. Kikuchi for information about experiments with ruby before publication, and for providing the ruby used in this work. They are grateful to E. F. McClain for help in planning this cooperative program and for his support throughout the course of the work, to J. Boland for his assistance in the installation and operation of the equipment, and to I. Beller, C. Dechert, A. Marshall, and R. Sloanaker for their contributions to the design and construction of the apparatus.

<sup>18</sup> L. E. Alsop, J. A. Giordmaine, C. H. Mayer, and C. H. Townes, "Observations using a maser radiometer at 3 cm wavelength," *Astronom. J.*, vol. 63, p. 301; September, 1958.  
<sup>19</sup> IAU-URSI Symposium on Radio Astronomy, Paris, France; July 30-August 6, 1958. (To be published in the Symposium report.)

# Tantalum Printed Capacitors\*

R. W. BERRY† AND D. J. SLOAN†

**Summary**—It has previously been considered either impossible or impractical to make tantalum oxide capacitors by applying a metal counter electrode directly to a tantalum oxide film. Attempts in this direction have led to either direct shorts or low breakdown strength. For this reason, it has been presumed that it was essential to use either an electrolyte or a semiconductor in the cathode structure.

Using sputtered tantalum films as the base for the anodized oxide film, however, excellent results have been achieved employing evaporated metal counter electrodes. Many of the properties of units made in this way are superior to those of other types of tantalum capacitors.

Capacitances obtained are comparable to the capacitance-area relationships for tantalum electrolytic capacitors formed to the same voltages. DC leakages, however, have been found to be much lower than values reported for those of tantalum electrolytic units. Another advantage is that these units are capable of withstanding higher voltages than will tantalum solid electrolytic capacitors formed to the same voltage. Indeed, voltages equal to the anodizing voltage may be maintained on the capacitor without impairment.

Capacitors have also been produced with thin films of other anodizable metals, evaporated aluminum in particular.

This type of capacitor should find many applications in the lower capacitance areas, and seems ideally suited for printed circuit applications.

## INTRODUCTION

ELECTROLYTIC capacitors find their major uses when large blocks of capacitance are needed in electrical circuits. No other capacitors provide such large capacitance to volume ratios at a low cost. For lower capacitance applications, however, it has been usual to use mica, paper, ceramics, or some other type capacitor which has a relatively inefficient capacitance to volume ratio. With the present trend toward microminiaturization, lower valued capacitors which maintain the higher capacitance-volume ratio of the electrolytic capacitors should find wide application, particularly if their properties are superior to most electrolytic units available today. Also toward this aim of miniaturization, printed circuit techniques are being utilized to a greater and greater extent, and a capacitor which could be "printed" onto a circuit board, still showing reasonable capacitance values, should be a real aid.

In our laboratories, a capacitor (to be referred to here as a tantalum printed capacitor) has been developed which meets both the above requirements. These units are fundamentally different from electrolytic capacitors in that there is no electrolyte or semiconductor involved in their cathode structure; instead, the electrodes are both metallic. Whereas larger capacitance requirements

might call for a capacitor of the solid electrolytic type described previously by McLean and Power,<sup>1</sup> these capacitors should find application in the capacitance range below about 1  $\mu$ f.

## STRUCTURAL FEATURES

The units to be discussed are made by putting a thin film of tantalum on an insulating substrate, electrochemically oxidizing a portion of the tantalum to form the dielectric layer, and then applying another layer of metal on top of this oxide film. Almost all experimental tests have been performed on units having standard microscope slides as their substrate, but capacitors have also been made on glazed ceramics and other smooth surfaces. The tantalum films were produced by the technique of cathodic sputtering, using appropriate masking to give the desired patterns of tantalum. The majority of the samples were prepared using a 4-inch diameter tantalum (Fansteel capacitor grade sheet) cathode approximately 2 inches from the substrate. The slides were sputtered at a potential of 5000 v for one hour in an argon atmosphere, held at a pressure of 20 microns. This produced films whose thicknesses were of the order of 5000 Angstrom units.

A layer of dielectric tantalum oxide was electrochemically formed on the tantalum surfaces by making the tantalum film the anode in an electrolytic bath. Although various electrolytic solutions were used successfully, most samples were prepared with an electrolyte consisting of one part by weight of oxalic acid, two parts water, and three parts ethylene glycol. The temperature during anodization was maintained at 105°C, and the current kept at about 1 ma per square cm of exposed tantalum until the desired voltage was reached. The anodization was continued for from 4 to 5 hours. The thickness of the oxide, which constitutes the dielectric of the capacitor, is controlled by the formation voltage, to which it is proportional.

The second, or counter, electrode was then put on by evaporating another metal onto the slide through appropriate masks. Both gold and aluminum have been used by the authors, and no difference in properties has been observed between these two different counter electrode materials. Fig. 1 is a diagrammatic representation of a structure comprising five individual test capacitors on a single substrate. Fig. 2 is a photograph of some representative units on both microscope slides and ceramic substrates.

\* Original manuscript received by the IRE, December 24, 1958; revised manuscript received, March 5, 1959.

† Bell Telephone Labs., Inc., Murray Hill, N. J.

<sup>1</sup> D. A. McLean and F. S. Power, "Tantalum solid electrolytic capacitors," *Proc. IRE*, vol. 44, pp. 872-878; July, 1956.

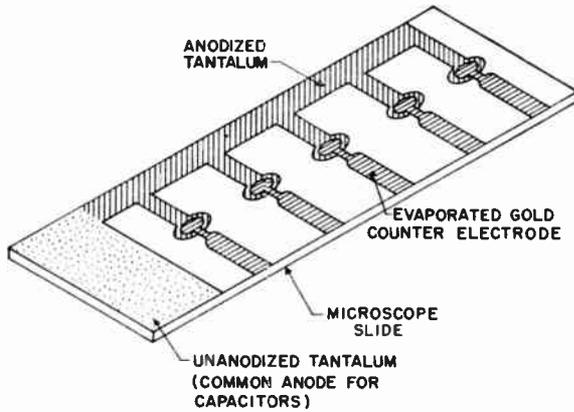


Fig. 1—Drawing showing five test capacitors on a microscope slide.

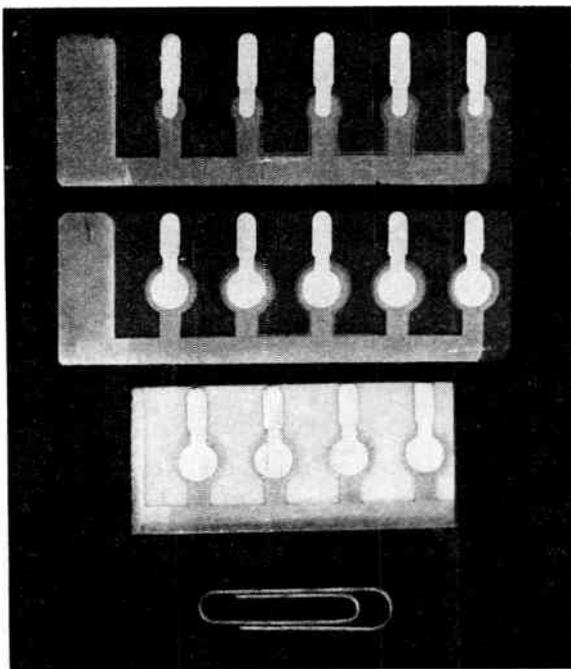


Fig. 2—Photograph of test capacitors on microscope slides and a ceramic substrate.

The capacitance values obtained with these units are, of course, proportional to the area of the electrodes, and above 50 v formation, inversely proportional to the formation voltage. Below 50 v, the capacitance is somewhat less than would be expected, probably because of the existence of the thermal oxide film initially on the samples. The following table gives approximate capacitances obtained with electrodes 250 mils in diameter for various formation voltages:

Formation Voltage	Capacitance ( $\mu\text{mf}$ )
5	250,000
10	185,000
20	92,000
40	68,000
100	30,000
150	20,000
200	15,000

These capacitors bear some resemblance to electrolytic units, but it is to be re-emphasized that no electrolyte or semiconductor is incorporated in their cathode structure. They, therefore, may be considered electrolytic only in the sense that the dielectric has been formed by standard electrolytic techniques. The use of a metal counter electrode has usually been considered impractical if not impossible, and it has generally been maintained that an electrolyte or semiconductor is required as a part of the cathode system. As discussed below, the success achieved with the structure described here probably results from the perfection of the sputtered tantalum films as compared, for example, with rolled or drawn tantalum pieces.

#### TEMPERATURE CHARACTERISTICS

These capacitors, containing no electrolyte or semiconductor, have superior temperature characteristics. The capacitance has been measured on several units at temperatures ranging from that of liquid nitrogen ( $-196^{\circ}\text{C}$ ) to  $250^{\circ}\text{C}$ . A curve illustrating the variation of capacitance with temperature is shown as Fig. 3. The temperature coefficient of capacitance is seen to remain essentially constant at  $+250$  parts per million per degree centigrade over the temperature range of  $-196^{\circ}\text{C}$  to  $+170^{\circ}\text{C}$ . Above  $170^{\circ}\text{C}$ , the temperature coefficient increases as is shown by the change in slope in Fig. 3.

The dissipation factor measured at a frequency of 1 kc shows a slight increase with temperature at temperatures up to  $170^{\circ}\text{C}$ , as shown in Fig. 4. With a further increase in temperature, the dissipation factor increases appreciably, while at lower temperatures, it should be noted that the change in dissipation factor is uniform and small to temperatures as low as that of liquid nitrogen.

The leakage current of these capacitors is considerably less than that obtained for electrolytic capacitors. Fig. 5 is a plot of leakage current vs temperature for a  $3000\text{-}\mu\text{mf}$  capacitor formed at 150 v. In this figure, two test potentials are shown, *i.e.*, 40 v and 10 v. The temperature scale is reciprocal, and the leakage current logarithmic; thus, the activation energy for conduction may be calculated, since the slope is equal to  $-E^*/k$ . The values calculated are 17.8 kcal/mole (0.77 ev) for the 40 v measurements and 17.1 kcal/mole (0.74 ev) for the 10 v measurements. A probable interpretation for the close agreement of these results is that the conduction is electronic in nature, since there is no dependence of the activation energy on potential, at least within these limits.<sup>2</sup> It should be noted here that the room temperature characteristics were unaltered by these excursions to temperature extremes.

<sup>2</sup> C. P. Bean, J. C. Fisher, and D. A. Vermilyea, "Ionic conductivity of tantalum oxide at very high fields," *Phys. Rev.*, vol. 101, pp. 551-554; January, 1956.

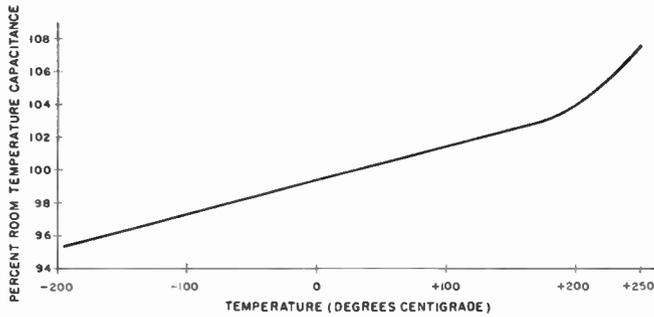


Fig. 3—Capacitance-temperature relationship typical of tantalum printed capacitors.

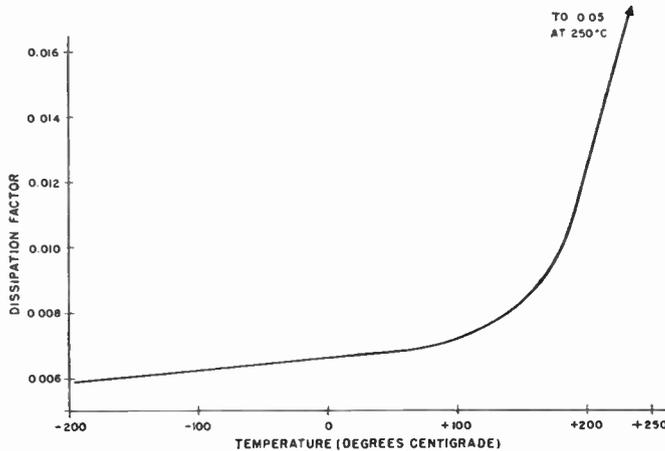


Fig. 4—Dissipation factor-temperature relationship typical of tantalum printed capacitors with 95-mil electrodes.

Room temperature leakages were measured by applying voltages to the capacitors and measuring the current flowing as a function of time. The units exhibit a long polarization time, particularly at the higher voltages. This is shown graphically in Fig. 6, which shows typical plots of the apparent resistances, as a function of time, for a 30,000- $\mu\text{mf}$  capacitor formed to 100 v when measured at different voltages. It is remarkable that although the full formation voltage is applied, (which represents a field strength of about 5 million v/cm) the apparent resistance rises to a value greater than 500,000 megohms after 4 hours. In terms of insulation resistance, this corresponds to approximately 15,000 ohm farads. At lower voltages, the apparent resistance increases, and at 60 v, it is of the order of 4,000,000 megohms, which corresponds to approximately 120,000 ohm farads. It was not possible to obtain accurate data at 50 v and below, because of insufficient sensitivity in the measuring equipment. It should be pointed out that these leakages are equal to or lower than the leakages of ordinary molded mica capacitors of the same capacitance values.

The apparent volume resistivities have been calculated from these data giving the following results at the various test potentials:

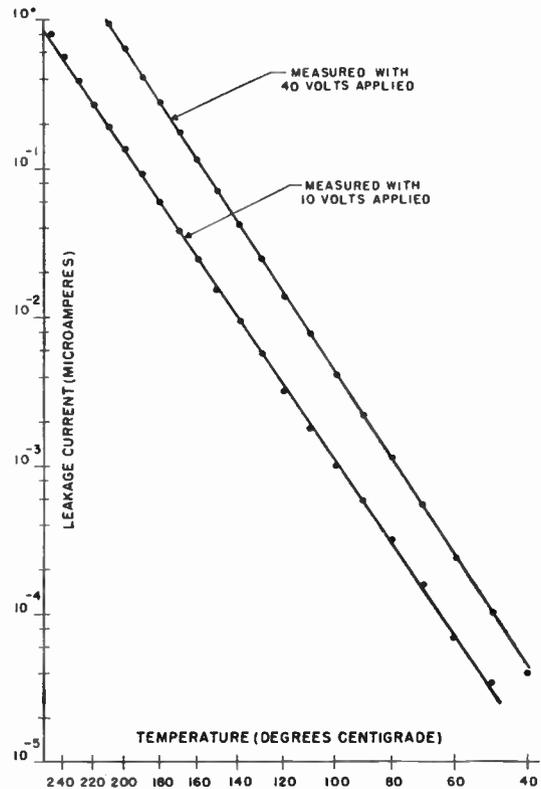


Fig. 5—Leakage current vs temperature for a 3000- $\mu\text{mf}$  tantalum printed capacitor formed at 150 v.

Potential	$\rho$ (ohm-cm)
100 v	$9 \times 10^{15}$
80 v	$2.8 \times 10^{16}$
60 v	$5.7 \times 10^{16}$

These values are in excess of any previously reported, but they probably do not represent intrinsic behavior.

In order to observe whether these units were polar, measurements were made in the following manner: Current was plotted on the y axis of an x-y recorder, while the potential, plotted on the x axis, was slowly increased. This was done both with the tantalum positive (anode) and negative. The majority of the current seen was the charging current of the capacitors, but by subtracting this, a comparison of the conduction properties in the two directions could be made. (It should be noted that this method will include any polarization current.) Such a plot with the instantaneous charging current subtracted is shown as Fig. 7. The capacitors are asymmetric, but rather than showing a sharp increase in current at very low voltages when the tantalum is cathodic, this increase in current is exhibited at anywhere from 10 to 40 per cent of the formation voltage. The actual behavior of the units is peculiar, however, in that they often may appear to be nonpolar even to voltages approaching that of formation. If allowed to remain at these high voltages, however, they usually begin to exhibit higher leakages, and a return trace shows its change in slope somewhere within the range mentioned above. It should be noted that the leakage character-

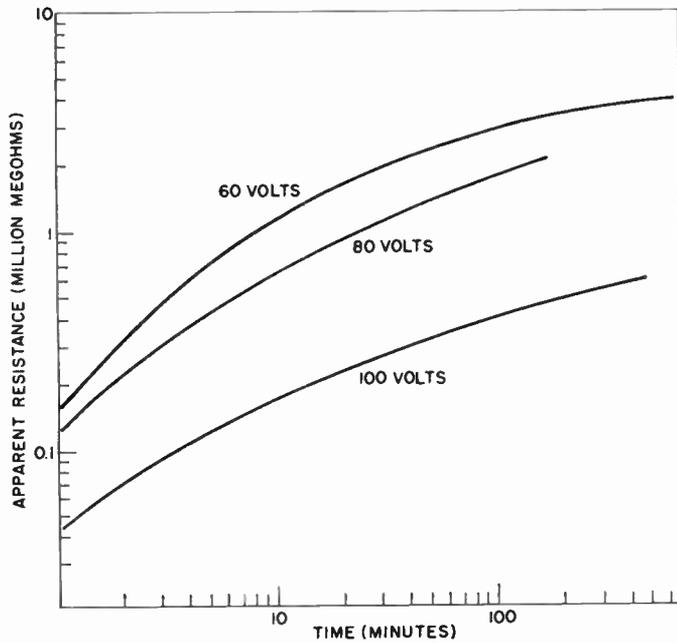


Fig. 6—Apparent resistance vs time, as measured at several potentials, for a 30,000- $\mu$ f tantalum printed capacitor formed to 100 v.

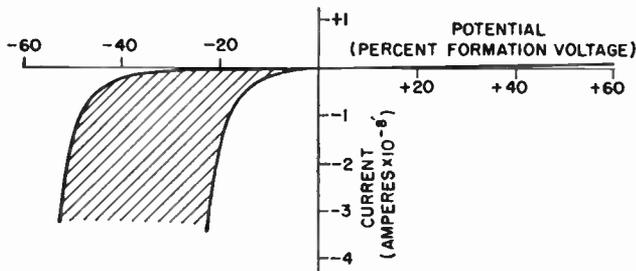


Fig. 7—Instantaneous leakage vs voltage for 95-mil tantalum printed capacitors formed to from 10 to 200 v.

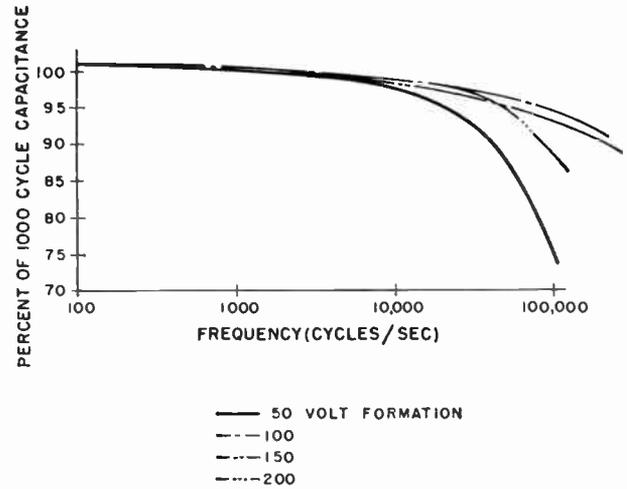


Fig. 8—Capacitance-frequency relationships for tantalum printed capacitors formed at various voltages.

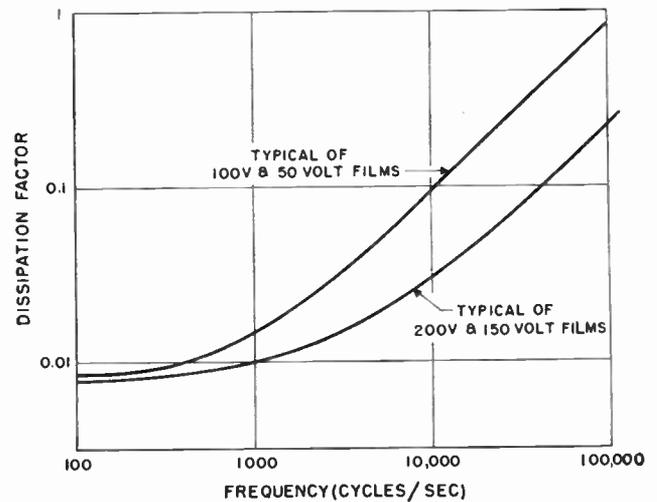


Fig. 9—Dissipation factor-frequency relationships for tantalum printed capacitors formed at various voltages.

istics with the tantalum positive are unchanged by this application of reverse potential. This indicates that a momentary application of substantial voltage in the reverse direction may be tolerated without damaging the capacitors.

In an attempt to make a nonpolar unit, a tantalum counter electrode was sputtered on, with the object of making a symmetrical capacitor. Although the capacitors produced were essentially nonpolar, their leakages were higher than the reverse leakages of the non-symmetrical units. It is believed that the sputtering process used in applying the tantalum counter electrode may have injured the oxide film, since the same high leakages are observed if a gold counter electrode be sputtered on instead of evaporated.

The breakdown strength of these capacitors approximates their formation voltage more closely than has been previously realized in truly dry tantalum units. It is not suggested, however, that the capacitors be used at these voltages. As it will be pointed out below,

operation at one half the formation voltage may well be possible at temperatures up to 65°C.

#### FREQUENCY CHARACTERISTICS

The high frequency limitations of these capacitors are less pronounced than in the case of normal aqueous electrolytic capacitors. The frequency characteristics are illustrated in Figs. 8 and 9. Fig. 8 shows several capacitance-frequency curves for different formation voltages. The pronounced decrease in capacitance at high frequency does not show up until frequencies of the order of 50 to 100 kc are reached. In Fig. 9, plots of dissipation factor vs frequency are shown. These curves show that the dissipation factor increases with frequency in such a manner as to indicate the series resistance of the units to be the major factor affecting the performance at high frequencies.

The logarithmic plots of Fig. 9 approach a 45° slope

at the higher frequencies, which is to be expected if the series resistance is the controlling factor in the dissipation factor at these frequencies. This follows from the fact that the dissipation factor for a series RC circuit is equal to  $\omega CR$ , where  $C$  is the capacitance,  $R$  the series resistance, and  $\omega$  is  $2\pi$  times the frequency. Furthermore, it will be noted that the capacitors formed at 150 and 200 v have better high-frequency characteristics than those formed at 50 and 100 v. This can be attributed primarily to the lower capacitance, the value of  $R$  remaining substantially the same. At low frequencies,  $\omega CR$  decreases, so that at 100 c, the dissipation factors of the capacitors are essentially equal, being controlled by the dissipation in the tantalum oxide film.

The high frequency characteristics can best be improved by decreasing the series resistance of the units. This may be done in a number of ways, the easiest being to shorten the length of the tantalum leads and to increase the thickness of the tantalum films. Alternatively, the tantalum might be put down on an underlying layer of another metal.

#### LIFE TESTS

It is difficult to collect life test data on new components during the period of active development, since samples used for life tests may be obsolete before the tests are complete. Nevertheless, eighteen units formed to 100 v have been on test at room temperature and at 75 v for a period of over 6 months, with no failures, and no important changes in electrical properties. At 85°C, however, with 70 v applied, eighteen of twenty-eight capacitors formed to 100 v had opened in 2 weeks. The failures apparently occurred via the process of momentary short circuits burning off the gold counter electrodes. The ten remaining capacitors did not change appreciably in any respect after 5 weeks when the test was stopped. With another group of 25 units also formed to 100 v, life testing at 85°C and 50 v is now in progress. No failures have occurred nor have any changes in electrical characteristics been observed over a 7-week period, but this is a very short time for any conclusions to be drawn. Many additional life tests must be performed on larger numbers of capacitors before any voltage ratings are assigned, although it appears that half the formation voltage would be a conservative rating up to 65°C.

#### OTHER STRUCTURES

Only single layer capacitors have been discussed thus far, but several multilayered units have been made as follows: A thin substrate (a microscope cover glass) is coated with tantalum, leaving one edge free of metal; this tantalum is then anodized, and gold evaporated on, leaving an edge of tantalum free from gold. Several of these are then stacked together, and a lead is put on each side with silver paste. The construction is shown schematically in Fig. 10. Utilizing such a structure, with substrates 10 mils thick (the approximate thickness of a

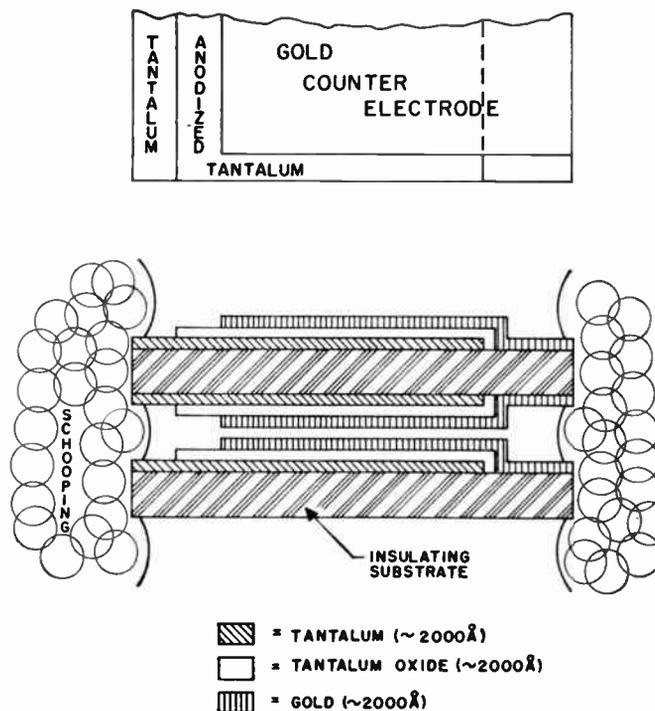


Fig. 10—Schematic representation of a multilayered capacitor.

microscope cover glass) and allowing one half the volume for margins, etc., the capacitance available per cubic inch for a 100 v formation would be in the order of 50  $\mu\text{f}$ . This could be increased substantially by decreasing the substrate thickness. Exploratory samples, only two layers thick, have a capacitance of 0.15  $\mu\text{f}$  and a volume of 0.009 cubic inch.

#### OTHER METALS

It has not previously been considered practical to make reliable capacitors by putting a metal electrode next to an electrochemically-formed oxide film. In the opinion of the authors, the limited success achieved with capacitors of this sort on bulk tantalum has probably been due to the presence of point concentrations of impurities on the surface, such as might be introduced in the mechanical forming of the metal. Indeed, the conduction properties of tantalum oxide formed on foil have been attributed to these impurity centers.<sup>3</sup> Any impurities in the sputtered films, however, would be distributed uniformly throughout the film, accounting for the much improved results reported here.

In the hope that conduction properties of anodic films of other metals might also be attributed to such centers, studies are now being made using evaporated films of aluminum and titanium, as well as sputtered films of niobium.

Good preliminary results have been obtained for capacitors made from aluminum, but the titanium and

<sup>3</sup> D. A. Vermilyea, "On the mechanism of electrolytic rectification," *J. Appl. Phys.*, vol. 27, pp. 963-964; August, 1956.

niobium films have produced units with low breakdown strength. Further work is being done with these other film-forming metals as well as with tantalum.

### CONCLUSIONS

A new tantalum capacitor, essentially two-dimensional in structure, has been made and has properties superior to other types of tantalum capacitors.

Capacitances obtained are comparable to the capacitance-area relationships for tantalum electrolytic capacitors formed to the same voltages. Using counter electrodes of 95 mils to 250 mils in diameter and the single-layered structure, capacitors have been prepared with capacitances ranging between about 2000  $\mu\text{f}$  and 0.25  $\mu\text{f}$ .

DC leakages measured at three-quarters the formation voltages are of the order of  $4 \times 10^{-11}$  a for 250-mil diameter electrodes. Expressed in terms of insulation resistance, this amounts to about 60,000 ohm farads.

Dissipation factors are in the neighborhood of 0.008 at 100 c and increase to between 0.1 and 0.8 at 100 kc.

The relatively high losses at the higher frequencies are caused by the high series resistance of the tantalum films. Thicker films of tantalum should reduce these losses.

Room temperature breakdown voltages approximate the formation voltages for these units, and successful models have been formed to 5, 10, 20, 30, 40, 50, 100, 150, and 200 v. A suggested working voltage is one half the formation voltage for temperatures up to 65°C. Voltage derating characteristics for elevated temperature operation have not been determined as yet. The units operate well at very low temperatures, however, even down to -196°C.

This type of capacitor should find many applications in the lower capacitance areas, and seems ideally suited for printed circuit applications.

### ACKNOWLEDGMENT

The authors are indebted to D. A. McLean and N. Schwartz for many helpful discussions, and to H. Basches for aid with the sputtering technique.

# Linear Diversity Combining Techniques\*

D. G. BRENNAN†, SENIOR MEMBER, IRE

**Summary**—This paper provides analyses of three types of diversity combining systems in practical use. These are: selection diversity, maximal-ratio diversity, and equal-gain diversity systems. Quantitative measures of the relative performance (under realistic conditions) of the three systems are provided. The effects of various departures from ideal conditions, such as non-Rayleigh fading and partially coherent signal or noise voltages, are considered. Some discussion is also included of the relative merits of predetection and postdetection combining and of the problems in determining and using long-term distributions. The principal results are given in graphs and tables, useful in system design. It is seen that the simplest possible combiner, the equal-gain system, will generally yield performance essentially equivalent to the maximum obtainable from any quasi-linear system. The principal application of the results is to diversity communication systems and the discussion is set in that context, but many of the results are also applicable to certain radar and navigation systems.

\* Original manuscript received by the IRE, April 21, 1958; revised manuscript received, January 14, 1959. The research reported in this paper was partly supported by the Army, Navy and Air Force under contract with the Massachusetts Institute of Technology.

† Lincoln Lab., Lexington, Mass., and Dept. of Mathematics, M.I.T., Cambridge, Mass.

### I. INTRODUCTION

WHEN a steady-state, single-frequency radio wave is transmitted over a long path, the envelope amplitude of the received signal is observed to fluctuate in time. This phenomenon is known as fading, and its existence constitutes one of the boundary conditions of radio system design. It is observed that if two or more radio channels are sufficiently separated in space, frequency, or time, and sometimes in polarization, then the fading on the various channels is more or less independent; *i.e.*, it is then relatively rare for all the channels to fade together. The standard techniques for reducing the effect of fading—known as diversity techniques—make use of this fact. The object of these techniques is to make use of the several received signals to improve the realized signal-to-noise ratio, or to improve some other performance criterion.

Several diversity combining and switching techniques are known, and there have been numerous papers on

this subject in recent years. (A sample of these, with comments, is indicated in a Bibliography at the end; these papers will be referenced by numbers in square brackets, running footnotes by superscript.) However, very few of these have provided quantitative comparative data on the relative performance of the various techniques, especially the two significant techniques (maximal-ratio and equal-gain) investigated since 1954. The major exception to this is a paper by Altman and Sichak [8], which is not widely known and even less understood.

Furthermore, there has been little attempt to explain the fundamental concepts and principles involved. For such reasons, therefore, it appeared desirable to provide an expository treatment of a comparative analysis, within a unified framework, of the three most promising diversity techniques presently known. An earlier memorandum<sup>1</sup> aimed at these objectives indicated that such a treatment might be of fairly general interest.

Of course, in an undertaking of this kind, several previously published results are naturally included as individual cases, though the available information will also be rounded out in a number of ways. Specifically, this paper includes the following material that the author has not seen published elsewhere:

1) A careful statement of the idealized circumstances required for canonical performance of coherent combiners (Section II),

2) Simple expressions for the mean signal-to-noise power ratios of various combiners [(18), (28), and (44); Fig. 8; Table I],

3) Probability distribution curves for equal-gain combiners for 3, 4, 6, and 8 channels (Figs. 10-13, Table II),

4) Estimates of the relative performance of three standard combiners for non-Rayleigh fading (Section VII),

5) A discussion of the relative performance of three standard combiners for correlated fading (Section VIII),

6) Estimates of the degradation of the average performance of equal-gain and maximal-ratio combiners caused by correlated noise voltages (Section X),

7) A bound (due to Stein) on the degradation of coherent-type combiners with imperfectly coherent signals (Section XII),

8) Certain aspects of the determination, meaning, and use of long-term distributions (Section XIII).

In addition, some previously published material has been simplified or otherwise clarified.

It should be mentioned that the criteria employed below are expressed entirely in terms of SNR. This has sometimes been taken to mean that the results were principally applicable to continuous signals, although they are also applicable to certain binary or other dis-

crete signals and can be translated into error rates once a suitable detection characteristic is either theoretically or experimentally known. But in the case of binary systems, it is possible to obtain more specific and precise results on error rates for specific systems. Such results have been extensively studied by Pierce [10], [15] and others and are not considered below. Neither is there a discussion of the considerable benefits obtainable by coding or other signal preprocessing techniques designed to counteract fading, several of which are currently under investigation by other workers.

On the other hand, it should be noted that radar and navigation systems in which a repetitive-addition signal-enhancement technique is employed are closely similar, in some respects, to certain diversity systems. Although radar and navigation systems are not discussed in detail below, many of the results and discussions set forth there are directly applicable to such systems.

## II. BASIC ASSUMPTIONS AND OTHER PRELIMINARIES

The principal background required of the reader is a basic acquaintance with certain elementary notions of probability and statistics, essentially equivalent to those developed in the first six pages of a highly readable tutorial paper given by Bennett.<sup>2</sup> No advanced techniques are required here. However, we shall make frequent use of a few ideas and techniques that were not particularly emphasized by Bennett, and a brief exposition of these is given in Appendix I. All probability distributions used in this paper will be interpreted as explained there.

We shall be concerned throughout with random variables given as functions of time (waveforms) in various intervals. In this setting, time and distribution averages are one and the same thing so  $\bar{f}$  or  $\langle f \rangle$  or  $\bar{x}$  or  $\langle x \rangle$  for such averages will be written interchangeably, but it is important to note at the outset that our averages will refer to intervals of different durations. Specifically, intervals of three different durations will be considered: 1) Short intervals, whose duration will be denoted by  $T$ . The requirement for  $T$  is that it must be short in comparison to the time required for the fading amplitude to change appreciably, but long in comparison to the period of the lowest frequency of interest in the signal. Specific representative values of  $T$  would range from a few microseconds to a few milliseconds. 2) Intermediate intervals, whose duration will be denoted by  $T_1$ . The requirements  $T_1$  must satisfy are rather complicated and will be explained at various points below. Specific suitable values of  $T_1$  would range from a few minutes to a few hours. 3) Long intervals, whose duration will be denoted by  $T_2$ . Values of  $T_2$  would range from one month to one year or more.

These intervals will be employed as follows. The short intervals of length  $T$  will be used to form "local" statis-

<sup>1</sup> D. G. Brennan, "Linear techniques in diversity communication," March, 1956. (Unpublished memorandum.)

<sup>2</sup> W. R. Bennett, "Methods of solving noise problems," *Proc. IRE*, vol. 44, pp. 609-638; May, 1956.

tics. For example, suppose  $e_1(t)$  is the instantaneous signal voltage and  $e_2(t)$  is the instantaneous noise voltage on some circuit. Then

$$x(t) = \left[ \frac{1}{T} \int_{t-T}^t [e_1(\tau)]^2 d\tau \right]^{1/2} = \sqrt{\langle e_1^2 \rangle} \quad (1)$$

and

$$y(t) = \left[ \frac{1}{T} \int_{t-T}^t [e_2(\tau)]^2 d\tau \right]^{1/2} = \sqrt{\langle e_2^2 \rangle} \quad (2)$$

would be the local rms signal and local rms noise, respectively, and  $x^2$  and  $y^2$  would be the local mean-square signal and noise voltages. Letting  $R$  denote the circuit resistance,  $x^2/R$  would be the local average signal power at time  $t$ , obtained by averaging  $e_1^2$  over the last  $T$  seconds to find  $x^2(t)$ . This averaging could be performed, for example, by feeding  $e_1^2$  into a suitable linear filter. Alternatively, one could determine the distribution of  $e_1$  in the interval  $[t-T, t]$  and obtain  $x^2(t)$  as the second moment of the distribution, though distributions in intervals of length  $T$  will not actually be of concern here.

Local statistics such as (1) and (2) will generally fluctuate in time because of fading and other effects. For example, the local rms SNR  $x(t)/y(t)$  and the local signal-to-noise power ratio

$$p(t) = \frac{x^2(t)}{y^2(t)} \quad (3)$$

will usually vary over wide limits, though they will be much better behaved than the (meaningless) instantaneous ratio  $e_1(t)/e_2(t)$ . The behavior of variables such as the local statistic (3) in intervals of length  $T_1$ , where  $T_1 \gg T$ , will be studied. In particular, various distributions and averages relative to intervals of length  $T_1$  will be considered. Such  $T_1$ -distributions and  $T_1$ -averages will also change with time, in ways discussed in Sections VII and XIII. Performance relative to  $T_1$ -intervals under standard conditions is summarized in Section VI.

Finally, the variability of certain  $T_1$ -averages will be considered in intervals of length  $T_2$ , where  $T_2 \gg T_1$ . This is done in Section XIII. It is usually assumed in system design that, for suitable values of  $T_2$ , all distributions for the system in question will be essentially the same in every corresponding interval of length  $T_2$ . (A suitable value might be one year, for example.) This is in marked contrast to the situation for  $T_1$ -distributions. However, it is found experimentally that this assumption is a reasonable first approximation; moreover, if this assumption were not satisfied, there would be no method available for predicting the performance of the system, at least at the present time.

By concentrating on system behavior relative to such prescribed lengths of intervals, it is possible to keep the relation between theory and experiment clearly visible, including, in particular, the practicable experiments required to verify theoretical predictions. This procedure is therefore vital to a complete and realistic analysis of

communication systems in general and diversity systems in particular.

In general, the term "diversity system" refers to a system in which one has available two or more closely similar copies of some desired signal. For example, certain radar systems operate by storing the signal received during one scan and adding this to the signal received during the next scan. If  $f_1(t)$  is written for the output of the storage device and  $f_2(t)$  for the signal currently being received, then the composite signal is simply  $f_1(t) + f_2(t) = f(t)$ . Now,  $f_1(t)$  may consist of a desired message component  $s_1(t)$  and an undesired additive noise component  $n_1(t)$ , so that  $f_1(t) = s_1(t) + n_1(t)$ , and similarly  $f_2(t) = s_2(t) + n_2(t)$ . Hence, the composite signal may be written

$$f(t) = [s_1(t) + s_2(t)] + [n_1(t) + n_2(t)], \quad (4)$$

*i.e.*, in the form of a resultant message component ( $s_1 + s_2$ ) plus a resultant noise component ( $n_1 + n_2$ ). If the message components  $s_1$  and  $s_2$  are closely similar, their sum  $s_1 + s_2$  will simply approximate an enlarged copy of either  $s_1$  or  $s_2$ . On the other hand, the noise components  $n_1$  and  $n_2$  may be quite dissimilar; one may be negative part of the time the other is positive, and vice versa, so they may partially cancel for part of the time. The sum (4) may then be a better signal than either  $f_1$  or  $f_2$  alone; in particular,  $f(t)$  may have a higher local SNR  $p(t)$ , defined as in (1)-(3), with  $e_1 = s_1 + s_2$ ,  $e_2 = n_1 + n_2$  than either  $f_1$  or  $f_2$  alone. Thus, one way of using two similar or suitably related copies,  $f_1$  and  $f_2$ , may be simply to add them together. Certain navigation systems in which a periodic signal is transmitted also use this storage-and-addition principle.

More generally, one may have  $N$  such copies  $f_1(t)$ ,  $f_2(t)$ ,  $\dots$ ,  $f_N(t)$ , each of the form  $f_j(t) = s_j(t) + n_j(t)$ , and one may form the sum

$$f(t) = f_1(t) + f_2(t) + \dots + f_N(t) = \sum_{j=1}^N f_j(t), \quad (5)$$

which may outperform, in some sense, the individual  $f_j$ . However, in view of the fact that the  $f_j$  will have fluctuating local statistics, it will be convenient to consider weighted sums of the  $f_j$ ; that is, the general linear combination will be considered:

$$f(t) = a_1 f_1(t) + \dots + a_N f_N(t) = \sum_{j=1}^N a_j f_j(t), \quad (6)$$

in which each  $f_j$  is weighted by a combining coefficient  $a_j$ , which is proportional to the channel gain and may be allowed to vary in accordance with the fluctuating local statistics of the  $f_j(t)$ . However, the cases to be considered will be those in which the  $a_j$  are locally constant, or at least approximately so. The adjective "linear" in the title of this paper stems from (6). Since the  $a_j$  may be allowed to vary, depending on the  $f_j$ , one should perhaps speak of (6) as locally linear or "quasi-linear." Evidently (4) is simply the case of (6) in which  $N = 2$ ,  $a_1 = a_2 = 1$ .

In diversity communication systems, there are several known methods of obtaining two or more signals  $f_j$ , and several known methods of combining these to obtain an improved signal. However, all of the latter methods in current practical use are special cases of (6). Let us first consider briefly methods of obtaining several suitable  $f_j$ . The simplest of these is that in which a single transmitting antenna furnishes a signal to several well-separated remote receiving antennas; this method is called space diversity. A variant of this, suitable for use in systems operating at UHF and above, uses two separated transmitting antennas, one of which transmits vertically polarized radiation and the other of which transmits horizontally polarized radiation, and a single receiving reflector with two feed horns or dipoles to separate the vertical and horizontal received signals. By combining these two methods, Altman and Sichak [8] obtained a fourth-order, bidirectional, full duplex space diversity system that requires only two reflectors at each end, as indicated in Fig. 1. (However, it should be added that recent experimental evidence indicates that the fading on the crossed pair of paths is more highly correlated than on the other pairs of paths.) In one form or another, space diversity has been the most commonly used form of diversity communication.

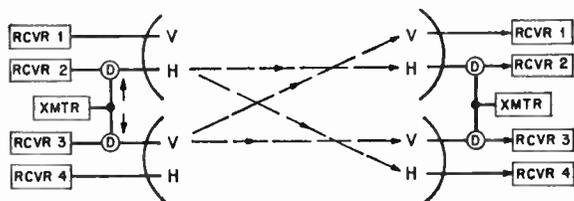


Fig. 1—Four-channel bidirectional space diversity system suitable for UHF and SHF systems. Signal paths are indicated for one direction only. The circles marked *D* denote diplexing filters. The transmitters are on different frequencies.

Another method, called frequency diversity, involves transmitting the same information on two or more carrier frequencies. If these are sufficiently separated, the fading on the various channels is approximately independent, as in the case of space diversity. This method is economical in terms of antennas and real estate, but expensive in terms of transmitters and required bandwidth. It has been discussed more often than used. (However, there are circumstances in which it is useful and has actually been used.) This is also true of the method called time diversity, so far as communication systems are concerned; however, it is not true of radar and navigation systems, as the method discussed in the opening paragraph of this section is essentially time diversity, although this terminology has not been much used in the radar field. In radio communication systems, time diversity involves transmitting the same information two or more distinct times. When this is instrumented for automatic operation, its chief disadvantage is equipment complexity; however, the simple

practice of sending each word twice, as used by many commercial CW stations, is actually a primitive but useful form of time diversity. At the other extreme, a very sophisticated communication system, currently under development,<sup>3</sup> which is designed to eliminate effects due to multiple transmission paths between fixed antennas, actually sorts out the various multipath contributions and recombines them with suitable delays and may be regarded as a form of time diversity in which the diversity is provided by the transmission medium itself.

A method that will sometimes yield two approximately independent fading signals is called polarization diversity. In normal ionospheric transmission at frequencies of a few megacycles, it is found that the received signal includes both vertically and horizontally polarized components, and the fading of these components is approximately independent.<sup>4,5</sup> However, in tropospheric transmission at UHF and above, the polarization of the transmitted signal is quite well preserved<sup>6</sup> and very little effect of this type takes place. Furthermore, even if both horizontal and vertical components are transmitted and separately received, the fading of the two components is far from independent if only a single transmission path is involved.<sup>6</sup>

Another method that has been used (infrequently) in the high-frequency region involves the combination of signals arriving with different angles of arrival (the Musa system).<sup>7</sup> A somewhat similar approach at UHF and above is currently under investigation by several workers,<sup>8-10</sup> but the efficacy of this technique is not yet firmly established.

Whichever of these methods is used, the signals obtained will initially be at radio frequency. The diversity combining techniques employed subsequent to this stage may be classed in two groups: predetection combining

<sup>3</sup> R. Price and P. E. Green, Jr., "A communication technique for multipath channels," *Proc. IRE*, vol. 46, pp. 555-570; March, 1958.

<sup>4</sup> J. L. Glaser and L. P. Faber, "Evaluation of polarization diversity performance," *Proc. IRE*, vol. 41, pp. 1774-1778; December, 1953.

<sup>5</sup> G. L. Gridale, J. G. Morris, and D. S. Palmer, "Fading of long-distance radio signals and a comparison of space- and polarization-diversity reception in the 6-18 mc range," *Proc. IEE*, pt. B, vol. 104, pp. 39-51; January, 1957.

<sup>6</sup> J. H. Chisholm, P. A. Portmann, J. T. deBettencourt, and J. F. Roche, "Investigations of the angular scattering and multipath properties of tropospheric propagation of short radio waves beyond the horizon," *Proc. IRE*, vol. 43, pp. 1317-1335; October, 1955. See especially Fig. 20, p. 1331.

<sup>7</sup> F. E. Terman, "Radio Engineers Handbook," McGraw-Hill Book Co., Inc., New York, N.Y., pp. 660-661; 1943. See also papers by Polkington and Friis and Feldman cited therein.

<sup>8</sup> R. Bolgiano, Jr., N. H. Bryant, and W. E. Gordon, "Diversity Reception in Scatter Communication with Emphasis on Angle Diversity," Cornell Univ., Ithaca, N. Y., Elec. Engrg. Res. Rep. 359; January, 1958.

<sup>9</sup> A. T. Waterman, Jr., "A rapid beam-swinging experiment in transhorizon propagation," *IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 338-340; October, 1958.

<sup>10</sup> J. H. Chisholm, L. P. Rainville, J. F. Roche, and H. G. Root, "Angular diversity reception at 2290 mcps over a 188-mile path," presented at Symp. on Extended Range and Space Communications, George Washington Univ., Washington, D. C.; October, 6-8, 1958. (Published in the Symp. Rec.)

methods and postdetection combining methods. In those methods in which, at any given time, only one of the  $a_j$  in (6) is different from zero, *i.e.*, a switch of some kind, the distinction is basically unimportant. However, important differences arise when the combining method is one in which two or more of the  $a_j$  may be different from zero at the same time. For example, it is clear that the simple addition scheme (4) can fail grievously if the message components  $s_1(t)$  and  $s_2(t)$  are not in the same phase, and RF or IF diversity signals will not usually be in the same phase unless special measures are taken to insure this. Consequently, such combining methods require special phase-control provisions when used in predetection applications, while this is not always the case in postdetection applications. An even more important difference arises in the case of FM or other bandwidth-exchange systems, where predetection combining can lead to substantial improvement over postdetection methods, as will be seen.

Once the method of providing a multiplicity of signals is decided, the basic problem confronting the designer of a diversity system becomes one of choosing the most appropriate method of combining these signals on the basis of reasonably accurate quantitative estimates of the performance of the various techniques. The balance of this paper is principally devoted to this problem. Instrumentation problems as such are not considered here; however, papers which describe certain instrumentation techniques are indicated.

We shall find it most economical to consider first a particular class of circumstances, and then indicate the way in which the results are modified by other circumstances or, in some cases, indicate where such modifications are treated elsewhere in the literature. The circumstances initially considered are those often applicable to postdetection combining in an AM system, or a single-sideband system in which provision is made for maintaining coherence of the postdetection signals.<sup>11</sup> These conditions are as follows: assume that  $N$  simultaneous functions,  $f_1(t)$ ,  $f_2(t)$ ,  $\dots$ ,  $f_N(t)$  represent the signals received in  $N$  different diversity channels as corrupted by noise and fading; each  $f_j$ ,  $j=1, 2, \dots, N$  represents the corrupted signal in the  $j$ th channel containing the originally transmitted signal  $m(t)$ . For convenience, suppose that  $m(t)$  is a steady test tone at a representative midband frequency, or some other steady test signal with a constant local mean square  $\overline{m^2}=1$ . That the following conditions are approximately satisfied is also assumed:

(A) The noise in each channel is independent of the signal, and additive:  $f_j(t) = s_j(t) + n_j(t)$  where  $s_j$  and  $n_j$  are the signal and noise components, respectively, in the  $j$ th channel.

(B) The signals  $s_j(t)$  are locally coherent; *i.e.*,  $s_j(t)$

$= x_j m(t)$ , where the  $x_j$  are positive real numbers that change with time because of fading, but at a rate that is very slow in comparison to the instantaneous variations of  $m(t)$ . More precisely, assume that the  $x_j$  do not change appreciably within any period of duration  $T$ , where  $T$  is the duration of the interval employed for the local averages. Then, since  $\overline{m^2}=1$ ,

$$\begin{aligned} \overline{s_j^2} &= \frac{1}{T} \int_{t-T}^t [s_j(\tau)]^2 d\tau = \frac{1}{T} \int_{t-T}^t x_j^2 [m(\tau)]^2 d\tau \\ &= x_j^2 \cdot \frac{1}{T} \int_{t-T}^t [m(\tau)]^2 d\tau = x_j^2, \end{aligned} \quad (7)$$

so that  $x_j = x_j(t)$  is simply the local rms value of  $s_j$ , taken over the last  $T$  seconds before the present time,  $t$ . It is clear that  $T$  must be short in comparison to the time required for the fading amplitude to change appreciably, but long in comparison to the period of  $m(t)$ .

(C) The noise components  $n_j(t)$  are locally incoherent (*i.e.*, uncorrelated) and have zero means:  $\overline{n_i n_j} = \overline{n_i} \overline{n_j}$  if  $i \neq j$ , where the duration of the averages is the same as in (7). We shall also assume that the local mean square noises  $\overline{n_j^2}$  are slowly varying, or, sometimes, constant.

(D) The local rms values of the signals,  $x_j(t) = \sqrt{\overline{s_j^2}}$ , are statistically independent. Note that this assumption automatically implies that at least two intervals are considered: first, the period  $T$  [of (7)] involved in the definition of the  $x_j$ ; and second, an interval of duration  $T_1$  in which we observe the  $x_j(t)$  as new random variables. Evidently  $T_1 \gg T$ ; in practical cases,  $T$  might be a few milliseconds and  $T_1$  approximately 30 minutes. A discussion of the requirements on  $T_1$  is provided in Section XIII. It is important that  $T_1$  cannot be too long. Assumption (D) is that, when observed in intervals with a duration on the order of  $T_1$ , the variables  $x_j(t)$  are statistically independent.

The circumstances characterized by assumptions (A)–(D) are illustrated in exploded fashion in Fig. 2 for  $N=2$ . By “exploded,” we mean that the actual signals given would be  $f_j(t) = s_j(t) + n_j(t)$ ,  $j=1, 2$ , while the  $s_j$  and  $n_j$  are shown separately. The meaning of the locally coherent assumption (B) is that, over periods of length  $T$ , the signals  $s_1$  and  $s_2$  are essentially identical except in amplitude, which is approximately constant over such periods. The local rms values  $x_j(t)$  are indicated by the dashed curves. Note that the assumption  $s_j(t) = x_j m(t)$  implies that the  $s_j(t)$  have the same zero crossings, and are in phase. If the  $s_j$  are RF or IF signals, the period  $T$  might be several microseconds or more, in which case no variation of the  $x_j$  would be perceptible within the scale of Fig. 2. If the  $s_j$  represent base-band signals,  $T$  might be a few milliseconds.

In contrast to the  $s_j$ , it will often be required that the noises  $n_j(t)$  be essentially different; this is the meaning of (C), as suggested by the waveforms  $n_1(t)$  and  $n_2(t)$  of Fig. 2. In particular, it will often be assumed that  $\overline{n_i n_j} = 0$  (if  $i \neq j$ ) over every interval of length  $T$ . In addition, however, it will sometimes also be as-

<sup>11</sup> W. E. Morrow, Jr., C. L. Mack, Jr., B. E. Nichols, and J. Leonard, “Single-sideband techniques in UHF long-range communications,” *Proc. IRE*, vol. 44, pp. 1854–1873; December, 1956.

sumed that the noises  $n_j$  have constant local average power, *i.e.*, that

$$\overline{n_j^2} = \frac{1}{T} \int_{t-T}^t [n_j(\tau)]^2 d\tau \quad (8)$$

is a constant, independent of  $t$  and  $j$ . This would be at least approximately true of the waveforms  $n_1$  and  $n_2$  of Fig. 2.

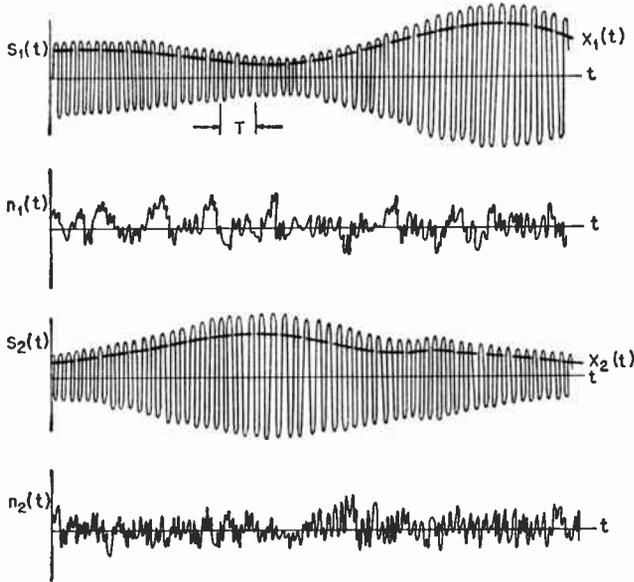


Fig. 2—Signals and noises in two diversity channels.

Assumption (D) is not particularly illustrated in Fig. 2, and could not be successfully illustrated there because the period  $T_1$  required for approximate independence of the  $x_j$  is much greater than the total scale length of Fig. 2. If the  $x_j$  were plotted throughout an interval of length  $T_1$  and the graph were then compressed to the length of Fig. 2, the resulting curves would resemble the waveforms illustrated for  $n_1$  and  $n_2$ , except that the  $x_j$  would be non-negative and would not usually be symmetric about their mean values. In particular, the  $x_j$  are not locally coherent in the sense of (B), where this “locally” refers to intervals of length  $T_1$ . Note that distributions or averages of the  $x_j$  or quantities derived therefrom, *e.g.*,  $\langle x_j^2 \rangle$ , refer to intervals of length  $T_1$ . Such averages could be distinguished by suitable notation, *e.g.*,  $\langle \rangle_1$ , but it will simplify the appearance of various expressions if the context is relied upon to make clear whether a short-term or intermediate-term average or distribution is meant.

Most of the work below is concerned with signal-noise-ratios, and from here on the word “ratio” is to mean “SNR.” This will be qualified as an amplitude ratio or a power ratio as the context requires.  $p_j = x_j^2/\overline{n_j^2}$  will be written for the local power ratio in the  $j$ th channel, and  $x_j/\sqrt{\langle n_j^2 \rangle}$ , is, similarly, the local amplitude ratio. We shall often take  $\overline{n_j^2} = 1$ ,  $j = 1, 2, \dots$ , in which

case the local amplitude ratio is simply  $x_j$  numerically, and  $p_j = x_j^2$ .

It will frequently be assumed that the variables  $x_j$  follow a Rayleigh distribution with density and distribution functions

$$p(x_j) = 2x_j e^{-x_j^2}, \quad P(x_j) = 1 - e^{-x_j^2}, \quad (9)$$

respectively. A plot of the Rayleigh density function is given in Fig. 15. All distributions considered in this paper are zero for negative values, and expressions such as (9) are to be understood as referring to positive values only. Writing the Rayleigh distribution in the form (9) implies a particular choice of scale; in particular, it implies that  $\langle x_j^2 \rangle = 1$ . The Rayleigh distribution is often written with an arbitrary scale factor, say

$$P(y) = 1 - e^{-(y/R)^2}, \quad p(y) = \frac{2y}{R^2} e^{-y^2/R^2}, \quad (10)$$

in which case  $\langle y^2 \rangle = R^2$ . However, the data below are given in a form that is completely independent of such scale factors, until Section XIII. This saves considerable cluttering of the landscape below. Similarly,  $\overline{n_j^2} = 1$  will often be taken instead of  $\overline{n_j^2} = n_0^2$ , for example. For, when  $\overline{n_j^2} = 1$ , then  $x_j$  is exactly the local amplitude ratio, which has the distribution (9), and  $p_j = x_j^2$  has the simple distribution

$$G(p_j) = 1 - e^{-p_j}, \quad g(p_j) = e^{-p_j}. \quad (11)$$

(Distribution functions will always be written with upper case letters, density functions with lower case letters.)

There are four principal types of diversity combining systems in practical use. Many of the combiners in actual use are not pure examples of one of these types; *i.e.*, they involve approximations to, or modifications of, one of these types. However, the effect of such modifications can often be estimated, at least roughly. (The terminology used here is not entirely standard—indeed, there is no generally accepted standard terminology—but is the result of careful consideration and discussion with several colleagues.) The four “pure” techniques are as follows:

1) Scanning Diversity. This technique is of the switched type; *i.e.*, at any time, only one of the  $a_j$  in (6) is different from zero, and that one is equal to 1. A selector device scans the channels in a fixed sequence until finding a signal above a preset threshold, uses that signal only until it drops below threshold, and then scans the other channels in the same fixed sequence until it again finds a signal above threshold. It is often applied to the case of two antennas supplying a single receiver through the switch, which is why it is sometimes called antenna selection diversity. It does not require a separate receiver for each channel, but at least one of the following techniques will always outperform it. We shall not consider this type in the present paper, confining our attention to the next three. This technique

has been analyzed (for  $N=2$ ) by Hausman [4] and most recently and most extensively by Henze [13].

2) Selection Diversity. This is also a switched technique, but of a more sophisticated sort. The design criterion here is that, at any given time, the system simply picks out the best of the  $N$  noisy signals  $f_1, f_2, \dots, f_N$ , and uses that one alone; the others do not then contribute to  $f(t)$ . More precisely, let  $k$  denote the index of a channel for which  $p_k \geq p_j, j=1, 2, \dots, N$ ; then this type of system is characterized by the design criterion

$$a_j = \begin{cases} 1, & \text{for } j = k, \\ 0, & \text{for } j \neq k, \end{cases} \quad (12)$$

in terms of (6). This is essentially the classical form of diversity communication [1], [2]. Very often, the selection in such systems is by electronic means (e.g., by using a common detector in such a way that the strongest signal cuts the others off) and is not quite as sharp as (12) would indicate; however, (12) is often a good approximation to such cases. A three-channel selection diversity system is depicted in Fig. 3.

3) Maximal-Ratio Diversity. This system is defined by the property that, among all systems of the type (6), it yields the maximum SNR of the output signal  $f(t)$ , provided assumptions (A), (B), and (C) are satisfied. More precisely, let  $p$  denote the local power ratio of  $f(t)$ . Then a maximal-ratio system realizes

$$p = \sum_{j=1}^N p_j; \quad (13)$$

i.e., the maximum power ratio realizable from any linear combination (6) is equal to the sum of the individual power ratios. Furthermore, the result (13) is equivalent to the requirement that the coefficients in (6) be proportional to

$$a_j = x_j/\bar{n}_j^2; \quad (14)$$

i.e., the maximum output ratio (13) is realized if and only if the gain of each channel is proportional to the rms signal and inversely proportional to the mean square noise in that channel, with the same proportionality constant for all channels. This will be proven below.

(This result has several times been quoted in the literature as requiring the weighting to be proportional to the amplitude ratio. It should be noted that this is correct only in the case where the local noise powers are all equal, in which case it would be less misleading to speak of weighting proportional to the rms signal.)

It is clear that (13) is a definite improvement over either scanning diversity or selection diversity, which can yield only one of the terms in the sum  $\sum p_j$  as the output power ratio. This observation is essentially due to Kahn [5], although the form stated here can be traced to [6], and closely similar results have been used in radar systems for some time. (It is quite possible that the diversity system discussed by Peterson, *et al.* [2]

was actually a maximal-ratio system. However, the authors made it clear that they were thinking in terms of selection diversity, whatever their actual instrumentation may have realized.) Maximal-ratio diversity has sometimes been called ratio squarer diversity, optimum diversity, and combiner diversity. Radar systems of the type discussed in connection with (4) and which employ square-law detection are essentially maximal-ratio systems. The general arrangement of a two-channel maximal-ratio system suitable for postdetection combining is shown in Fig. 4; a predetection combiner would require the addition of phase-control circuitry to satisfy assumption (B).

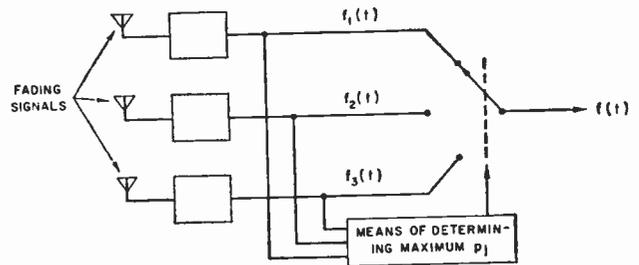


Fig. 3—Selection diversity. The  $f_j$  may be predetection or postdetection signals.

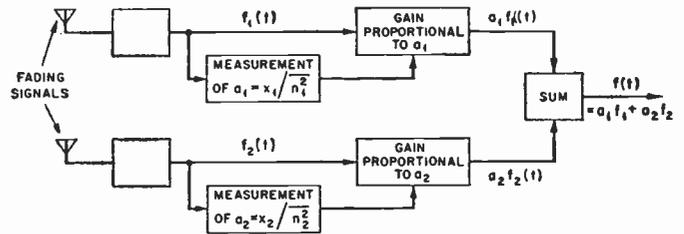


Fig. 4—Maximal-ratio diversity.

4) Equal-Gain Diversity. This is probably the simplest possible linear diversity technique; it is characterized by the property that all channels have exactly the same gain. Thus, in terms of (6),

$$a_j = 1, j = 1, 2, \dots, N, \quad (15)$$

i.e., the noisy signals  $f_j(t)$  are simply added together. In applications of this technique, the channel gains can be made to vary in such a way that the resultant signal level is approximately constant; however, this is irrelevant to the performance of the system. The important feature is that the channel gains are all equal. Note: it is important to observe that this is not the case with conventional common-detector type diversity systems; a common-detector combiner is essentially a selection diversity system, and an equal-gain system is decidedly different both in instrumentation and performance. However, an equal-gain system may well use a common AGC detector, but not a common signal detector of the usual type, as is also the case with maximal-ratio sys-

tems [5]. A basic two-channel equal-gain system is illustrated in Fig. 5. Note that the blank boxes representing receivers must have the same gains, including conversion and detection gains, which, therefore, must be fixed; they could not include separate, independent AGC systems. Also, they could not be conventional FM (or similar) receivers, as the detection gain of an FM receiver depends on the signal level. However, it is possible to instrument unconventional FM detectors for postdetection equal-gain combining.<sup>12</sup> An arrangement suitable for use with AGC is shown in Fig. 6. As in the case of Fig. 4, the application of an equal-gain combiner before detection would require the addition of phase control provisions to Figs. 5 and 6.

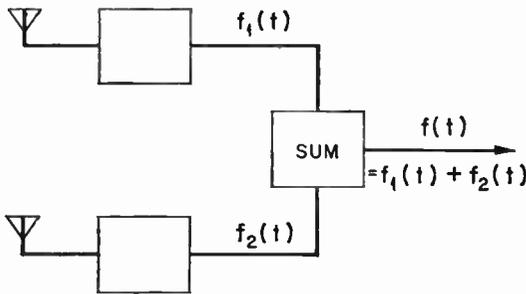


Fig. 5—Basic equal-gain diversity.

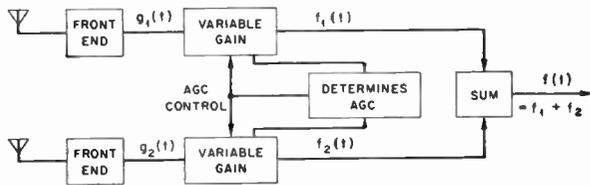


Fig. 6—Equal-gain diversity. The boxes “variable gain” must have the same gain, which may include conversion and detection gain.

It has been pointed out by Sichak [8]<sup>13</sup> that, under conditions often occurring in practice, equal-gain systems will outperform selection diversity, and will perform almost as well as maximal-ratio systems. In view of the simplicity of the instrumentation required for (15) as compared to (14) (equivalently, see Figs. 5 and 4), this fact is of great practical importance. The conditions required are that assumptions (A)–(D) must be satisfied; in addition, the  $x_j$  must be Rayleigh-distributed, and the local mean square noises  $\overline{n_j^2}$  must be approximately constant. Under other conditions, it may not perform as well as selection diversity, as will be seen. Since, however, these conditions are often (approximately) satisfied, it follows that equal-gain diversity should be more widely known than is presently the case.

In the following three sections, the principal features of diversity combiners of types 2)–4) above will be de-

veloped. In particular, distribution functions will be obtained for the local power ratio  $p$  of the composite signal  $f(t)$ , and mean values of  $p$ , under the conditions discussed. Then, in Sections VII to XII the results will be compared and evaluated, and the way in which the results are altered by various modifications of the conditions as they occur in practice will be indicated.

### III. SELECTION DIVERSITY

The distribution function for an  $N$ -channel selection diversity system is particularly simple to obtain, provided the local noise powers  $\overline{n_j^2}$  are constant. Let  $\overline{n_j^2} = 1, j = 1, 2, \dots, N$ , and assume that the  $x_j$  are Rayleigh-distributed. Then the individual channel power ratios  $p_j$  have the distribution  $G(p_j)$  of (11). By (12), the output power ratio of the combiner is simply the largest of the individual  $p_j$ . Now, if the largest power ratio is  $\leq p$ , then the power ratio of every channel is  $\leq p$ ; conversely, if the power ratio of every channel is  $\leq p$ , then so is the power ratio of  $f$ . Hence, the probability of having the power ratio of  $f$  be  $\leq p$  is precisely the probability of having the individual channel ratios all  $\leq p$  simultaneously. Since the  $x_j$  are independent, by (D), so are the  $p_j = x_j^2$ , and, hence, the probability that all channels have power ratio  $\leq p$  is simply the product of the separate probabilities that each channel individually has a power ratio  $\leq p$ . Thus,

$$S_N(p) = G(p) \cdot G(p) \cdot \dots \cdot G(p) = [G(p)]^N \quad (16)$$

$$= (1 - e^{-p})^N$$

is the distribution function of  $p$ , the realized local power ratio, for an  $N$ -order selection diversity system.

The average value  $\bar{p}$  of  $p$  will be required for this system. In this case, it is most easily obtained from the distribution (16). Thus,

$$\bar{p}(N) = \int_{-\infty}^{\infty} p dS_N(p) = \int_0^{\infty} p \cdot N(1 - e^{-p})^{N-1} e^{-p} dp \quad (17)$$

This integral is evaluated in Appendix VI, where it is shown to reduce to the remarkably simple form

$$\bar{p}(N) = \sum_{k=1}^N \frac{1}{k} \quad (18)$$

Thus,  $\bar{p}(2) = 1 + \frac{1}{2} = \frac{3}{2}$ ,  $\bar{p}(3) = 1 + (\frac{1}{2}) + (\frac{1}{3}) = 11/6$ , etc.; these values will be used below. It is clear at once from (18) that increasing the number of channels in a selection diversity system yields rapidly diminishing returns; adding the  $N$ th channel increases  $\bar{p}$  by only  $1/N$ . It will be seen that the next two systems to be considered can perform much better in this respect, in consequence of (B) and (C). However, it should be noted here that neither the functioning of a selection diversity system nor the statistics developed in this section depend on assumptions (B) or (C), which are not required here. The significance of this fact will be discussed.

<sup>12</sup> R. T. Adams, private communication; December, 1958.

<sup>13</sup> W. Sichak, Fed. Telecommun. Labs., Nutley, N. J., private communication; August 19, 1955.

IV. MAXIMAL-RATIO DIVERSITY

The first order of business is to establish (13) and (14). In order to do this, it will be convenient to use a mathematical device known as the Schwarz inequality. This is not specifically related to statistics, but is a general result of great importance in many fields of pure and applied mathematics. One form of this states that if  $u_1, u_2, \dots, u_N$  are any  $N$  real numbers and  $v_1, v_2, \dots, v_N$  are any  $N$  real numbers, then

$$\left[ \sum_{j=1}^N u_j v_j \right]^2 \leq \left[ \sum_{j=1}^N u_j^2 \right] \left[ \sum_{j=1}^N v_j^2 \right]. \tag{19}$$

The proof of this, which is quite short, is given in Appendix II. Note that if  $u_j = a_j \sqrt{\langle n_j^2 \rangle}$ ,  $v_j = x_j / \sqrt{\langle n_j^2 \rangle}$ , then (19) takes the form

$$\left[ \sum_{j=1}^N a_j x_j \right]^2 \leq \left[ \sum_{j=1}^N a_j^2 \overline{n_j^2} \right] \left[ \sum_{j=1}^N (x_j^2 / \overline{n_j^2}) \right], \tag{20}$$

which, since  $p_j = x_j^2 / \overline{n_j^2}$ , can be written

$$\left[ \sum_{j=1}^N a_j x_j \right]^2 \leq \left[ \sum_{j=1}^N a_j^2 \overline{n_j^2} \right] \left[ \sum_{j=1}^N p_j \right]. \tag{21}$$

Now, in (6), let us write

$$s(t) = \sum_{j=1}^N a_j s_j(t), \quad n(t) = \sum_{j=1}^N a_j n_j(t), \tag{22}$$

so that  $f(t) = s(t) + n(t)$ , and  $p = \overline{s^2} / \overline{n^2}$  is the local power ratio of  $f$ . But (all sums from 1 to  $N$ )

$$\overline{s^2} = \langle [\sum a_j s_j]^2 \rangle$$

and by assumption (B)

$$\begin{aligned} &= \langle m^2 [\sum a_j x_j]^2 \rangle \\ &= \langle m^2 \rangle \cdot [\sum a_j x_j]^2 \end{aligned}$$

and since  $[\sum a_j x_j]^2$  is locally constant and can be taken outside the average, and since  $\overline{m^2} = 1$ ,

$$\overline{s^2} = [\sum a_j x_j]^2, \tag{23}$$

Furthermore,

$$\begin{aligned} \overline{n^2} &= \langle [\sum a_j n_j]^2 \rangle \\ &= \sum a_j^2 \overline{n_j^2}, \end{aligned} \tag{24}$$

by (61) and assumption (C). Thus, using (21),

$$p = \frac{\langle s^2 \rangle}{\langle n^2 \rangle} = \frac{[\sum a_j x_j]^2}{\sum a_j^2 \overline{n_j^2}} \leq \sum p_j, \tag{25}$$

which proves that  $p$  cannot exceed  $\sum p_j$ . On the other hand, if  $a_j = x_j / \overline{n_j^2}$ , then

$$\begin{aligned} p &= \frac{[\sum x_j^2 / \overline{n_j^2}]^2}{\sum x_j^2 / \overline{n_j^2}} = \frac{[\sum p_j]^2}{\sum p_j} \\ &= \sum p_j, \end{aligned} \tag{26}$$

so that  $p = \sum p_j$ , if  $a_j = x_j / \overline{n_j^2}$ , and similarly if  $a_j$

$= k(x_j / \overline{n_j^2})$  for any  $k \neq 0$ , thus proving (13) and (14). [Readers acquainted with the Schwarz inequality for complex numbers will recognize that it may be used to include the case of positive or negative  $x_j$ , or even complex  $x_j$ . This is, however, only a more formal way of including assumption (B).]

It is interesting to note that the only purely statistical fact used in this development is that averaging is a linear operation, as discussed in Appendix I. No use was made of distribution functions or any similar apparatus. However, it is important to observe that each of the assumptions (A)-(C) entered in a very vital way.

We now consider the statistical properties of the local power ratio  $p$ . The first point to be noticed is that  $p = \sum p_j$  implies

$$\overline{p} = \sum_{j=1}^N \overline{p}_j, \tag{27}$$

without regard to the distribution of the  $p_j$  or the possible dependence of these variables. If, in particular,  $\overline{p}_j = 1, j = 1, 2, \dots, N$ , then

$$\overline{p}(N) = N. \tag{28}$$

This behavior is in marked contrast to the corresponding relationship (18) for selection diversity, which increases much less rapidly with  $N$  than (28) does.<sup>14</sup> (It should be clear that the notation  $\overline{p}(N)$  is used in a somewhat flexible way:  $\overline{p}(N)$  is a different function of  $N$  in (18) than it is here.) The average value of the local power ratio of the output of an  $N$ -order maximal-ratio system is simply  $10 \log_{10} N$  db above a single channel.

In order to obtain an explicit distribution of  $p$ , we shall employ the same assumptions used for the selection diversity case, namely, that the  $x_j$  are independent Rayleigh variables and that  $\overline{n_j^2} = 1, j = 1, 2, \dots, N$ , so that the  $p_j$  have the exponential distribution (11). Thus we are interested in the distribution of the sum  $p = \sum p_j$  of  $N$  independent random variables, each with the distribution (11). This problem can be treated by a simple application of characteristic functions,<sup>2</sup> as indicated in Appendix III. Alternatively, it can easily be solved by using the geometric approach mentioned in Appendix I, without reference to characteristic functions. (One integrates the joint density function

$$\exp [ - (p_1 + p_2 + \dots + p_N) ]$$

over the  $N$ -dimensional volume bounded by the hyperplane  $p_1 + p_2 + \dots + p_N = p$  and the coordinate hyperplanes.) In either case, the result, writing  $G_N(p)$  for the desired distribution function and  $g_N(p)$  for the associated density function, is

$$g_N(p) = \frac{1}{(N-1)!} p^{N-1} e^{-p}, \tag{29}$$

$$G_N(p) = \frac{1}{(N-1)!} \int_0^p y^{N-1} e^{-y} dy. \tag{30}$$

<sup>14</sup> For large  $N$ , (18) is approximately  $\log_e N$ .

By using  $G_1(p) = 1 - e^{-p}$  and the recursion relation  $G_N(p) = G_{N-1}(p) - g_N(p)$ , easily verified by an integration by parts, we have

$$\begin{aligned} G_2(p) &= 1 - (1 + p)e^{-p}, \\ G_3(p) &= 1 - \left(1 + p + \frac{p^2}{2!}\right)e^{-p}, \\ G_4(p) &= 1 - \left(1 + p + \frac{p^2}{2!} + \frac{p^3}{3!}\right)e^{-p}, \end{aligned}$$

and, in general,

$$G_N(p) = 1 - \left(\sum_{k=0}^{N-1} \frac{p^k}{k!}\right)e^{-p}, \quad (31)$$

which can also be written

$$G_N(p) = \left(\sum_{k=N}^{\infty} \frac{p^k}{k!}\right)e^{-p}. \quad (32)$$

The utility of the form (32) is that it indicates the approximations

$$G_N(p) \cong \frac{p^N}{N!}e^{-p} \cong \frac{p^N}{N!} \quad (33)$$

are accurate for sufficiently small  $p$ . The distribution (31), known as the gamma distribution, is easily computed for the integral values of  $N$  of interest here and has also been tabulated.<sup>15</sup> It can also be identified with the chi-square distribution with  $2N$  degrees of freedom.<sup>16</sup>

The origin of the maximal-ratio distribution (31) has sometimes been incorrectly attributed (in Pierce [15] and Packard [14] among others). That (31) is the distribution function of sums of squares of Rayleigh variables has been known in radar circles for quite some time. In the context of maximal-ratio diversity combiners, the result (31) for arbitrary  $N$  was first published in March, 1956, by Altman and Sichak.<sup>17</sup> (It also appeared independently in an unpublished memorandum<sup>1</sup> at about the same time.) Curves of (31) for several values of  $N$  were subsequently published by Staras [9].

## V. EQUAL-GAIN DIVERSITY

Recalling that the relations  $\bar{s}^2 = [\sum a_j x_j]^2$  of (23) and  $\bar{n}^2 = \sum a_j^2 \bar{n}_j^2$  of (24) did not depend on a choice of the

<sup>15</sup> K. Pearson (ed.), "Tables of the Incomplete  $\Gamma$  Function," Cambridge University Press, Cambridge, Eng.; 1946. In his notation, he tabulates

$$I(u, p) = \frac{1}{p!} \int_0^u \int_0^{\sqrt{p+1}} v e^{-t} dt,$$

so that his  $p$  is here  $N-1$ , and his  $u$  is here  $p/\sqrt{N}$ .

<sup>16</sup> E. S. Pearson and H. O. Hartley (eds.), "Biometrika Tables for Statisticians," Cambridge University Press, Cambridge, Eng., Table 7, p. 122 ff.; 1954. Short tables of the  $\chi^2$  distribution are given in many other statistical tables and in most textbooks on statistics. See, for example, H. Cramér, "Mathematical Methods of Statistics," Princeton University Press, Princeton, N. J.; 1946.

<sup>17</sup> Altman and Sichak, [8], p. 55, middle of right-hand column, with  $(aL)^2 = p$ .

$a$ , [i.e., they hold for any combiner of the type (6), provided assumptions (A)-(C) are satisfied, and hence hold in particular for  $a_j = 1, j = 1, 2, \dots, N$ ], we have

$$\bar{s}^2 = \left[ \sum_{j=1}^N x_j \right]^2 \quad (34)$$

and

$$\bar{n}^2 = \sum_{j=1}^N \bar{n}_j^2 \quad (35)$$

for an equal-gain system. The relation  $\sqrt{\bar{s}^2} = \sum x_j$  from (34) is simply the well-known fact that the rms value of a sum of coherent signals is equal to the sum of the individual rms values, while (35) similarly expresses the fact that the average power of a sum of uncorrelated signals is equal to the sum of the individual average powers. Some communication engineers express (34) by saying that coherent signals "add linearly"; however, this language is both formally meaningless and conducive of an imperfect understanding of the situation and is better replaced by "add coherently" if some such expression is necessary.

From (34) and (35), we have

$$b = \frac{[\sum x_j]^2}{\sum n_j^2} \quad (36)$$

for an equal-gain system. In order to develop comparative statistics for this, it is again assumed that the  $\bar{n}_j^2 = 1, j = 1, 2, \dots, N$ , and that the  $x_j$  are independent and Rayleigh-distributed. Since  $\bar{n}_j^2 = 1, p = [\sum x_j]^2/N$ . Put

$$u = \sqrt{N}b = \sum_{j=1}^N x_j. \quad (37)$$

It is clear that the distribution function of  $p$  will follow immediately from that for  $u$ . The distribution of a sum of  $N$  Rayleigh variables, each with the distribution (9), is accordingly of interest. Unfortunately, this problem is not nearly as tractable as in the maximal-ratio case. The characteristic function of a Rayleigh variable is not expressible in an immediately useful form. We are here essentially forced to rely on the geometric approach mentioned in Appendix I. Let  $B_N(u)$  denote the distribution function of (37). For  $N=2$ , say  $u = x + y$ , it can be seen that  $B_2(u)$  is given by the integral of the joint density function  $4xye^{-(x^2+y^2)}$  over the region of the  $x-y$  plane bounded by the line  $x+y=u$  and the coordinate axes (Fig. 7). (We can stay in the first quadrant because the density function is zero in all other quadrants.) It is easy to see that this is

$$\begin{aligned} B_2(u) &= \int_0^u \int_{x=0}^{x=u-y} 4xye^{-(x^2+y^2)} dx dy \\ &= 2 \int_0^u ye^{-y^2} [1 - e^{-(u-y)^2}] dy. \end{aligned} \quad (38)$$

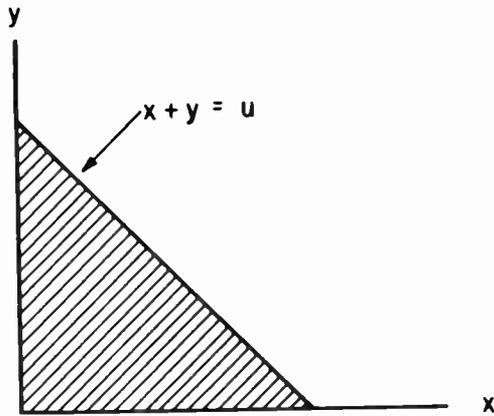


Fig. 7—Region of integration for (38).

By completing the square in the exponent and making a few other routine manipulations, this becomes [8]

$$B_2(u) = 1 - e^{-u^2} - (\sqrt{\pi}/2)ue^{-u^2/2}I\left(\frac{u}{\sqrt{2}}\right), \quad (39)$$

where

$$I(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

is the error function and is tabulated.<sup>18</sup> Thus the distribution function of  $p = u^2/2$ ,  $A_2(p)$ , is  $B_2(\sqrt{2p})$ , i.e.,

$$A_2(p) = 1 - e^{-2p} - \sqrt{p}pe^{-p}I(\sqrt{p}), \quad (40)$$

and is readily plotted.

Corresponding to (38), it is easy to see that the distribution function of the sum of  $N$  independent Rayleigh variables is

$$B_N(u) = 2^N \int_0^u \int_0^{u-x_N} \dots \int_0^{u-\sum_{k=3}^N x_k} \int_0^{u-\sum_{k=2}^N x_k} x_1 x_2 \dots x_N e^{-(x_1^2 + \dots + x_N^2)} dx_1 \dots dx_N, \quad (41)$$

which is simply the integral of the joint density function over the  $N$ -dimensional volume bounded by the hyperplane  $x_1 + x_2 + \dots + x_N = u$  and the coordinate hyperplanes, as in Fig. 7 for  $N=2$ . Unfortunately, the integral (41) is quite as frightful as it appears; numerous workers—going back to Lord Rayleigh himself—have tried to express  $B_N(u)$  in terms of tabulated functions, but with no success if  $N \geq 3$ . However,  $B_N(u)$  has recently been tabulated,<sup>19</sup> and curves of  $A_N(p) = B_N(\sqrt{Np})$

have been constructed from these tables for the present paper (in Section VI) for  $N=2, 3, 4, 6$ , and  $8$ . An outline of the method of computation is sketched in Appendix IV.

We shall next obtain the average values  $\bar{p}(N)$  for the equal-gain case. Although these averages depend on the distribution of  $p$  and the distribution of  $p$  is not given in a particularly explicit form, it is nevertheless easy, following Appendix I, to obtain the  $\bar{p}(N)$ . Since  $\bar{n}_j^2 = 1$ , (36) becomes

$$p = \frac{1}{N} \left[ \sum_{j=1}^N x_j \right]^2 = \frac{1}{N} \left[ \sum_{j=1}^N x_j^2 + \sum_{i \neq j} x_i x_j \right]. \quad (42)$$

Since the  $x_j$  are independent,  $\overline{x_i x_j} = \bar{x}_i \bar{x}_j$  if  $i \neq j$ , so

$$\begin{aligned} \bar{p} &= \frac{1}{N} \left[ \sum_{j=1}^N \overline{x_j^2} + \sum_{i \neq j} \bar{x}_i \bar{x}_j \right] \\ &= 1 + \frac{1}{N} \sum_{i \neq j} \bar{x}_i \bar{x}_j, \end{aligned} \quad (43)$$

using the fact that  $\overline{x_j^2} = 1$ ,  $j=1, 2, \dots, N$ . Let  $\bar{x}_j = r$ ,  $j=1, 2, \dots, N$ . By considering the terms of the sum  $\sum_{i \neq j} \bar{x}_i \bar{x}_j$  as the entries from an  $N$  by  $N$  matrix with the main diagonal deleted, it is seen that there are  $N^2 - N = N(N-1)$  such terms, each equal to  $r^2$ , and so

$$\bar{p}(N) = 1 + (N-1)r^2, \quad (44)$$

the desired average value. The constant  $r^2 = (\bar{x}_j)^2$  depends on the distribution of the  $x_j$ . For the Rayleigh distribution,  $r^2 = \pi/4 \approx 0.785$ . For any distribution,  $r^2 = (\bar{x})^2/\overline{x^2}$  is a dimensionless constant between 0 and 1, but values of  $r^2$  much less than 0.785 are relatively infrequent in observed fading distributions.

It is thus seen that  $\bar{p}$  increases linearly with  $N$ , as was also the case for maximal-ratio systems. The only difference is that (28) increases with slope 1 while (44) increases with slope  $r^2 = \pi/4$  for Rayleigh fading. But the absolute maximum by which (28) can exceed (44) is  $10 \log_{10}(4/\pi) = 1.05$  db, and this only in the limit of an infinite number of channels.

### VI. CANONICAL ONE-HOUR PERFORMANCE

The three systems will first be compared simply on the basis of the average values of the local power ratio  $p$  of the output. This is done first in Fig. 8, for  $N=1, 2, 3, \dots, 10$  channels. The maximal-ratio points are values of  $10 \log_{10} N$  from (28), the equal-gain points are  $10 \log_{10} [1 + (N-1)(\pi/4)]$  from (44), and the selection diversity values are

$$10 \log_{10} \left[ \sum_{k=1}^N (1/k) \right]$$

from (18). Since  $\bar{p}_j = 1$ , these give the increase in decibels in the average local power ratio over a single channel. The data of Fig. 8 for  $N=2, 3, 4, 6$ , and  $8$  are presented from a different point of view in Table I, which gives

<sup>18</sup> "Tables of the Error Function and its Derivative," Natl. Bur. Stand., Washington, D. C., Appl. Math. Ser. No. 41; 1954. The error function or, for that matter, (39) itself, can also be expressed in terms of the much tabulated normal (Gaussian) probability distribution function. See "Tables of the Normal Probability Functions," Natl. Bur. Stand., Appl. Math. Ser. No. 23; 1953. Brief tables of the normal distribution function appear in most statistical texts, e.g., Cramér, *op. cit.*

<sup>19</sup> W. C. Mason, M. Ginsburg, and D. G. Brennan, "Tables of the distribution functions of sums of Rayleigh variables," to be published.

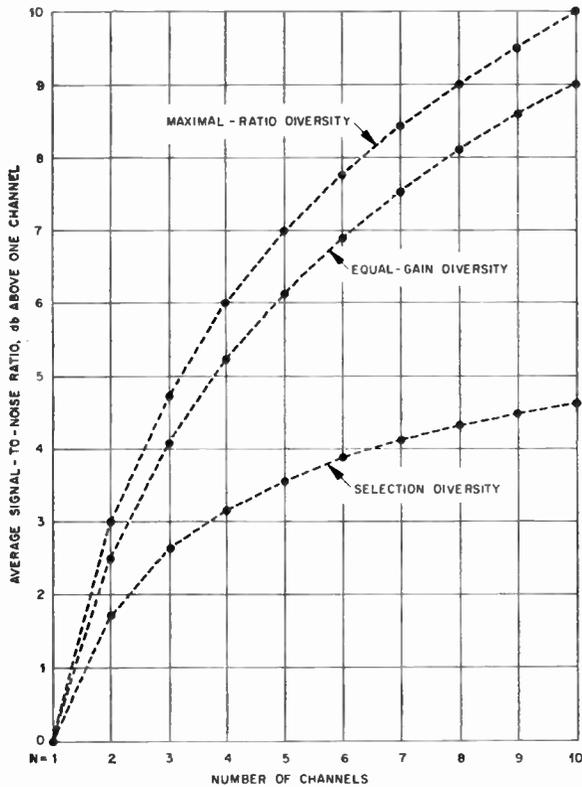


Fig. 8—Diversity improvement (in decibels) in average SNR, for independently fading Rayleigh-distributed locally coherent signals in locally incoherent noises with constant local rms values.

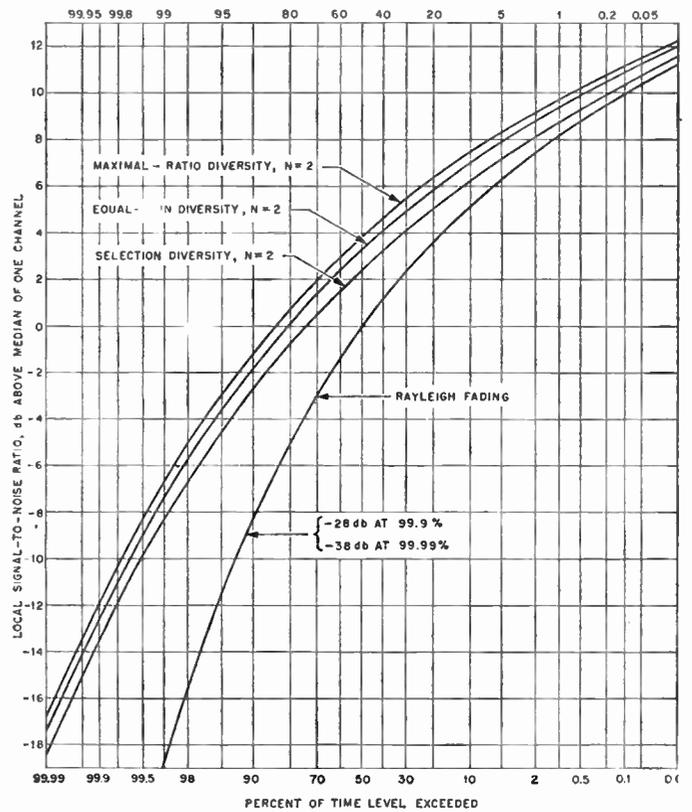


Fig. 9—Dual diversity distributions, conditions of Fig. 8.

TABLE I

COMPARATIVE AVERAGE SNR (SAME CONDITIONS AS IN FIG. 8)

Number of Channels <i>N</i>	Number of DB by which Maximal-Ratio Exceeds		
	Equal-Gain	Selection	One Channel
2	0.49	1.25	3.01
3	0.67	2.14	4.77
4	0.76	2.83	6.02
6	0.85	3.89	7.78
8	0.90	4.69	9.03
∞	1.05	∞	∞

the differences between the maximal-ratio values and the lower curves of Fig. 8, counting the zero axis as a curve. The last entry in the equal-gain column is essentially the assertion that no matter how far the curves of Fig. 8 were continued, the top two would never differ by more than 1.05 db, although they would get farther and farther away from the selection diversity curve and the base axis.

A brief discussion of the significance of these data is in order. These results are useful, for example, in estimating relative average system capacities, or in other circumstances where the average value alone of  $p$  is of interest. Most recent diversity systems have been designed for a specified percentage of reliability, *i.e.*, a

specified percentage of time during which the system performance will exceed some given criterion. This requires information about probability distributions. This approach is appropriate whenever high reliability is a primary requisite, *e.g.*, in important military communication systems, or in relay systems carrying commercial television programs. However, it should be pointed out here that some systems do not require very high local reliability or they may effectively achieve it by other means, such as coding. In such circumstances, the data of Table I may be more meaningful than results based on the distributions to be presented.

Let us next compare the probability distributions of  $p$  realized by the three systems for different orders of diversity. The case  $N=2$  (dual diversity) is illustrated in Fig. 9, together with the distribution of  $p$  for a single channel with Rayleigh fading for comparison. The term "median" in the designation of the ordinate scale of Fig. 9 refers to a value  $x_0$  of a random variable  $x$  for which  $P(x_0) = \frac{1}{2}$ , *i.e.*, a value  $x_0$  for which  $x \leq x_0$  for 50 per cent of the time and  $x \geq x_0$  for 50 per cent of the time. Thus, the median  $p_0$  of the one-channel distribution (11) is obtained by setting  $G(p_0) = \frac{1}{2}$  and solving for  $p_0$ , from which  $p_0 = \log_e 2 \approx 0.693$ . The ordinate scale of Fig. 9 is expressed in decibels relative to this  $p_0$ . That is, the  $N=2$  curves of Fig. 9 are plots of  $10 \log_{10}(p/p_0)$  vs  $100 [1 - D_2(p)]$ , where  $D_2(p)$  is, respectively,  $G_2(p)$  of (31) (maximal-ratio),  $A_2(p)$  of (40) (equal-gain) and

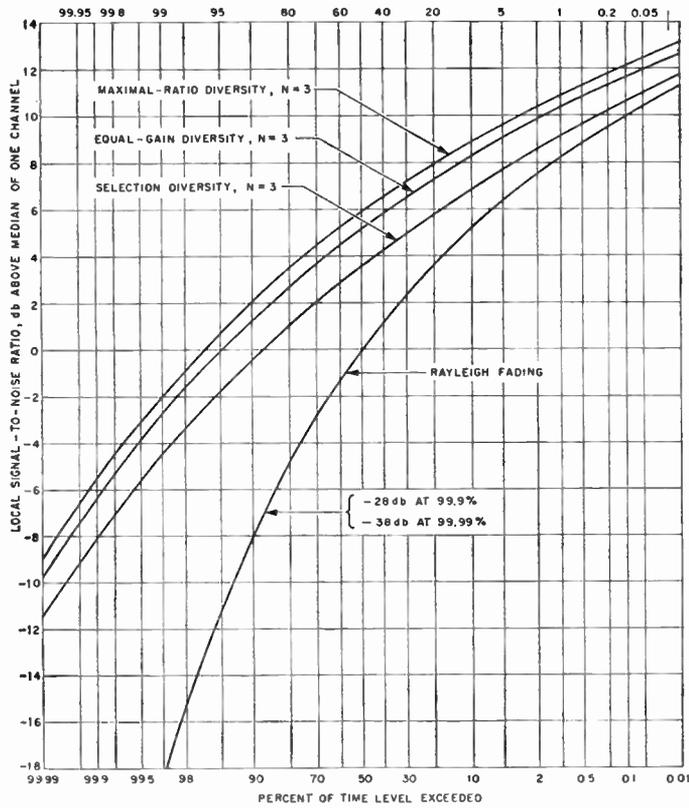


Fig. 10— $N=3$ .

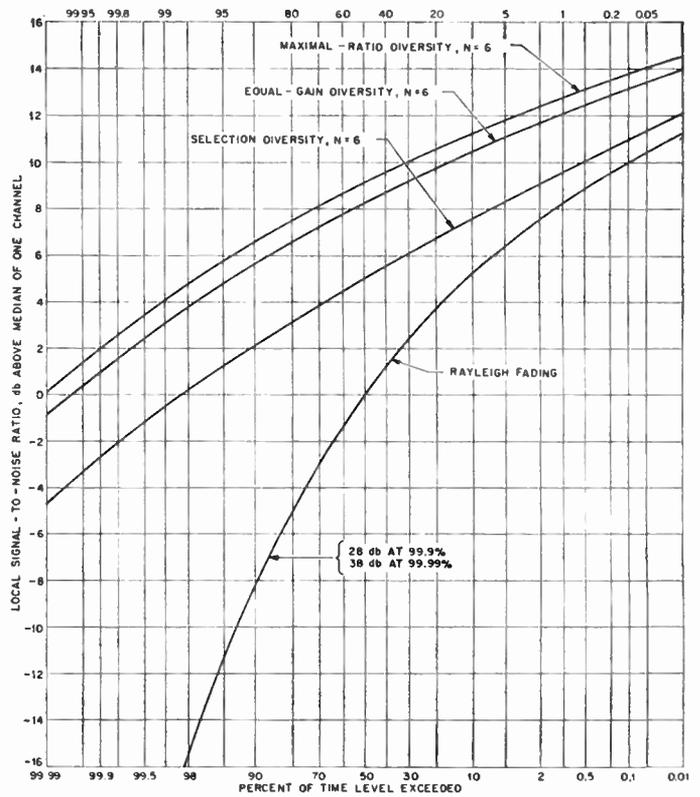


Fig. 12— $N=6$ .

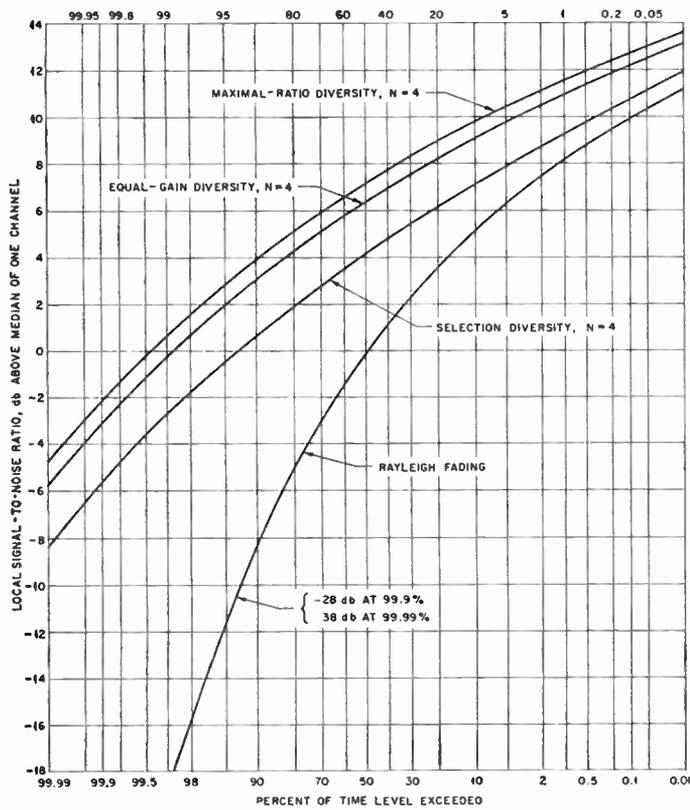


Fig. 11— $N=4$ .

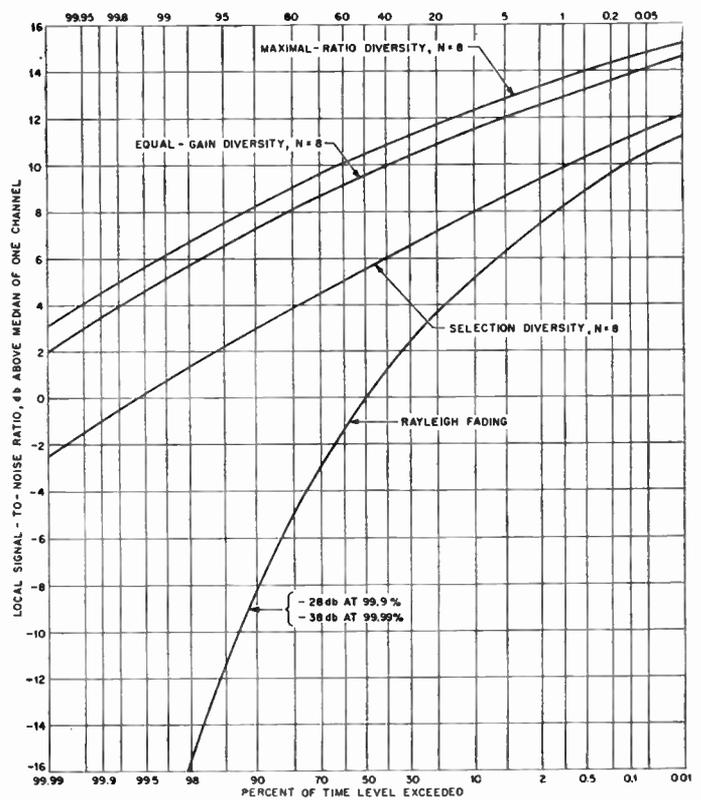


Fig. 13— $N=8$ .

$S_2(p)$  of (16) (selection).  $100 [1 - D_2(p)]$  is the per cent of time ordinate exceeded. The Rayleigh fading curve is  $100 [1 - G(p)]$  of (11). (For the distributions considered here, the median values of  $p$  do not differ from the corresponding mean values by more than about 1.6 db. The reason for using the median value of the Rayleigh distribution as a reference here is that this is commonly presented as an experimental datum, since median values can be read directly from the distribution function determined by a totalizer.)

It is clear that the differences between the various dual diversity curves of Fig. 9 are quite small, especially in comparison to the difference between any one of them and the Rayleigh fading curve. For example, the 99.99 per cent exceeded level of the selection diversity curve is almost 20 db above the Rayleigh curve, while the maximal-ratio curve is only 1.4 db above the selection curve at the 99.99 per cent point. Evidently one would choose among the three types of two-channel systems on the basis of Fig. 9 only if one were fighting for the last decibel. Even then, one would wish to make very sure that that last decibel could actually be realized, the selection diversity curve does not depend on the important assumptions (B) and (C), which must be satisfied for the equal-gain and maximal-ratio systems to work properly, as will be seen.

However, the differences in the performance of the various combining techniques become more important as the number  $N$  of channels is increased. The maximal-ratio and equal-gain systems improve much more rapidly than selection diversity does, as can be seen in Figs. 10-13, which give the distributions for  $N=3, 4, 6,$  and  $8$ , respectively.<sup>20</sup> (Note that the ordinate scales of Figs. 9-13 cover different ranges.) However, the maximal-ratio and equal-gain curves remain quite close together; indeed, the difference between them is hardly significant even for  $N=8$ . This is one of the facts that makes equal-gain diversity quite attractive and suggests that there are many applications where it should be exploited. It can be seen that the maximal-ratio and equal-gain curves differ approximately by the constants in the equal-gain column of Table I; that is, the equal-gain diversity distributions can be approximated quite well by translating the maximal-ratio distributions downward by the values in the second column of Table I.

The data of Figs. 9-13 are useful in the design of radio communication systems and radar and navigation systems of the type discussed in the Introduction. One such application is as follows. Suppose a high-reliability communication system is to be designed for a fixed information rate, which cannot be maintained whenever the received local power ratio  $p$  drops below a certain value. That is, it is desired to maintain the local power

ratio above a certain value for, say, 99 per cent of the time during an interval of length  $T_1$ , for which the curves of Figs. 9-13 are applicable if the relevant conditions are satisfied. Referring to the 99 per cent exceeded values of Fig. 9, it can be seen that the difference between the Rayleigh fading curve and the dual selection diversity curve is about 10 db at the 99 per cent point. But this implies that whatever transmitter power was required for a single-channel system, a transmitter of 10 db less power would be adequate if dual selection diversity were employed at the receiving terminal. Of course, part of the reduction could be applied to the antenna gains, etc. Similarly, reference to the 99 per cent values of Fig. 11 shows that the use of fourth-order maximal-ratio diversity would enable a reduction in transmitter power of 19 db relative to that required for a single-channel system.

This reduction in transmitter power required for a given grade of local reliability has been called "diversity gain." The term was apparently introduced by Jelonek, *et al.* [3]. Here the term "local reliability diversity gain," or simply "local reliability gain," is used to emphasize the fact that it is not a gain in the usual sense and that it depends very heavily on the local reliability percentage chosen. The dependence of the local reliability gains on the percentage selected can be seen in Table II, which gives the values realized by the three types of systems for  $N=2, 3, 4, 6,$  and  $8$  corresponding to local reliability percentages of 99 per cent and 99.9 per cent.

TABLE II  
LOCAL RELIABILITY GAINS (IN DB), CONDITIONS OF FIG. 8, FOR  
99 PER CENT AND 99.9 PER CENT LOCAL RELIABILITY

N	Selection		Equal-Gain		Maximal-Ratio	
	99 per cent	99.9 per cent	99 per cent	99.9 per cent	99 per cent	99.9 per cent
2	10	14.5	11	15.5	12	16
3	14	20	16	21.5	16.5	22.5
4	16	22.5	18.5	25	19.5	26
6	18	25.5	21.5	29	22.5	30
8	19	27	23.5	31.5	24.5	32.5

It can be seen that the values of Table II are much larger than those of Fig. 8. They can be made to appear even larger by computing the local reliability gains corresponding to 99.99 per cent or higher percentages; however, considering the present or immediately foreseeable state of the art, such values would not be meaningful. Among other things, the Rayleigh distribution does not provide an accurate model for actual fading distributions outside of the 0.1 per cent to 99.9 per cent range.

Various modifications and extensions of these considerations as they occur in practice will be considered next. However, it should be noted that there are many practical situations in which the conditions assumed above are realistic approximations, and for which the results can be used without significant modifications.

<sup>20</sup> High-resolution graphs of the curves of Figs. 9-13 are available from the author to those having serious need of such graphs. Letters requesting the same should describe the nature of said need. Requests on postal cards or form letters will not be honored. This offer may be withdrawn at any time.

VII. NON-RAYLEIGH FADING DISTRIBUTIONS

Only in the case of long-range UHF and SHF tropospheric transmission does it appear that observed fading distributions are most often Rayleigh, when observed in intervals of length  $T$ . For short-range UHF circuits and normal or scatter ionospheric transmission at VHF and below, other distributions are often observed. Indeed, at frequencies of a few megacycles and below, an accurate fit to the Rayleigh distribution is more nearly the exception than the rule. It is therefore of interest to discuss the way in which these results are modified by other distributions, assuming that the other conditions still hold.

Certain of these results are easily discussed. The maximal-ratio curve of Fig. 8 and the last column of Table I do not in any way depend on the fading distribution and are not modified at all. In order to discuss the effect on the distributions of Figs. 9-13, it will be convenient to return to the geometric approach mentioned in Appendix I and consider the case  $N=2$  channels. Let  $x=x_1$  and  $y=x_2$  be the local amplitude ratios of the two channels. The probability that a maximal-ratio system has a local power ratio  $\leq p$ , i.e., the maximal-ratio distribution function  $G_2(p)$ , is obtained by integrating the joint density function of  $x$  and  $y$  over the interior of the quarter circle  $x^2+y^2=p$  in the  $x-y$  plane. Similarly, the equal-gain distribution function  $A_2(p)$  is obtained by integrating the same density function over the triangle bounded by the line  $x+y=\sqrt{2p}$ . The corresponding region for the selection diversity distribution  $S_2(p)$  is the square bounded by  $x=\sqrt{p}$ ,  $y=\sqrt{p}$ . These three regions are shown together in Fig. 14. Now, the fact that the maximal-ratio system outperforms the other two is intimately connected with the fact that, for any fixed  $p$ , the probability that the maximal-ratio output ratio is  $\leq p$  is smaller than it is for the others, that is,  $G_2(p) < A_2(p)$  and  $G_2(p) < S_2(p)$ . This is reflected in Fig. 9 in the fact that the maximal-ratio curve is strictly above the others. The reason for this can be seen at once in Fig. 14; the region of integration for the maximal-ratio system is smaller than it is for the others and interior to both of the others. Hence, no matter what fading distribution is involved, the joint density function will still be non-negative and, therefore, its integral over the maximal-ratio region of Fig. 14, i.e.,  $G_2(p)$ , will still be smaller than for the others. Thus, the maximal-ratio curve of Fig. 9 would be above the others for any fading distribution.

Of course, this result could also be seen from the fact that the maximal-ratio system yields an output power ratio that is indeed maximal. But there is no similar fact to use as a guide in comparing the other two, for which we must rely on Fig. 14. It can be seen there that the areas of the selection and equal-gain regions are identical and that neither region includes the other. This would lead one to suspect that the relative performance of selection diversity and equal-gain diversity

depends on the form of the fading distribution. In order to discuss this, consider the nature of the possible departures from the Rayleigh distribution.

For purposes here, two cases may be distinguished: fading distributions more disperse (broader or more smeared-out) than the Rayleigh distribution, which are associated with frequent or persistent deep fades, and distributions less disperse (narrower or shallower) than the Rayleigh distribution, which are associated with shallow fading. These cases are illustrated in Fig. 15, together with the Rayleigh distribution. Curve (b), one of a family of distributions given by Rice, illustrates the less disperse or shallow fading often encountered at frequencies below UHF.<sup>21</sup> Curve (c) illustrates the more disperse case sometimes found in short-range UHF circuits and in high- and medium-frequency systems.

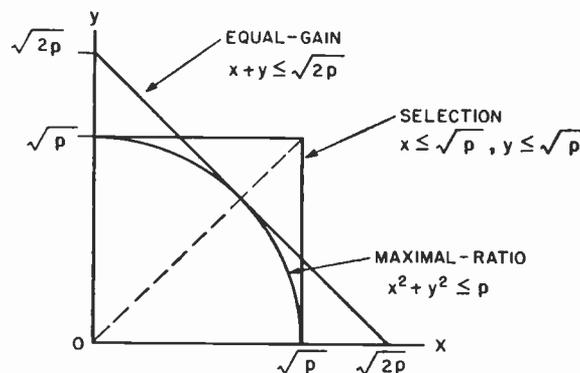


Fig. 14—Regions of integration for three types of dual diversity systems, after Altman and Sichak [8].

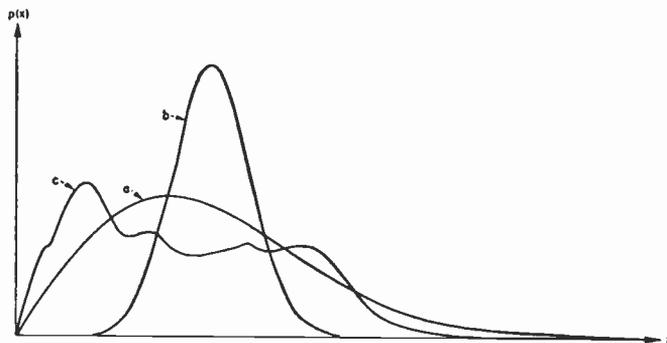


Fig. 15—Representative fading distributions. (a) Rayleigh density function. (b) Representative Rice distribution. (c) Typical distribution of the unpleasant sort often observed at frequencies below UHF.

Returning now to Fig. 14, it is not difficult to see that independent shallow fading will tend to improve the performance of an equal-gain system. This is because the height of the joint density function will be small in the region near the origin common to both the equal-gain and selection regions, and the bulk of the density function will be "pushed out" along the diagonal where

<sup>21</sup> R. W. E. McNicol, "The fading of radio waves of medium and high frequencies," *Proc. IEE*, vol. 96, pp. 517-524; October, 1949.

it will contribute more to the integral over the selection region than to the integral over the equal-gain region. Thus, an equal-gain combiner will continue to outperform selection diversity in the presence of shallow fading; indeed, its performance will more nearly approximate a maximal-ratio system. This can also be seen directly by considering the basic operation of a two-channel equal-gain system, and visualizing the case where the two signals are approximately constant.

However, this is not true for the more disperse distributions. Consider first the case where the individual amplitude ratios  $x$  and  $y$  are rectangularly distributed,<sup>2</sup> say on  $0 \leq x \leq A$  and  $0 \leq y \leq A$ , with a joint density function  $p(x, y) = 1/A^2$  on the square  $x \leq A, y \leq A$ . It is then easy to see that for values of  $p$  for which both the equal-gain and selection regions fit inside this square (*i.e.*, for  $\sqrt{2p} < A$ , or  $p < A^2/2$ ), their distribution functions are identical. That is, with respect to the smaller values of  $p$ , the equal-gain and selection systems yield identical performance. Next, suppose the independent amplitude ratios are exponentially distributed, say  $e^{-x}$  and  $e^{-y}$ , so that their joint density function is  $e^{-(x+y)}$ . Noting that the contours of constant height of this density function are the lines  $x+y=\text{constant}$ , parallel to the boundary of the equal-gain region, it is easy to see that the integral of this density function over the equal-gain region is strictly larger than it is over the selection region. This can also be verified by direct computation, as the relevant distribution functions are easily evaluated. Hence, for exponential amplitude fading, the local reliability gain of dual equal-gain diversity is, for any percentage, strictly less than it is for dual selection diversity.

It is thus seen that the relative performance of selection diversity and equal-gain diversity depends to some extent on the fading distribution involved. Consequently, the application of equal-gain diversity should be viewed with a modicum of caution in cases where very disperse fading distributions might be encountered. However, the exponential distribution used above is probably extreme in this respect,<sup>22</sup> and even for this case, the equal-gain system is not significantly poorer than selection diversity. For high reliability percentages, the local reliability gain of the dual maximal-ratio system over either the selection system or the equal-gain system is exactly (approximately)  $10 \log_{10} (4/\pi) = 1.05$  db for rectangular (exponential) fading.

It was noted above that the maximal-ratio data of Fig. 8 were independent of the fading distribution. However, the mean power ratios of equal-gain systems do depend on the distribution, but only to the extent of the parameter  $r^2 = (\bar{x})^2/\bar{x}^2$  of (44). For the rectangular and exponential distributions considered above,  $r^2 = \frac{3}{4}$  and  $r^2 = \frac{1}{2}$ , respectively, indicating that the average local

power ratio of an equal-gain system is not substantially degraded by even very disperse fading distributions. Unfortunately, no such simple and clear dependence of the selection diversity mean values on the form of the distribution exists. The result  $\bar{p} = \sum_{k=1}^N (1/k)$  of (18) is intimately wrapped up with the Rayleigh distribution, not merely the first two moments. But it is certainly clear that moderate changes in the form of the fading distribution could not lead to substantial changes in the selection diversity values of Fig. 8.

### VIII. RELATIVE EFFECTS OF CORRELATED FADING

Two smoothly varying random variables such as the  $x_j$  cannot, in general, be strictly independent. Of course, they may fail to be even approximately independent. It is therefore of interest to estimate the effect of dependent fading.

It is convenient to estimate this in terms of a parameter called the correlation coefficient. For two random variables  $x$  and  $y$  with positive variances<sup>2</sup>  $\sigma_x^2$  and  $\sigma_y^2$ , this is defined by

$$\rho_{xy} = \frac{\langle (x - \bar{x})(y - \bar{y}) \rangle}{\sigma_x \sigma_y}, \quad (45)$$

which reduces readily to

$$\rho_{xy} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}. \quad (46)$$

If  $x$  and  $y$  are independent, then  $\overline{xy} = \bar{x}\bar{y}$ .<sup>2</sup> Hence, if  $x$  and  $y$  are independent, then  $\rho_{xy} = 0$ . It is known<sup>23</sup> that  $-1 \leq \rho_{xy} \leq 1$ , and  $\rho_{xy} = \pm 1$  if and only if  $y = \pm ax + b$  ( $a > 0$ ).  $x$  and  $y$  are said to be correlated if  $\rho \neq 0$ , uncorrelated if  $\rho = 0$ , and partially correlated if  $0 < |\rho| < 1$ .

The problem of correlated fading in selection diversity systems has been studied by Staras [7] and others [3], [11], [13]. (See Appendix V for certain questions related to this subject.) Packard [14] and Bolgiano, *et al.*,<sup>8</sup> have studied this problem for two-channel maximal-ratio systems. Quite recently, Pierce<sup>24</sup> and Stein<sup>25</sup> independently studied correlated fading in  $N$ -channel maximal-ratio systems, and their results will be published in the near future. Some of Staras' results will simply be reproduced here in Fig. 16, in a form suitable for direct comparison with Fig. 9. The curve  $\rho = 1$  is the Rayleigh fading curve of Fig. 9, while  $\rho = 0$  denotes the dual selection diversity curve of that figure. It can be seen that approximately half of the uncorrelated local reliability gain is realized even for  $\rho = 0.8$ , and that the effect is negligible for  $0 < \rho < 0.3$ .

To consider the relative effect of correlated fading on

<sup>22</sup> Cramér, *op. cit.*, p. 265, or other standard sources. It is also known that the vanishing of  $\rho_{xy}$  does not necessarily imply that  $x$  and  $y$  are independent.

<sup>24</sup> J. N. Pierce, Air Force Cambridge Res. Center, Bedford, Mass., private communication.

<sup>25</sup> S. Stein, Hycon Eastern, Inc., Cambridge, Mass., private communication.

<sup>22</sup> So far as conventional applications are concerned. It should be noted that it is not extreme, or even sufficient, for postdetection distributions in FM systems, or special applications, such as that of Price and Green, *op. cit.*

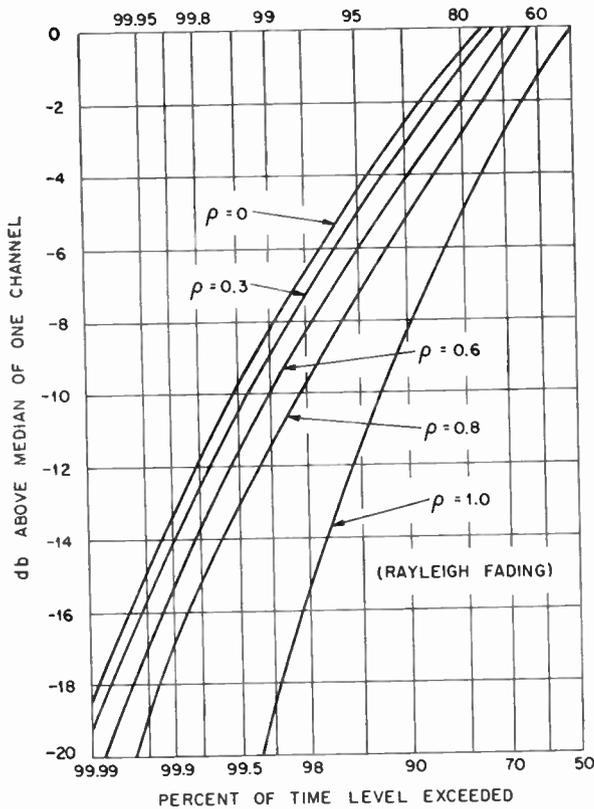


Fig. 16—Dual selection diversity distributions, Rayleigh fading, for various degrees of correlation.

the other systems, refer once again to Fig. 14. Now, in terms of the joint density function, there are two major effects of correlation: first, the mass of the density function tends to concentrate around the diagonal line  $y=x$ ; second, the mass tends to be pulled back nearer the origin. The first effect is simply an expression of the fact that as the correlation increases, the probability that  $y$  can differ appreciably from  $x$  necessarily decreases. (Variables with the same distribution are being considered here.) The second fact can be inferred from the behavior of Fig. 16. Given these facts, it is not difficult to see from Fig. 14 that appreciable correlation will, if anything, tend to improve the performance of equal-gain diversity, relative to selection diversity. (Of course, all three systems degrade in an absolute sense with increasing correlation.) Indeed, as  $\rho$  approaches 1, when the density function approaches zero except on the line  $y=x$ , it is clear that the equal-gain system approaches the maximal-ratio system in performance. This can also be seen by considering the basic operation of the two systems. From these considerations, it is not difficult to visualize the way in which the maximal-ratio and equal-gain curves of Fig. 9 follow the selection diversity curves of Fig. 16. At  $\rho=1$ , the dual maximal-ratio and equal-gain curves coincide and are uniformly 3 db above the selection curve for  $\rho=1$ .

With respect to the average values of Fig. 8, the maximal-ratio data are unaltered by correlated fading.

The equal-gain values actually increase toward the maximal-ratio values with increasing correlation. This can be seen either from physical considerations, or by noting that the terms  $\overline{x_i x_j}$  of (42) are replaced by  $x_i x_j$ ; for  $\rho > 0$ ,  $x_i x_j > \overline{x_i x_j}$ , and in fact  $x_i x_j$  approaches 1 as  $\rho_{ij}$  approaches 1. In contrast, the selection diversity values of Fig. 8 approach zero as the  $\rho_{ij}$  approach 1, as is easily seen.

In space-diversity communication systems, an antenna separation of 30 to 50 wavelengths is typically required to obtain correlation coefficients consistently less than 0.3. However, 10 to 15 wavelengths will often yield coefficients less than 0.6. Van Wambeek and Ross<sup>26</sup> measured the performance of certain III' selection diversity systems directly, without measuring correlation coefficients, and apparently found that even shorter spacings led to useful results. More recently, Grisdale, *et al.*,<sup>5</sup> have obtained numerous data bearing on this question in the 6- to 18-mc region.

#### IX. VARIABLE LOCAL NOISE POWERS

Many of the data above were obtained on the assumption  $\overline{n_j^2} = \text{constant}$ . This will not be usually strictly true and in many cases will not even be approximately true. If the noises are principally due to interference from remote sources, the  $\sqrt{\langle n_j^2 \rangle}$  themselves may well follow the Rayleigh distribution, a case that has recently been studied by Bond and Meyer [12] for dual selection diversity. Related material has also been given by Clarke and Cohn.<sup>27</sup> If the  $n_j$  are principally due to receiver front end noise, then the  $\overline{n_j^2}$  may be approximately constant; the actual amount of fluctuation to be expected is a function of the noise bandwidth and the duration  $T$  of the local averages. This fluctuation has been studied by Rice,<sup>28</sup> whose results are quite useful in determining a suitable value of  $T$ .

In terms of the analysis above, the principal effect of variable  $\overline{n_j^2}$  is to modify the distribution of the  $p_j = x_j^2 / \overline{n_j^2}$  with results as discussed in Section VII above. (The distribution of the  $p_j$  becomes more disperse as the noise power fluctuations increase.) It is not difficult to obtain quantitative estimates of the degradation in particular cases. It should be pointed out that extreme fluctuations in noise power can lead to very poor performance of an equal-gain system, which has no provision for cutting off a very noisy channel, in contrast to the maximal-ratio and selection systems.

<sup>26</sup> S. H. Van Wambeek and A. H. Ross, "Performance of diversity receiving systems," *Proc. IRE*, vol. 39, pp. 256-264; March, 1951.

<sup>27</sup> K. K. Clark and J. Cohn, "Carrier-to-noise statistics for various carrier and interference characteristics," *Proc. IRE*, vol. 46, pp. 889-895; May, 1958.

<sup>28</sup> S. O. Rice, "Filtered thermal noise—fluctuation of energy as a function of interval length," *J. Acoust. Soc. Amer.*, vol. 14, pp. 216-227; April, 1943. For additional results, see also, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 24, pp. 46-156; January, 1945. (Section 3.9, p. 87 ff.) These results have recently been extended by D. Slepian, "Fluctuations of random noise power," *Bell Sys. Tech. J.*, vol. 37, pp. 163-184; January, 1958.

## X. FAILURE OF THE NOISES TO BE LOCALLY INCOHERENT

The failure of assumption (C) would have no effect on selection diversity systems, for which the assumption  $\overline{n_i n_j} = 0$  if  $i \neq j$  has no relevance whatever. However, this assumption is of vital importance for the maximal-ratio and equal-gain systems, as will be seen.

There are essentially two ways in which  $\overline{n_i n_j}$  may fail to be identically zero, the first of which is simply due to the fact that the average  $\overline{n_i n_j}$  is over a short interval of duration  $T$  and the local average  $\overline{n_i n_j}$  will fluctuate about zero if the noises are basically unrelated. The amount of fluctuation will decrease as  $T$  is increased and will be small if the lowest frequency of the noise is large in comparison to  $1/T$ . In this case, the  $\overline{n_i n_j}$  terms will be negative as often as positive and will simply contribute a small perturbation to the output noise power  $\overline{n}^2$  of a maximal-ratio or equal-gain system. This case is not troublesome.

The troublesome case arises when the noises have a definite positive correlation, as can happen, for example, in a postdetection combiner when the noises stem largely from sources of external interference. To consider this, let  $\overline{n_i^2} = 1$  and  $\overline{n_j} = 0$ ; then  $\rho_{ij} = \overline{n_i n_j}$  is the correlation of  $n_i$  and  $n_j$ . (Note that the local correlation over intervals of length  $T$ , in contrast to the correlation over length  $T_1$  of the  $x_j$  discussed in Section VIII above, is considered here.) Let  $\rho_{ij} = \rho$  if  $i \neq j$ . Then the output local noise power of an equal-gain system becomes

$$\begin{aligned} \overline{n}^2 &= \left\langle \left[ \sum_{j=1}^N n_j \right]^2 \right\rangle \\ &= \sum_{j=1}^N \overline{n_j^2} + \sum_{i \neq j} \overline{n_i n_j} \\ &= N[1 + (N-1)\rho], \end{aligned} \quad (47)$$

instead of  $\overline{n}^2 = N$ . Hence, the local noise power is increased by the factor  $[1 + (N-1)\rho]$ , which is to say the output power ratio  $\bar{p}$  of (42) is decreased by this factor, which may not be trivial. To see how untrivial it can be, consider  $\bar{p}$  for an equal-gain system. Eqs. (45) and (46) become

$$\begin{aligned} \bar{p} &= \frac{\sum x_j^2 + \sum_{i \neq j} \overline{x_i x_j}}{N[1 + (N-1)\rho]} \\ &= \frac{N + N(N-1)r^2}{N[1 + (N-1)\rho]} \\ &= \frac{1 + (N-1)r^2}{1 + (N-1)\rho}, \end{aligned} \quad (48)$$

which reduces to (46) when  $\rho = 0$ . Hence, if  $\rho > r^2$ —a situation by no means impossible—it would follow that  $\bar{p}(1) > \bar{p}(2)$ ; *i.e.*, the average local power ratio of a two-channel equal-gain system would be less than for a single channel, and the performance gets worse as the number of channels is increased. It is probably gratuitous to point out explicitly that, in such a case, it would

be much better to use a selection diversity system, for which (18) would still hold. Similar considerations show that the average local noise power of a maximal-ratio system—by which is meant one for which the coefficients are given by (14), though this is no longer “maximal”—is increased by the factor  $[1 + (N-1)\rho r^2]$ .

It follows that the use of maximal-ratio or equal-gain diversity in circumstances where the noise voltages may be highly correlated is hazardous.

## XI. PREDETECTION VS POSTDETECTION COMBINING

In systems where the power ratio at the output of the detector is essentially the same as at the input, there is no fundamental change required in the conclusions developed above. Of course, there are always practical differences between predetection and postdetection combining; *e.g.*, a predetection maximal-ratio or equal-gain combiner will require the addition of phase-control circuitry in order to satisfy the local-coherence assumption (B). On the other hand, predetection selection diversity will sometimes produce smaller switching transients than postdetection selection. Once phase control is established, it is easier to satisfy the conditions (B) and (C) required for maximal-ratio and equal-gain combiners in the case of a predetection system.

However, substantial changes are required in the case of FM systems with a large deviation ratio, or in other bandwidth-exchange systems with a pronounced threshold effect. In such systems, an SNR at the detector input that is more than a few db above threshold yields a large output ratio, while an input ratio that is more than a few db below threshold yields a very small output ratio. That is, the output ratio changes from “completely useful” to “completely useless” with a few db change of input ratio. This fact has important consequences.

To begin with, a Rayleigh distribution of input signal strength for an FM system will emphatically not lead to a Rayleigh distribution of the postdetection amplitude ratio. Hence, the distribution-sensitive results of Figs. 8–13 and Tables I and II are not realistic for postdetection combining in FM systems. Furthermore, equal-gain combiners are not even suitable for postdetection combining in conventional FM systems; this may be regarded as a consequence of the fact that the detection gain of such systems is not constant. An alternative point of view would be that the distribution of the amplitude ratios at the input of the combiner would be such as to eliminate the equal-gain combiner from consideration; cf. (36), and note the unfortunate effect if any one of the  $\overline{n_j^2}$  becomes large.

Of course, a maximal-ratio system can be used for postdetection combining in an FM system. The requirement  $a_j = x_j / \overline{n_j^2}$  for the coefficients insures that any channel with large  $\overline{n_j^2}$  contributes very little to the output. However, a maximal-ratio system will not yield much improvement over selection diversity in such circumstances. It will eliminate switching transients, but

otherwise will not usually make a significant difference in the operation of the system. This can be seen on the basis of various qualitative considerations. When dealing with postdetection combination in sharp-threshold FM systems, at least for  $N \leq 8$ , it would probably be best to use only the selection diversity values of Table II, whether selection or maximal-ratio diversity is actually used. In any event, the actual local reliability gains of such maximal-ratio systems—which could be computed from specific detector characteristics, such as those given by Middleton<sup>29</sup> or obtained by measurement—would certainly be less than the maximal-ratio values in Table II. A specific distribution computed on the basis of a highly simplified detector characteristic has been given. [16]

If the local reliability gain is defined in terms of the transmitter power required to maintain the input level of the detector above a certain value for more than a specified percentage of time, then the selection diversity values of Table II are applicable whether the selection is predetection or postdetection. It is clear that the operation is identical in either case. Furthermore, the maximal-ratio and equal-gain data of Table II are completely applicable to predetection combining, as is easily seen. Accordingly, the full advantages of maximal-ratio and equal-gain combiners can be realized in FM systems when and only when they are employed before detection. Taking the selection values of Table II as being the gains obtained by a postdetection maximal-ratio combiner, the differences between the maximal-ratio and selection values of Table II then illustrate the added advantage of predetection maximal-ratio or equal-gain combining. This may be regarded as due to an effective lowering of the detector threshold resulting from these techniques.

An additional advantage of predetection combining in FM systems is that FM multipath distortion can be reduced by this method. It has been shown by Adams and Mindes [16], both theoretically and experimentally, that a predetection equal-gain combiner yields substantially less multipath distortion than is obtained with a postdetection maximal-ratio combiner, when both are operated under the same circumstances.

Instrumentation for postdetection maximal-ratio combining has been discussed by Kahn [6] and by Morrow, *et al.*,<sup>11</sup> for what amount to AM systems, and by Mack<sup>30</sup> for FM systems. An ingenious predetection maximal-ratio combiner has been devised by Granlund.<sup>31</sup> A particularly elegant predetection equal-gain combiner has been developed by the Federal Telecommunication Laboratories (now the ITT Laboratories),

<sup>29</sup> D. Middleton, "On theoretical signal-to-noise ratios in FM receivers: a comparison with amplitude modulation," *J. Appl. Phys.*, vol. 20, pp. 334-351; April, 1949.

<sup>30</sup> C. L. Mack, "Diversity reception in UHF long-range communications," *Proc. IRE*, vol. 43, pp. 1281-1289; October, 1955.

<sup>31</sup> J. Granlund, "Topics in the Design of Antennas for Scatter," Lincoln Lab., M.I.T., Lexington, Mass., Tech. Rep. No. 135, pp. 105-113; November, 1956. See also recent Quarterly Progress Reports of the M.I.T. Res. Lab. of Electronics.

indicated in Fig. 17. This combiner, called simply a phase combiner in FTL literature, is the same one used in the experiments reported by Adams and Mindes [16]. The phase control and adder circuits require only two semiconductor diodes and 16 passive linear elements. Phase control is established via a phase discriminator, the output of which is applied as a bias voltage to one of the local oscillators. This corrects the phase of the local oscillator via Miller-effect changes in the oscillator tube capacity.

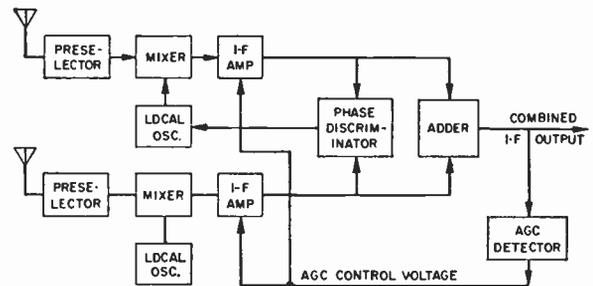


Fig. 17—FTL predetection equal-gain combiner. This can be used with any type of modulation.

The problem of adequate phase control for predetection maximal-ratio or equal-gain combining leads naturally to the next topic, namely:

### XII. FAILURE OF THE LOCAL-COHERENCE CONDITION (B)

It is obviously of interest to estimate the possible degradation in performance of maximal-ratio and equal-gain combiners when the local-coherence condition (B) is not satisfied. The following treatment is due to Stein.<sup>32</sup>

Recall that (B) was the assumption  $s_j(t) = x_j m(t)$  where  $x_j$  was the slowly varying local rms value of  $s_j$ . If the  $s_j$  are not all in phase, we must write  $s_j(t) = x_j m_j(t)$ , where the  $m_j$  have different phases. Consider the case  $m_j(t) = \sqrt{2} \cos(\omega t - \phi_j)$  where the  $\phi_j$  are locally constant in the sense that the  $x_j$  are. Then  $\langle m_j^2 \rangle = 1$  and  $\langle s_j^2 \rangle = x_j^2$ , as before, averaging over a few cycles (or more) of  $\omega t$ . Then, for any locally linear combiner of the type (6),

$$\begin{aligned} \overline{s^2} &= \left\langle \left[ \sum_{j=1}^N a_j s_j \right]^2 \right\rangle \\ &= \left\langle \sum a_j^2 x_j^2 2 \cos^2(\omega t - \phi_j) \right. \\ &\quad \left. + \sum_{i \neq j} a_i a_j x_i x_j 2 \cos(\omega t - \phi_i) \cos(\omega t - \phi_j) \right\rangle \\ &= \sum a_j^2 x_j^2 + \sum_{i \neq j} a_i a_j x_i x_j (2 \cos(\omega t - \phi_i) \cos(\omega t - \phi_j)) \\ &= \sum a_j^2 x_j^2 + \sum_{i \neq j} a_i a_j x_i x_j \cos(\phi_i - \phi_j), \end{aligned} \tag{49}$$

where the last step used  $2 \cos A \cos B = \cos(A+B)$

<sup>32</sup> S. Stein, private communication; August, 1957.

$+\cos(A-B)$  and the fact that  $\langle \cos(2\omega t - \phi_i - \phi_j) \rangle = 0$ .  
Eq. (49) may also be written

$$\overline{s^2} = \sum_{i,j} a_i a_j x_i x_j \cos(\phi_i - \phi_j) \tag{50}$$

since  $\cos(\phi_j - \phi_j) = 1$ , and this reduces to

$$\overline{s^2} = \left[ \sum_{i,j} a_j x_j \right]^2 = \sum_{i,j} a_i a_j x_i x_j \tag{23'}$$

when  $(\phi_i - \phi_j) = 0$ . Let  $p$  denote the output power ratio of the general combiner (6) when (23') holds, and  $p'$  denotes the same for the phase-degraded case (50). Then (assuming (C) still holds, so that  $\overline{n^2} = \sum a_j^2 \overline{n_j^2}$ )

$$\begin{aligned} p' &= \frac{\sum_{i,j} a_i a_j x_i x_j \cos(\phi_i - \phi_j)}{\sum a_j^2 \overline{n_j^2}} \\ &= \frac{\sum_{i,j} a_i a_j x_i x_j \cos(\phi_i - \phi_j)}{\left[ \sum a_j x_j \right]^2} \frac{\left[ \sum a_j x_j \right]^2}{\sum a_j^2 \overline{n_j^2}} \\ &= kp, \end{aligned} \tag{51}$$

where

$$k = \frac{\sum_{i,j} a_i a_j x_i x_j \cos(\phi_i - \phi_j)}{\left[ \sum a_j x_j \right]^2} \tag{52}$$

is the "phase degradation ratio"  $p'/p$ .

Apart from the fact that  $0 \leq k \leq 1$ , not much can be said about  $k$  in the general case (52) in the absence of additional information about the  $\phi_j$ . It is easy to see that  $k$  may actually vanish; e.g.,  $N=2$ ,  $a_1=a_2=x_1=x_2=1$ ,  $\phi_1-\phi_2=180^\circ$ . Then  $p'=k=0$ , which is entirely to be expected when adding two signals of equal magnitudes and opposite phases. This illustrates the fact that the condition (B) cannot be ignored. On the other hand, it is not necessary that it should be satisfied with great precision. Suppose that the magnitudes of the phase differences,  $|\phi_i - \phi_j|$ , do not exceed  $90^\circ$ , and let  $\Delta = \text{maximum of } |\phi_i - \phi_j|$ ,  $i, j=1, 2, \dots, N$ ;  $0 \leq \Delta \leq 90^\circ$ . Then  $0 \leq \cos \Delta \leq \cos(\phi_i - \phi_j)$ , so

$$k \geq \frac{\sum_{i,j} a_i a_j x_i x_j \cos \Delta}{\left[ \sum a_j x_j \right]^2} = \cos \Delta, \tag{53}$$

or

$$p' \geq p \cos \Delta. \tag{54}$$

That is, the local power ratio is not reduced by more than  $\cos \Delta$ , or  $-10 \log_{10} \cos \Delta$  db, in any combiner whatever of the general type (6), provided  $\Delta \leq 90^\circ$ . In particular, this conclusion holds for equal-gain and maximal-ratio combiners. Thus, to restrict the reduction in  $p$  due to imperfect phase control to 1 db or less, it is only necessary to maintain the phases within  $37.5^\circ$  of each other, while  $51^\circ$  is sufficient to guarantee a maximum loss of 2 db. Furthermore, it is clear that the estimate  $p \cos \Delta$  is actually conservative.

### XIII. LONG-TERM VARIABILITY

Recall that the distributions of the  $x_j$  and  $p$ , and mean values of these quantities, were to be determined in intervals of length  $T_1$ , relative to which the  $x_j$  were assumed to be approximately Rayleigh-distributed and approximately independent. It is important to understand the nature of this situation.

It is an experimental fact that, for a suitable choice of  $T_1$ , both of these assumptions are often satisfied. It is also an experimental fact that if  $T_1$  is made too long or too short, neither assumption is satisfied. Hence, the approach used above and all of the results developed above depend entirely on the use of finite intervals of observation that are neither too long nor too short.

Specific suitable values of  $T_1$  depend on the circumstances, primarily the carrier frequency and transmission distance. Of course, it is necessary to understand that the results of Figs. 8-13, etc., refer only to intervals of length  $T_1$ , whatever this may be. Specific values are roughly as follows for long-range transmission. At frequencies below VHF, intervals of 30 minutes to an hour are usually suitable. In VHF ionospheric scatter systems, values of one or two minutes usually seem to be appropriate; in UHF and SHF tropospheric systems, intervals of five to 30 minutes are often used.

It is manifestly necessary to consider the behavior of diversity systems over longer intervals than those of length  $T_1$ . This may be done as follows. The previous results were obtained using a Rayleigh distribution (9) of unit mean square, with a definite median value  $M_0 = \sqrt{\log_e 2}$ . Now, the experimental fact is that the fading distributions observed over different intervals of length  $T_1$  will not usually have the same median values. However, the medians obtained in two adjacent or overlapping intervals of length  $T_1$  will not usually differ by very much. One way to represent this fact is to let  $M = M(t)$  denote the median of the distribution obtained in the interval from  $t - T_1$  to the present time  $t$ . Then this median function is a continuous function of time and the experimental fact is that  $M(t)$  is usually approximately constant over intervals of length  $T_1$ . It should be clear that this does not depend on having the distributions, for which the values of  $M(t)$  are the medians, all be of the same form, Rayleigh or other.

This may be used as follows. If  $x$  is any random variable with a nonzero median  $M_0$  and  $M$  is any nonzero constant, then it is easy to see that  $y = (M/M_0)x$  has a distribution of the same form as that of  $x$ , differing only in the scale factor  $M/M_0$ , and that the median value of  $y$  is  $M$ . (The only reason for writing this scale factor in terms of the medians is that these are easily determined experimental quantities.) Then, instead of taking the local rms signals to be the  $x_j$  with a fixed median  $M_0$ , the actual local rms signals may be written as  $y_j = (M_j/M_0)x_j$ , with median values  $M_j$ , relative to a period of length  $T_1$ . Here, however, another experimental fact enters. The medians  $M_j$  are usually approximately the same for different channels, and we

may write  $M_j = M$ ,  $j = 1, 2, \dots, N$ . If the median function  $M(t)$ , above, is approximately constant over intervals of length  $T_1$ , we may take  $M = M(t)$ .

To apply this to our previous results, note that the local linearity of (6) implies that any common scale factor multiplying the signal components  $s_j$  may be taken outside the sum  $s = \sum s_j$ . Hence, the combined output signal  $s(t)$  is simply multiplied by  $M/M_0$  and the local power ratio becomes  $(M/M_0)^2 p$  wherever  $p$  was before. This becomes even simpler when expressed in decibels. Let

$$\begin{aligned} w &= 10 \log_{10} [(M/M_0)^2 p] \\ &= 20 \log_{10} M + 10 \log_{10} (p/M_0^2) \\ &= u + v \end{aligned} \quad (55)$$

be the local power ratio in db, where  $u = 20 \log_{10} M$  and  $v = 10 \log_{10} (p/M_0^2)$ . Then this expresses the actual local power ratio delivered by any combiner of the type (6) as the sum of a variable  $v$  whose  $T_1$  median does not depend on time and a variable  $u$  that is approximately constant over every interval of length  $T_1$ . Now, the distributions plotted in Figs. 9–13 are precisely the distributions of the variable  $v$  for the conditions of Fig. 8, for different combiners and orders of diversity. Hence, to apply the results of Figs. 9–13 to any particular interval of length  $T_1$ , it is only necessary to translate their ordinate scales by  $u = 20 \log_{10} M$  where  $M$  is the median of the single-channel fading distribution for the interval concerned.

In order to describe the long-term variability of the actual local power ratio  $w$ , it is necessary to have information about the long-term variability of the ( $T_1$ -) median  $u$ . Distributions of  $u$  are usually studied in intervals of length  $T_2 =$  one month to one year; such distributions are often called distributions of hourly medians, though they should properly be distributions of  $T_1$ -medians. Several such distributions for frequencies at VHF and above have been given.<sup>33</sup> Unfortunately, no comparable single source of information for MF and HF systems presently exists; the relevant data are largely scattered in (generally unobtainable) Signal Corps reports, FCC hearing transcripts, and URSI and CCIR documents, though a few such data have been published. Observed distributions of  $u$  are sometimes approximately Gaussian (normal) in form, especially at the higher frequencies, which is why the distributions of  $M$  are often said to be log-normal.

Once a distribution of  $u$  pertinent to the proposed circuit is available, there are two ways it may be used. The first method, which is applicable to high-reliability systems at VHF and above, is to estimate the lowest  $T_1$ -median likely to be encountered on the circuit. (The great virtue of these systems is that this minimum value of  $u$  is not  $-\infty$ .) The ordinate scales of Figs. 9–13 are then translated to this value, after which a rational

choice among the various possibilities of transmitter power, order and type of diversity system, etc., may be made on the basis of economic and other factors, and on the basis of the local reliability percentage it is desired to maintain during such worst hours. Of course, the data of Figs. 9–13 must be modified in accordance with the discussion of Sections VII to XII if the circumstances so dictate.

The second method is applicable in circumstances where all distributions of  $v$  in all intervals of length  $T_1$  are approximately the same.<sup>34</sup> It is then easy to see that the distribution of  $v$  in an interval of length  $T_2 =$  one year would also be the same; furthermore, the variables  $u$  and  $v$  of (55) would then be independent, relative to  $T_2$ , to a very high degree of accuracy. Then the  $T_2$ -distribution of the local power ratio  $w$  would be the distribution of the sum of two independent variables with known distributions, and could be computed. It would usually be found that the "exact" determination of the  $T_2$ -distribution of  $w$  would require numerical methods of integration in (67). Such distributions have been computed by Shimony<sup>35</sup> and Sichak<sup>36</sup> among others. However, Staras [9] has observed that, for the larger values of  $N$ , the relevant distributions of  $v$  are approximately normal (cf. the  $N=6$  and  $N=8$  curves of Figs. 12 and 13, on which figures a normal distribution would be a straight line) and since the distribution of  $u$  is approximately normal, the  $T_2$ -distribution of  $w$  would therefore<sup>2</sup> approximate a normal distribution with a median and variance respectively equal to the sums of the medians and variances of the  $u$  and  $v$  distributions. But this approximation is not very accurate for  $N \leq 4$ .

It should be added, however, that this second method has not been universally accepted by designers of high-reliability systems, for the following reasons. Computing the long-term distribution of  $w$  serves to obscure the question of whether the periods of very low signal are a few long periods or many short ones. This question can be important; e.g., there are systems in operation in which the loss of two hours in a year would not be troublesome if split into a number of separated intervals of a minute or two each, but which could be disastrous if concentrated in a single interval of two hours. Since two hours in a year corresponds to the 99.98 per cent exceeded level, it would be necessary to compute the  $T_2$ -distribution of  $w$  down to something like the 99.999 per cent exceeded level in order to insure approximately the reliability obtained by the first method. However, the empirical distributions on which this computation must ultimately rest are not known to anything approaching this degree of accuracy, and the validity of such a computation would seem to be open to question. In addition, the problem mentioned in footnote 34

<sup>34</sup> It should be noted that such circumstances, however, are not too common in practice.

<sup>35</sup> A. Shimony, Final Rep. MU-156, Evans Signal Lab., Fort Monmouth, N. J., (cited in footnote 36).

<sup>36</sup> W. Sichak, Fed. Telecommun. Labs., Nutley, N. J., Tech. Memo. 619; December, 1956. See Appendix D, "Diversity Theory."

<sup>33</sup> Scatter Propagation Issue, PROC. IRE, vol. 43; October, 1955.

would often infect such a computation.

Two additional points should also be noted in connection with long-term distributions of  $w$ . The first is that the variability or dispersion of  $w$  will increase as the dispersion of the  $T_1$ -median  $u$  increases. In other words, the variability of  $v$  will tend to be obscured by that of  $u$ . However, it is precisely the variability of  $v$  that can be reduced by diversity techniques, while that of  $u$  cannot. It has been noted<sup>36</sup> that  $w$  distributions for different orders of diversity show less difference than would be indicated by Table II, but this is simply a reflection of the dispersion contribution by  $u$ . Thus, computing long-term distributions of  $w$  tends to obscure the gains (Table II) that actually are realized by diversity techniques.

The second point to be noted is that any long-term distribution of  $w$  is highly specific to the circuit for which it was computed, because  $u$  distributions are highly specific. This is indicated in Fig. 18,<sup>37</sup> which shows percentage points of  $u$  distributions as a function of distance at 400 mc. (Note that Fig. 18 gives only the distributions of the  $T_1$  median "scatter" loss; the free-space loss has been subtracted out.) It can be seen that the dispersion of  $u$  decreases with distance; e.g., the interdecile range is about 12.5 db at 200 miles, but only 5 db at 600 miles. This indicates that a long-term distribution of  $w$  would only be valid for the distance on which the distribution of  $u$  was based.

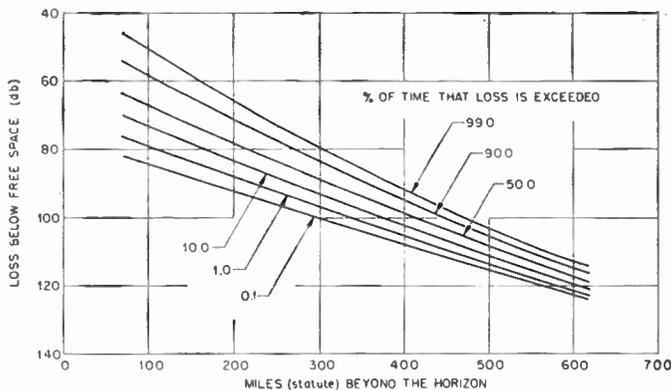


Fig. 18—Distributions of hourly medians as a function of distance, wintertime propagation at 400 mc, experimental data from several Lincoln Laboratory circuits. (After Morrow<sup>37</sup>.)

#### XIV. CASE OF UNEQUAL MEDIAN SIGNALS

Most of the material above presupposed that the  $T_1$ -medians for the several channels were all the same. Experimentally, this is found to be a reasonable approximation in most cases, provided that the interval length  $T_1$  is not made too short. However, there are cases, most especially angle diversity,<sup>8-10</sup> in which it is not a reasonable approximation. The first treatment of this problem apparently was given in [3]. It is quite simple to plot

selection diversity distributions for unequal medians or even for dissimilar distributions; the identical factors  $G(p)$  of (16) are simply replaced by the proper distribution functions. Maximal-ratio distributions for unequal-median Rayleigh signals can also be expressed explicitly. Unfortunately, this is not true of complete equal-gain distributions (cf. Appendix IV). However, it would be possible to obtain the low-signal ends of such distributions by taking the first few terms of a power series, but a detailed analysis of this problem would appear to be premature at the present time.

#### XV. EXPERIMENTAL RESULTS

Experimental data relating to diversity systems have been given by several workers, including Glaser and Van Wambeek,<sup>38</sup> Van Wambeek and Ross,<sup>26</sup> Glaser and Faber,<sup>4</sup> and Grisdale, *et al.*<sup>5</sup> It is unfortunate for our present purposes that most of these data relate to selection diversity only; clear-cut and unambiguous experimental data bearing on the comparative performance of the three combining techniques studied above are so rare to be as essentially nonexistent. It is hoped that some comparative experimental results will be available within the next year.

Perhaps the best single datum presently available was obtained in unpublished experiments conducted by the Signal Corps a few years ago. Two high-frequency systems were compared, one of which used dual maximal-ratio diversity and the other used something approximating selection diversity. It appeared that the maximal-ratio system yielded an average power ratio of 1.0 to 1.5 db above the selection system when averaged over periods of about 30 minutes.<sup>39</sup> This compares very favorably with the value 1.25 db entered in the "selection" column of Table I for  $N=2$ . However, there were many periods during which the performance of the maximal-ratio system was inferior to the other. This could probably be traced to the failure of one or both of the conditions (B) and (C) during such periods, which would not affect a selection system.

#### XVI. CONCLUSIONS

Perhaps the most important conclusion to be drawn is that, all things considered, no one of the diversity combining techniques studied deserves to be called "the optimum system." All three have their merits and defects, and the one to be used will depend on the circumstances. However, the simplicity and efficacy of the equal-gain system suggest that this may well become the principal standard of the art. In addition to the data set forth above, it should be especially noted that the instrumentation for an equal-gain combiner is completely independent of what one chooses to think of as a SNR. The importance of this fact is considerable.

<sup>38</sup> J. L. Glaser and S. H. Van Wambeek, "Experimental evaluation of diversity receiving systems," *PROC. IRE*, vol. 39, pp. 252-255; March, 1951.

<sup>39</sup> F. E. Bond and H. F. Meyer, Signal Corps Eng. Labs., Fort Monmouth, N. J., private communication; June, 1957.

<sup>37</sup> W. E. Morrow, Jr., "Etude de systemes de radiocommunication troposphérique UHF a longue distance," *Onde Elect.*, vol. 37, pp. 444-449; May, 1957.

APPENDIX I

CERTAIN FACTS ABOUT PROBABILITY THEORY

It will be recalled<sup>2</sup> that, in general, a distribution function  $P(x)$  is the probability that (some random variable) is less than or equal to  $x$ . A particularly simple special case of this arises when the random variable in question is some voltage or current waveform given as a function of time, say  $f(t)$ , and the probability that  $f \leq x$  is simply the fraction of some interval  $t_1 - T \leq t \leq t_1$  in which  $f \leq x$ . In this case,  $P(x)$  is determined, for any given value of  $x$ , simply by adding up the lengths of the  $t$  intervals for which  $f(t) \leq x$  and dividing their sum by the total duration of the observation, as indicated in Fig. 19. Instruments for measuring  $P(x)$  at selected values of  $x$  are known as "totalizers" or "level distribution recorders" and exist in various forms. Another method of obtaining the distribution function of some random variable is to sample it at discrete intervals and count the fraction of sampled values that are  $\leq x$ ; however, it is not difficult to see that this will lead to the same  $P(x)$  as defined above. The associated density function<sup>2</sup>  $p(x) = dP(x)/dx$ , so that  $p(x)dx = dP(x)$ .

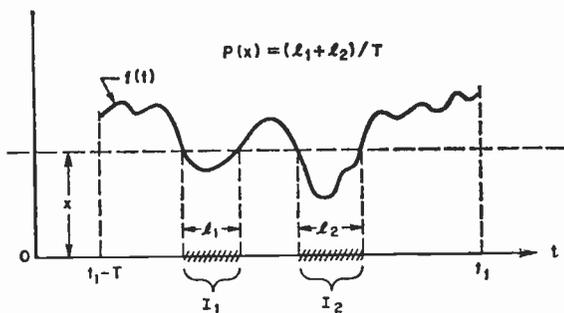


Fig. 19—Definition of  $P(x)$  = fraction of the time of observation that  $f(t) \leq x$ .

The central purpose in using such distribution functions in radio engineering stems from the fact that many different individual waveforms have approximately the same distribution function, or at least have distribution functions that differ in describable ways, as in Section XIII. This is an experimental fact, no more, but, what is important, no less. Thus, in circumstances where the distribution function of some waveform can be approximately predicted from either theoretical or empirical grounds, one has available a method of predicting many important facts about the situation.

One such fact, of the first importance, is that all time averages of  $f$  in the interval  $t_1 - T \leq t \leq t_1$  are given by the moments<sup>2</sup> of the distribution function  $P(x)$ . For example, suppose one is interested in the average value of  $f(t)$ . Then

$$\frac{1}{T} \int_{t_1-T}^{t_1} f(t) dt = \int_{-\infty}^{\infty} x p(x) dx = \int_{-\infty}^{\infty} x dP(x); \quad (56)$$

i.e.,  $\bar{f}$  is the first moment of the distribution. More generally, for any value of  $n$ , the time average of

$[f(t)]^n$  is given by the  $n$ th moment of the distribution:

$$\frac{1}{T} \int_{t_1-T}^{t_1} [f(t)]^n dt = \int_{-\infty}^{\infty} x^n dP(x), \quad (57)$$

a result that is especially useful in computing the average power when  $f$  is a voltage or current and  $n=2$ . In the light of (57),  $\bar{f}^n$  or  $\langle f^n \rangle$  or  $\overline{x^n}$  or  $\langle x^n \rangle$  can be and is written interchangeably for such averages, using whichever notation seems most convenient for the expression involved.

It is important to understand the sense in which (57) is applicable. No theory or representation or mathematical model whatever can predict the particular distribution function of a particular waveform in a particular interval exactly, but to the extent it can be predicted by whatever means, (57) is applicable. In many applications, one uses a specific mathematical model distribution (e.g., the Rayleigh distribution, much used in the body of this paper) for predicting facts such as averages of the form (57), but with a clear understanding that any realized distribution function can only approximate the Rayleigh distribution, however long or short the interval of observation. But the approximation may be very close. In cases where the nature of the possible departures from the model distribution can be estimated, and there are many such cases, the possible departures in the corresponding time averages can similarly be estimated via (57). [Some statisticians and noise theorists may be bothered by the absence from this discussion of any reference to the classical notions of sample and population. The reason for this is that a parent population in the classical sense does not usually exist in this environment. (See Section XIII.) No fixed distribution can serve as a population distribution for any non-stationary process. In the notation of Section XIII, it would be possible, but not necessarily desirable, to discuss parent populations for  $T_2$ -distributions, but certainly not  $T_1$ -distributions. One may, however, discuss a "distribution" of distributions, as is done in engineering language in that section. It would be a simple matter to provide a more formal framework for this material by defining suitable classes of functions  $f$ ; e.g., all those  $f$  whose half-hour local distribution functions (in the sense of Fig. 19) were all within a specified distance (in the sense of Lévy's metric) of some Rayleigh distribution function, and whose half-hour medians (*ipso facto* unique) had yearly distribution functions within a specified distance of a fixed log-normal distribution function. This class is non-empty for positive distances, and would suffice for most of the purposes of Sections II through VI. One could similarly replace our heuristic language about approximate constants with a more formal treatment that was liberally seasoned with epsilons and deltas and rigorous inequalities. However, there is probably little to be gained by this formalism in the present context.]

These considerations above extend directly to several random variables given as in Fig. 1 and their corre-

sponding multidimensional distribution functions. Thus, all of the multidimensional probability theory given by Bennett<sup>2</sup> can be directly applied to our present circumstances. Of course, such distributions will generally depend on the present time  $t_1$  and the duration  $T$ , but for suitable (not necessarily long) values of the duration  $T$ , this dependence may be considered to be negligible for certain engineering purposes.

It is quite well known that averaging is a linear operation; *i.e.*, if  $x$  and  $y$  are random variables and  $a$  and  $b$  are constants, then  $\langle ax+by \rangle = a\bar{x} + b\bar{y}$ . This is clear when considered as time averages and, as an immediate consequence of (56), also holds for the corresponding distribution averages. (Let  $P_1$ ,  $P_2$ , and  $P_3$ , denote the distribution functions of  $f$ ,  $g$ , and  $f+g$ , respectively, and write (56) three times. This does not require independence of  $f$  and  $g$ .)

Although this fact is well known, such extensive use of it is made in the body of the paper that it is advisable to mention a few consequences here. First, if  $x_1, x_2, \dots, x_n$  are random variables, not necessarily independent, and  $a_1, a_2, \dots, a_n$  are constants and

$$y = \sum_{k=1}^n a_k x_k \tag{58}$$

then

$$\bar{y} = \sum_{k=1}^n a_k \bar{x}_k \tag{59}$$

In order to consider higher moments than the first, note the simple algebraic fact that  $(a_1x_1+a_2x_2)^2 = a_1^2x_1^2 + a_2^2x_2^2 + 2a_1a_2x_1x_2$  can be written  $a_1^2x_1^2 + a_2^2x_2^2 + a_1a_2x_1x_2 + a_2a_1x_2x_1$ . More generally, one can write

$$y^2 = \left[ \sum_{k=1}^n a_k x_k \right]^2 = \sum_{k=1}^n a_k^2 x_k^2 + \sum_{i \neq j} a_i a_j x_i x_j \tag{60}$$

Hence, by (59),

$$\overline{y^2} = \sum_{k=1}^n a_k^2 \overline{x_k^2} + \sum_{i \neq j} a_i a_j \langle x_i x_j \rangle \tag{61}$$

If the  $x_i$  are independent, then  $\langle x_i x_j \rangle = \bar{x}_i \bar{x}_j$  if  $i \neq j$ .<sup>40</sup> Then (61) becomes

$$\overline{y^2} = \sum_{k=1}^n a_k^2 \overline{x_k^2} + \sum_{i \neq j} a_i a_j \bar{x}_i \bar{x}_j \tag{62}$$

Hence, if the first two moments  $\bar{x}_i$  and  $\overline{x_i^2}$  of the  $x_i$  are known, the average square of  $y$  can be computed without even knowing the distribution of the  $x_i$ , much less the explicit distribution of  $y$ . It will be seen in several sections of this paper that these simple facts can lead to interesting and important results, some of which are by no means obvious.

Another well-known fact is that a joint density func-

tion<sup>2</sup>  $p(x, y)$  of two random variables can be integrated over a region in the plane to obtain the probability of the region. Thus, the probability that  $x_1 < x \leq x_2, y_1 < y \leq y_2$  is given by

$$\int_{x_1}^{x_2} \int_{y_1}^{y_2} p(x, y) dy dx, \tag{63}$$

and the joint distribution function  $P(x, y)$  is simply

$$P(x, y) = \int_{-\infty}^x \int_{-\infty}^y p(s, t) dt ds. \tag{64}$$

Notice that both (63) and (64) can be written in the form

$$\iint_E p(x, y) dx dy \tag{65}$$

of an integral of  $p(x, y)$  over a certain region  $E$  in the  $x$ - $y$  plane. In the case of (63), the region is an ordinary rectangle, while in the case of (64), it is a semi-infinite rectangle. The virtue of this geometric approach is that it often enables the expression of an event of practical interest in terms of such a region, not necessarily a "rectangle," after which the probability of the event in question can be computed by integrating the joint density function over the region. Joint density functions of three or more variables can similarly be integrated over regions in space of three or more dimensions. This is used at several places in the body of the paper.

Finally, a few words on computing the distribution functions of sums of independent random variables may be useful. It was pointed out by Bennett<sup>2</sup> that if  $x$  and  $y$  are independent random variables with density functions  $p_1$  and  $p_2$ , respectively, the density function  $p_3$  of the sum  $z = x + y$  is given by

$$p_3(z) = \int_{-\infty}^{\infty} p_1(z - y) p_2(y) dy, \tag{66}$$

which is called the "convolution" or "composition" of the density functions  $p_1$  and  $p_2$ . (This has sometimes been referred to as "combining" the  $x$  and  $y$  distributions.) However, in many—possibly most—practical applications, the distribution function of  $z$ , *i.e.*, the probability that  $z \leq u$ , is of more interest than the density function. This can be expressed in terms of the component distribution functions as

$$P_3(u) = \int_{-\infty}^{\infty} P_1(u - y) dP_2(y), \tag{67}$$

which can be seen from (66) by writing  $p_2(y) dy = dP_2(y)$  and integrating (66) on  $z$  from  $-\infty$  to  $u$ . The integral (67), which is well known among mathematicians,<sup>41</sup> may be defined as the limit of approximating sums of the form

$$\sum_k P_1(u - y_k) [P_2(y_k) - P_2(y_{k-1})], \tag{68}$$

<sup>40</sup> Bennett, footnote 2, p. 619.

<sup>41</sup> Cramer, *op. cit.*, p. 190, (15.12.2), and other sources.

where the  $y_k$  form a suitably fine partition of the range of interest. This and other numerical integration techniques may be used to evaluate (67) numerically, at least as easily as a numerical evaluation of (66). (See Appendix IV.) Hence, if the distribution functions of two variables are given numerically and the distribution function of the sum is required, there is no need to transform the given distribution functions into approximate density functions, evaluate (66) numerically, and then sum the resulting approximate density function, a procedure unnecessarily involved, but fairly commonly used.

APPENDIX II

THE SCHWARZ INEQUALITY

There are several ways of proving the inequality

$$\left[ \sum_{j=1}^N u_j v_j \right]^2 \leq \left[ \sum_{j=1}^N u_j^2 \right] \left[ \sum_{j=1}^N v_j^2 \right], \tag{19}$$

of which the simplest is perhaps to notice the algebraic identity

$$\begin{aligned} \left[ \sum_{k=1}^N u_k v_k \right]^2 &= \left[ \sum_{k=1}^N u_k^2 \right] \left[ \sum_{k=1}^N v_k^2 \right] \\ &\quad - \frac{1}{2} \sum_{i,j=1}^N (u_i v_j - v_i u_j)^2, \end{aligned} \tag{69}$$

which can be verified at once simply by expanding both sides. It is clear that (19) follows immediately from (69), for the term

$$\sum_{i,j=1}^N (u_i v_j - v_i u_j)^2$$

is obviously non-negative. However, more precise information can be extracted from this. Evidently there is equality holding in (19) if and only if every term in the double sum on the right in (69) vanishes, *i.e.*, if and only if

$$u_i v_j - v_i u_j = 0, \quad i, j = 1, 2, \dots, N, \tag{70}$$

and it is not difficult to see that this will happen if and only if there are constants  $a$  and  $b$ , not both zero, such that  $au_i = bv_i$ ,  $i = 1, 2, \dots, N$ . That is, equality holds in (19) if and only if the  $u$ 's and  $v$ 's are proportional. It is this fact that accounts for the "and only if" assertion following (14); the material of Section IV does not justify this assertion.

APPENDIX III

THE MAXIMAL-RATIO DISTRIBUTION

The characteristic function<sup>2</sup> of the  $p_j$  is

$$\begin{aligned} \phi(t) &= \int_{-\infty}^{\infty} e^{it p_j} g(p_j) dp_j = \int_0^{\infty} e^{-p_j + it p_j} dp_j \\ &= \frac{1}{1 - it}, \end{aligned} \tag{71}$$

where  $i = \sqrt{-1}$  is not an index. The characteristic function of  $p$  is then simply

$$\phi_N(t) = [\phi(t)]^N = \frac{1}{(1 - it)^N}, \tag{72}$$

so that the density function of

$$p = \sum_{j=1}^N p_j$$

is

$$g_N(p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ip t} \phi_N(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ip t}}{(1 - it)^N} dt. \tag{73}$$

The integral (73) is easily evaluated by contour integration and the residue theorem. The result is

$$g_N(p) = \frac{1}{(N - 1)!} p^{N-1} e^{-p} \tag{74}$$

for  $p > 0$ , while  $g_N(p) = 0$  for  $p < 0$ . This is precisely the result (29).

APPENDIX IV

COMPUTATION OF THE EQUAL-GAIN DISTRIBUTION

The function  $B_N(u)$  of (41) is given recursively by

$$\begin{aligned} B_N(u) &= \int_0^u B_{N-1}(u - t) dB_1(t) \\ &= \int_0^u B_{N-1}(u - t) B_1'(t) dt, \end{aligned} \tag{75}$$

as can be seen from (67), where  $B_1'(t) = 2te^{-t^2}$  is the Rayleigh density function. Tables of  $B_N(u)$  have been constructed<sup>19</sup> from (75) for  $N = 2, 3, \dots, 8$ , using an IBM 704 computer. Tables were constructed for various increments  $\Delta$  of  $u$ , ranging from  $\Delta = 0.2$  down to  $\Delta = 0.02$ , and for the range  $0 \leq u \leq 17$ . The Rayleigh density and distribution functions were generated in the computation program by rational approximations accurate to six decimals. Each value  $B_N(u_k)$ , where  $u_k = k\Delta$ , was then computed from

$$B_N(u_k) = \sum_{l=1}^k \int_{(l-1)\Delta}^{l\Delta} B_{N-1}(l) B_1'(u_k - t) dt, \tag{76}$$

where each integral over the range of length  $\Delta$  was computed by a 16-point Gaussian quadrature formula. The values of  $B_{N-1}(l)$  for this integration were obtained from the previously constructed table of  $B_{N-1}$  by a modified Tchebycheff-Everett interpolation formula. Tables of  $B_2(u)$  constructed by this method agreed with (39) (separately tabulated) to six decimals. For all  $N \leq 8$ , tables for the smaller values of  $\Delta$  were consistent to four decimals for the entire range of  $u$ .

An additional check was provided. The function defined by (41) for all complex values of  $u$  is an entire function, which therefore admits a power series,

$$B_N(u) = \sum_{k=0}^{\infty} b_k^N u^k, \tag{77}$$

valid for all values of  $u$ , and which coincides with the desired distribution function for positive real values of  $u$ . It is not especially difficult to show that  $b_k^N = 0$  for all odd  $k$  and for  $k < 2N$ . It can be shown on the basis of extensive computations from (41) that for even  $k \geq 2N$ ,

$$b_k^N = (-1)^{k/2-N} \frac{2^{k/2-l} l = (k/2)+1-N}{k!} \sum_{l=1} \rho_l \times \sum_{\substack{j_1 \geq 1 \\ j_1 + \dots + j_{N-1} = k/2-l}} \rho_{j_1} \rho_{j_2} \dots \rho_{j_{N-1}}, \tag{78}$$

where

$$\rho_l = (2l-1)(2l-3) \dots 5 \cdot 3 \cdot 1. \tag{79}$$

In particular, the coefficient  $b_{2N}^N$  of the leading term is  $b_{2N}^N = 2^N / (2N)!$ , i.e.,

$$B_N(u) = \frac{2^N}{(2N)!} u^{2N} + \dots, \tag{80}$$

but this term alone is not sufficiently accurate for useful values of  $u$ . For larger values of  $k$ , (78), which has resisted strenuous attempts at simplification, is not as useful for the explicit computation of coefficients as the recursion relation

$$b_k^N = (-1)^{k/2-N} \frac{1}{k!} \sum_{j=1}^{j=(k/2)-1} 2^j \rho_j (k-2j)! |b_{k-2j}^{N-1}|, \tag{81}$$

which can also be established from (41). In connection with (81), one uses  $b_k^1 = (-1)^{(k/2)-1} / (k/2)!$ .

It can be shown from (78) that for  $|u| \leq u_0$  and for

$$k \geq 2 \left( N - 1 + \frac{1}{N} \right) u_0^2 - 2, \tag{82}$$

the terms of (77) are monotonically decreasing in magnitude. Since the terms alternate in sign, this means that the error in terminating (77) at the  $k$ th term is less than the magnitude of the  $k$ th term, provided (82) is satisfied. This was used to construct a table of  $B_N(u)$  for  $N=2, 3, \dots, 8$  for  $0 \leq u \leq 1.5$  with a guaranteed accuracy of six decimals. This table agreed in this range with that constructed from (76) to six decimals or better.

The results (78), (81), and (82) are principally due to Michael Ginsburg.

#### APPENDIX V

##### CERTAIN QUESTIONS RELATED TO THE PROBLEM OF CORRELATED FADING

All of the presently published treatments of correlated fading known to the present writer rely on a result due to Uhlenbeck,<sup>42</sup> which was reproduced in a

<sup>42</sup> G. E. Uhlenbeck, Rad. Lab., M.I.T., Cambridge, Mass., Rep. No. 454; October 15, 1943.

paper by Booker, Ratcliffe, and Shinn.<sup>43</sup> Uhlenbeck's result rests in turn on the joint distribution of two Rayleigh variables given by Rice.<sup>44</sup> However, meaningful sufficient conditions under which this distribution is applicable to correlated Rayleigh fading do not appear to be known. It is essentially certain that it is applicable to narrow-band random noise of the type originally studied by Rice and Uhlenbeck, but it is far from clear that it is equally applicable to fading radio waves in general. For example, if Uhlenbeck's result always held, then the correlation of two Rayleigh variables could not be negative; however, several investigators, including Grisdale, *et al.*,<sup>45</sup> and McNicol<sup>46</sup> have found such negative correlation. It seems very probable that Uhlenbeck's result is satisfactory as a first approximation for engineering purposes. This is why Fig. 16 was unhesitatingly included in Section VII; however, such results should be understood as representative, rather than absolute. In other words, the correlation coefficient does not uniquely determine diversity performance, even when  $\rho$  and the separate input distributions are known.

A closely related problem that sometimes arises in this connection is the assumption, which has not always been recognized as such, that two random variables that are individually Gaussian or normal have a joint distribution that is a two-dimensional normal distribution. This assumption is the basis of the common statement that "uncorrelated normal variables are independent." As a mathematical matter, this need not be true; it is not difficult to give counter-examples. The prevalence of the quoted statement probably stems in part from some insufficiently explicit language of Cramér.<sup>46</sup> The two-dimensional form of the central limit theorem suggests that this assumption would often be very reasonable, but should be recognized as an assumption.

#### APPENDIX VI

##### SELECTION DIVERSITY MEAN POWER RATIOS

In the integral (17) for  $\bar{p}(N)$ ,

$$\bar{p}(N) = N \int_0^{\infty} p(1 - e^{-p})^{N-1} e^{-p} dp, \tag{17}$$

we make the change of variable  $y = 1 - e^{-p}$ , obtaining

$$\bar{p}(N) = N \int_0^1 [-\log(1-y)] y^{N-1} dy. \tag{83}$$

<sup>43</sup> H. G. Booker, J. A. Ratcliffe, and D. H. Shinn, "Diffraction from an irregular screen with applications to ionospheric problems," *Phil. Trans. Royal Soc. (London) A*, vol. 242, pp. 579-607; September, 1950.

<sup>44</sup> S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 24, pp. 46-156; January, 1945. See (3.7-13).

<sup>45</sup> McNicol, *op. cit.*, Fig. 4.

<sup>46</sup> Cramér, *op. cit.*, p. 289.

Using the series

$$\begin{aligned}
 -\log(1-y) &= \sum_{k=1}^{\infty} y^k/k \text{ for } |y| < 1, \\
 \bar{p}(N) &= N \int_0^1 \left[ \sum_{k=1}^{\infty} \frac{y^{k+N-1}}{k} \right] dy \\
 &= \sum_{k=1}^{\infty} \frac{N}{k(k+N)} \\
 &= \sum_{k=1}^{\infty} \left( \frac{1}{k} - \frac{1}{N+k} \right) \\
 &= \lim_{m \rightarrow \infty} \left[ \sum_{k=1}^m \left( \frac{1}{k} - \frac{1}{N+k} \right) \right] \\
 &= \lim_{m \rightarrow \infty} \left[ \sum_{k=1}^N \frac{1}{k} - \sum_{k=1}^m \frac{1}{m+k} \right] \\
 &= \sum_{k=1}^N \frac{1}{k}, \tag{84}
 \end{aligned}$$

which is the result (18). The termwise integration from the first to second line of (84) is easily justified.

The result  $\bar{p}(N) = \sum_{k=1}^N 1/k$  was originally found essentially by accident and verified by induction on  $N$ .<sup>47</sup> Another direct approach was also suggested by Stein, who pointed out that (17) could be written

$$\bar{p}(N) = -\frac{\partial}{\partial x} \int_0^{\infty} N(1-e^{-p})^{N-1} e^{-xp} dp \Big|_{x=1}, \tag{85}$$

which, integrating by parts  $(N-1)$  times, becomes

$$\bar{p}(N) = -\frac{\partial}{\partial x} \left[ \frac{(x-1)N!}{(x+N-1)!} \right] \Big|_{x=1}. \tag{86}$$

Stein remarked that this is

$$\bar{p}(N) = -N \frac{\partial}{\partial x} \beta(N, x) \Big|_{x=1} \tag{87}$$

where  $\beta(N, x)$  is the Beta function [substitute  $t = e^{-p}$  in (85)], and that higher moments of  $p$  are given by successive derivatives of the same function. The differentiation indicated in (86) is straightforward.

ACKNOWLEDGMENT

I am indebted to the several individuals who either provided helpful comments on footnote 1, furnished data or results for inclusion in the present paper, carefully read and criticized the manuscript, or did one or more of the foregoing. These included W. R. Bennett of the Bell Telephone Laboratories, E. J. Baghdady and John Granlund of the Department of Electrical Engineering, M.I.T., F. E. Bond of the Signal Corps Engineering Laboratories, H. F. Meyer of SCEL, William Sichak of the ITT Laboratories, Harold Staras

of the RCA Laboratories, Princeton, Seymour Stein of Hycon Eastern, Inc., and several colleagues at Lincoln Laboratory. Among these, I should especially like to thank William Sichak, who contributed very generously to this paper under each of the indicated headings.

I am also indebted to my colleagues W. C. Mason and Michael Ginsburg for the use of the equal-gain distributions in Figs. 9-13 before the complete tables of Mason, *et al.*, have been published. Phyllis Bloom and Marguerite Glynn of the Division 3 Computing Section of Lincoln Laboratory provided substantial assistance with the material in Appendix IV and prepared several of the curves.

BIBLIOGRAPHY

The following list includes most of the theoretical papers related to diversity techniques published since 1947, together with two papers of historical interest. The comments following the references from [3] onwards are those of the present author and are not intended to be complete summaries of the papers involved. (Additional references dealing primarily with experimental results or instrumentation techniques are included among the footnotes.)

- [1] H. H. Beverage and H. O. Peterson, "Diversity receiving system of RCA Communications, Inc., for radiotelegraphy," *PROC. IRE*, vol. 19, pp. 531-561; April, 1931.
- [2] H. O. Peterson, H. H. Beverage, and J. B. Moore, "Diversity telephone receiving system of RCA Communications, Inc.," *PROC. IRE*, vol. 19, pp. 562-584; April, 1931.
- [3] Z. Jelonek, E. Fitch, and J. H. H. Chalk, "Diversity reception, statistical evaluation of possible gain," *Wireless Engr.*, vol. 24, pp. 54-62; February, 1947. Selection diversity for the cases 1) uncorrelated Rayleigh fading and 2) partially correlated shallow Gaussian fading for two channels is analyzed. The first case is also considered at the low end of the distribution for unequal mean signals.
- [4] A. H. Hausman, "An analysis of dual diversity receiving systems," *PROC. IRE*, vol. 42, pp. 944-947; June, 1954. A comparison of two-channel selection diversity and scanning diversity in terms of unspecified general fading distributions is presented but no concrete results.
- [5] L. R. Kahn, "Ratio squarer," *PROC. IRE*, vol. 42, p. 1704; November, 1954. This is a note pointing out that two-channel maximal-ratio diversity could outperform two-channel selection diversity. Instrumentation and comparative tests are discussed.
- [6] D. G. Brennan, "On the maximal signal-to-noise ratio realizable from several noisy signals," *PROC. IRE*, vol. 43, p. 1530; October, 1955. Statement and proof of the theorem are embodied in (13) and (14) of the present paper.
- [7] H. Staras, "Diversity reception with correlated signals," *J. Appl. Phys.*, vol. 27, pp. 93-94; January, 1956. An ingenious method of evaluating the effect of partially correlated Rayleigh fading in two-channel selection diversity systems is presented with results that are also useful in measuring such correlation.
- [8] F. J. Altman and W. Sichak, "A simplified diversity communication system for beyond-the-horizon links," *IRE TRANS. ON COMMUNICATIONS SYSTEMS*, vol. CS-4, pp. 50-55; March, 1956. Several comparative results on selection, maximal-ratio and equal-gain diversity systems, including a demonstration that equal-gain systems perform almost as well as maximal-ratio systems in the presence of Rayleigh fading, are presented. The importance of this paper has not yet been widely recognized, perhaps because it is in a journal with a very limited circulation and is so concise that even experts have failed to appreciate its contents. Apart from the papers by Adams and Mindes, and Morrow, *et al.*, and a passing reference by Staras, it is not even mentioned in any of the following references of this paper.
- [9] H. Staras, "The statistics of combiner diversity," *PROC. IRE*, vol. 44, pp. 1057-1058; August, 1956. Curves of the distribution functions for maximal-ratio diversity and a method of approximating very-long-term distributions of the local SNR are presented. This method would be most accurate for systems with a large number of channels. Combiner diversity here means maximal-ratio diversity. See the end of Section IV following (33) in this paper.
- [10] J. N. Pierce, "Diversity Improvement in Frequency-Shift Keying for Rayleigh Fading Conditions," *Air Force Cambridge Res. Center, Bedford, Mass., Tech. Rep. No. 56-117*; September, 1956. Results applicable to certain types of binary signals and four combining methods, including maximal-ratio and selection diversity, are considered. See [15].

<sup>47</sup> The procedure used here was suggested by one of the IRE reviewers.

[11] K. H. Schmelovsky, "Einfluss der Korrelation zwischen Empfangsfeldstärken bei Diversity-Empfang," *Hochfrequenz. Elektr.*, vol. 65, pp. 74-76; November, 1956. (Translations of Friedman available at the John Crerar Library.) A method of evaluating the effect of partially correlated Rayleigh fading in two-channel selection diversity systems is discussed.

[12] F. E. Bond and H. F. Meyer, "The effect of fading on communication circuits subject to interference," *PROC. IRE*, vol. 45, pp. 636-642; May, 1957. An analysis of two-channel selection diversity for the case of Rayleigh-variable signals in Rayleigh-variable noise is presented. This case often arises in practice in communication systems operating at frequencies below 30 mc. This case was also considered for a selection diversity system in which the selection is of the channel with the greatest signal-plus-noise, in contrast to the greatest SNR.

[13] E. Henze, "Theoretische Untersuchungen über einige Diversity-Verfahren," *Arch. Elekt. Übertragung*, vol. 11, pp. 183-194; May, 1957. (Translations by M. D. Friedman available from the John Crerar Library, 86 E. Randolph St., Chicago 1, Ill.) A consideration of two-channel selection diversity and scanning diversity, including both uncorrelated and partially correlated Rayleigh fading

cases, is presented. Several of the derivations are extraordinarily complicated and can be greatly simplified.

[14] K. S. Packard, "Effect of correlation on combiner diversity," *PROC. IRE*, vol. 46, pp. 362-363; January, 1958. Combiner diversity here means maximal-ratio diversity. A method of estimating the effect of partially correlated Rayleigh fading in two-channel maximal-ratio diversity systems is discussed.

[15] J. N. Pierce, "Theoretical diversity improvement in frequency-shift keying," *PROC. IRE*, vol. 46, pp. 903-910; May, 1958. Error rates for various combining techniques are derived and the effects of correlated fading in dual-diversity binary systems are considered. See also additional references pertinent to binary signal systems given in this paper.

[16] R. T. Adams and B. M. Mindes, "Evaluation of IF and baseband diversity combining receivers," *IRE TRANS. ON COMMUNICATION SYSTEMS*, vol. CS-6, pp. 8-13; June, 1958. A comparison of predetection equal-gain combining with postdetection maximal-ratio combining for a two-channel FM system is presented, with a theoretical and experimental demonstration of the reduction in FM multipath distortion achieved by using predetection equal-gain combining.

## Physical Principles of Avalanche Transistor Pulse Circuits\*

D. J. HAMILTON†, ASSOCIATE MEMBER, IRE, J. F. GIBBONS†, AND W. SHOCKLEY†, FELLOW, IRE

**Summary**—A simple physical theory is developed which permits a calculation of the significant points of avalanche transistor transient behavior.

A model for the transistor is defined in terms of charge variables and the physical parameters of the device. The transient performance of the model is calculated by focusing attention on the minority carrier charge stored in the base region and the influence of base-width modulation upon this stored charge. In the charge formulation of the problem, the physical details of the avalanche multiplication process need not be considered; multiplication is accounted for by the boundary conditions which it imposes upon the stored charge.

Good agreement has been obtained between calculated and experimentally observed data for a simple avalanche transistor relaxation oscillator.

### I. INTRODUCTION

TRANSISTORS exhibiting avalanche multiplication have recently been shown<sup>1</sup> to be useful for the generation of millimicrosecond pulses. These devices thus provide a new and simple solution to a problem which previously taxed the ingenuity of both circuit and device designers.

\* Original manuscript received by the IRE, December 28, 1959; revised manuscript received, March 6, 1959. This work was supported in part by the U. S. Army Signal Corps, the U. S. Air Force, and the U. S. Navy, through ONR Contract Nonr 225(24) at Stanford Electronics Laboratories.

† Stanford Electronics Labs., Stanford University, Stanford, Calif.

<sup>1</sup> G. B. B. Chaplin, "A method of designing transistor avalanche circuits with application to a sensitive transistor oscilloscope," paper presented at 1958 Solid-State Circuits Conference, Philadelphia, Pa.; February, 1958.

As is frequently the case, however, a simple empirical solution poses difficult analytical problems. These analytical difficulties arise primarily from a failure to recognize the important physical principles which govern the terminal behavior of the device. When the problem is properly formulated, many of the analytical complications are removed, and a simple unified theory is obtained. It is the purpose of this paper to present such a theory.

The most significant aspect of the theory is the concept of minority carrier charge stored in the base region during the transient period, a concept which results in a considerable simplification of the problem by permitting time to be eliminated in several of the calculations.<sup>2</sup> A relaxation oscillator (Fig. 1) is used to illustrate the theory.

Section II of the paper deals with the circuit model for the relaxation oscillator, together with a physical model for the avalanche transistor; the concept of stored minority carrier charge is also introduced.

In Section III, the stored charge concept is used to determine the two critical values of external capacitance: the capacitance required to start regeneration and the capacitance required to forward bias the collector junction.

<sup>2</sup> W. Shockley and J. Gibbons, "Theory of transient build-up in avalanche transistors," *Trans. AIEE (Commun. and Electronics)*, no. 40, pp. 993-998; January, 1959.

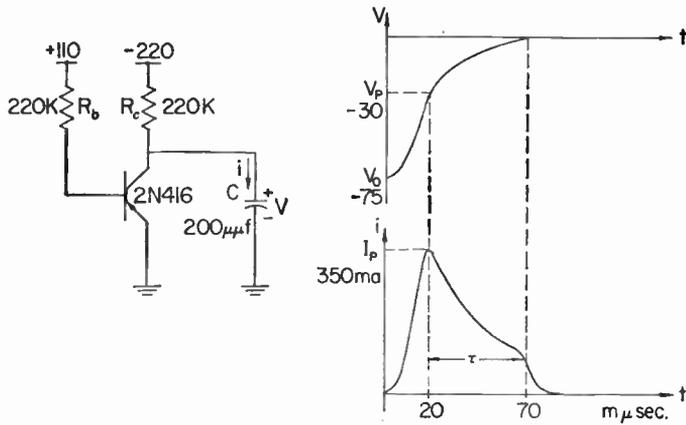


Fig. 1—Circuit and waveforms showing typical values for the avalanche transistor relaxation oscillator.

Section IV describes the detailed shape of the current pulse. Once again stored charge is used to calculate the peak current; a very simple method of approximating the rise-to-peak current is given; and the effect of base-width modulation is used to calculate the time after peak current required for the collector to become forward biased.

In Section V, observed values are compared with values calculated from the theory.

The significant results are:

- 1) The capacitance required to start regeneration varies inversely with the initial multiplication. However, the capacitance required to forward bias the collector junction depends critically upon the exponent  $n$  in

$$M = \frac{1}{1 - \left(\frac{V_c}{V_B}\right)^n}$$

and has a non-zero value even for infinite initial multiplication.

- 2) Regarding the detailed shape of the current pulse:
  - a) The maximum current is a function of the minority carrier charge that has been stored when the effective multiplication reaches unity.
  - b) The rise to peak current is approximately a short exponential rise followed by a linear rise. The exponential portion is given by the solution of the diffusion equation; a lower bound on the rate of rise of the linear portion is found by assuming the entire rise to be linear.
  - c) The fall time is determined by the extent to which base-width modulation occurs. It is constant over a fairly wide range of capacitances.
- 3) The agreement between calculated and observed values is good in most cases. Adequate qualitative reasons can be given for existing departures.

## II. THE AVALANCHE MODEL

The operation of the relaxation oscillator is best visualized by assuming that the initial voltage across the capacitor is zero. The collector and emitter junctions are reverse biased with one or both junctions exhibiting avalanche multiplication. For simplicity it is assumed here that the emitter current is zero, only the collector junction exhibits multiplication, and the collector-to-base voltage is very nearly  $V_B$ , the breakdown voltage. Resistor  $R_c$  supplies the current required by the collector junction, and additional current to charge  $C$  negatively. As  $C$  charges, the collector to base voltage remains approximately  $V_B$ , so that when the voltage across  $C$  is  $V_B$ , the emitter junction becomes forward biased and injection into the base begins. It is at this time that the transient shown in Fig. 1 takes place. A regenerative action ensues during which the transistor-capacitor loop current builds up rapidly and the capacitor discharges. If the capacitor is not large enough, it will not be able to "bottom" the transistor; *i.e.*, drive the collector voltage to zero or forward bias.

In the circuit under consideration, the bias currents flowing in  $R_b$  and  $R_c$  are orders of magnitude smaller than the current in the transistor-capacitor loop during the period of interest. The function of these bias currents is merely to set the stage for the regenerative process; they contribute little to its actual development, and may therefore be omitted in studying the essential features of the transient build-up. The circuit model to be studied is shown in Fig. 2 and consists simply of a transistor and a capacitor.

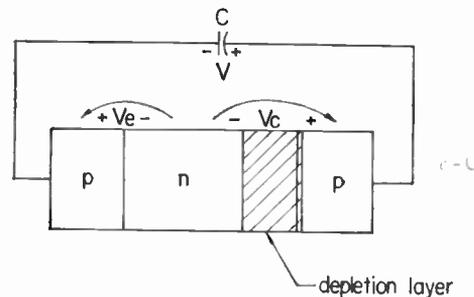


Fig. 2—Circuit model for the relaxation oscillator.

The plan of attack for the analysis of the circuit during the transient period is as follows: the transistor is represented by a diffusion model; *i.e.*, a model in which current flow in the base region is governed by diffusion of minority carriers. Four basic equations are then written in terms of charge variables for the diffusion model. These equations do not contain time. From these equations, the charge of minority carriers stored in the base region during the transient period is calculated. The stored charge is then used in the calculation of critical values of circuit capacitance, peak current, rate of rise, and fall-time.

General assumptions will be made to simplify the analysis. These are:

- 1) No carrier recombination occurs in the base during the time interval of interest.
- 2) Avalanche multiplication occurs only at the metallurgical junction and is described by the relation

$$M = 1 / 1 - \left( \frac{V_c}{V_B} \right)^n$$

- 3) All of the voltage appearing across the load capacitor  $C$  appears across the collector depletion region; *i.e.*,

$$v_c \cong V, \quad v_e \cong 0.$$

Under these conditions, a very simple physical model may be formulated for dealing with the detailed nature of the transients. The model is illustrated in Fig. 3. As is apparent from the figure, major emphasis is placed on charge variables. The following increments in the charge variables may be defined:

- 1)  $dQ$  is the incremental charge flowing in the external circuit.
- 2)  $dQ_s$  is the incremental charge of minority carriers stored in the base.
- 3)  $dQ_{mc}$  is the incremental charge of minority carriers arriving by diffusion at the base side of the depletion layer.
- 4)  $dQ_D$  is the incremental charge deposited in the collector side of the depletion layer.

With these definitions, the following equations may be written:

$$dQ_s = \gamma dQ - dQ_{mc} \tag{1}$$

$$M dQ_{mc} = dQ + dQ_D \tag{2}$$

$$dQ_D / C_c = dQ / C \tag{3}$$

$$dQ_{mc} / \gamma dQ \equiv \beta \tag{4}$$

where  $C_c$  is the collector capacitance and  $\beta$  is the transport factor under transient conditions for reverse collector voltage.

Eq. (1) expresses charge conservation in the base and relies for its validity on assumption 1); (2) expresses charge conservation in the depletion region; (3) is the result of assumption 3); and (4) is the definition of the transport factor.

These four simple equations are the basis of the theory to be developed.

### III. ROLE OF THE CIRCUIT CAPACITANCE

As is implied in the introduction, the stored charge concept may be used conveniently to calculate certain capacitances which are critical in the operation of the relaxation oscillator circuit. The capacitances of interest are the minimum circuit capacitance required to start regeneration,  $C_R$ , and the minimum circuit capacitance,  $C_{min}$ , required to forward bias the collector junction

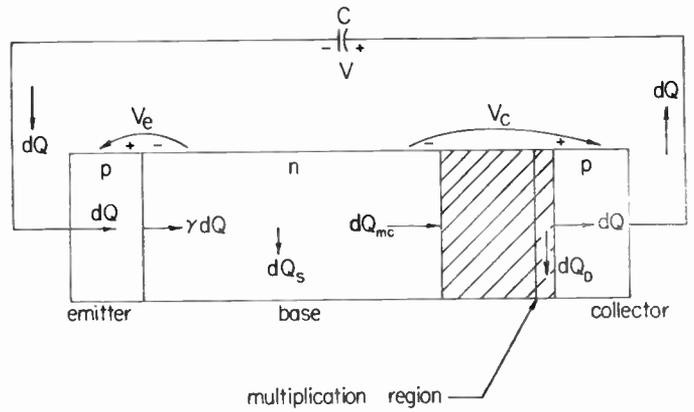


Fig. 3—Diffusion model showing charge variables for the avalanche transistor.

once regeneration is started. The calculation of these capacitances proceeds directly from (1)–(3), and does not require a consideration of time, a fact which results in considerable simplification of the analysis.

Using (2) and (3) to eliminate  $dQ_{mc}$ , (1) may be written as

$$dQ_s = \left( \gamma - \frac{1}{M_e} \right) dQ \tag{5}$$

where

$$M_e = \frac{MC}{C + C_c} \tag{6}$$

is called the effective multiplication factor.

It may be seen immediately from (5) that when

$$\gamma M_e = 1, \tag{7}$$

$dQ_s$  is zero. It is possible that  $dQ_s$  may equal zero for  $M_0$ , the initial value of  $M$ , and, hence, no charge would ever be stored, and the regeneration process would never start.

It therefore follows from (7) that a circuit capacitance of

$$C \geq C_R = \frac{C_c}{\gamma M_0 - 1} \tag{8}$$

is required to start regenerative action.

Eq. (5) may also be used to calculate the “bottoming” capacitance  $C_{min}$ . This may be done by realizing that

$$M_e = \frac{CM(V)}{C + C_c(V)}$$

and

$$dQ = CdV$$

so that

$$Q_s = \int_{V_0}^V dQ_s = \int_{V_0}^V \left( \gamma - \frac{C + C_c(V)}{CM(V)} \right) CdV. \tag{9}$$

Assuming that the collector-base junction is a step transition from uniform *p*-type material to uniform *n*-type material,  $C_c$  may be written approximately as

$$C_c = C_{cB}(V_B/V)^{1/2} \tag{10}$$

where  $C_{cB}$  is the collector capacitance at the breakdown voltage. Using this and the form of  $M$  given in assumption 2),

$$Q_s = \frac{CV_B}{n+1} \left[ \left( \frac{V}{V_B} \right)^{n+1} - \left( \frac{V_0}{V_B} \right)^{n+1} \right] - 2C_{cB}V_B \left[ \left( \frac{V}{V_B} \right)^{3/2} - \left( \frac{V_0}{V_B} \right)^{3/2} \right] + \frac{2}{2n+1} C_{cB}V_B \left[ \left( \frac{V}{V_B} \right)^{n+1/2} - \left( \frac{V_0}{V_B} \right)^{n+1/2} \right] + (\gamma - 1)C(V - V_0) \tag{11}$$

where  $V_0$  = initial voltage across  $C$ . The following reasoning is now used to obtain  $C_{min}$ . When  $Q_s$  is zero, no current can flow in the circuit because there can be no minority carrier charge gradient in the base. There exists a minimum value of  $C$  which will make  $Q_s$  in (11) zero just when  $V$ , the capacitor voltage, becomes zero. This must be  $C_{min}$ , the minimum value of capacitance required to forward bias the collector junction. Setting  $Q_s$  and  $V$  equal to zero in (11) and noting that in general  $n > 3$

$$C_{min} \cong \frac{2C_{cB}}{\frac{(V_0/V_B)^{n+1/2}}{n+1} - (1-\gamma)\left(\frac{V_0}{V_B}\right)^{1/2}} \tag{12}$$

An interesting case results if  $V_0 = V_B$  and  $\gamma = 1$ :

$$C_{min} \cong (2n + 1)C_{cB}.$$

This shows that a non-zero value of  $C_{min}$  exists even when  $M_0 = \infty$  ( $V_0 = V_B$ ), and it depends critically on  $n$ .<sup>3</sup> A qualitative understanding of this is easily obtained by considering the behavior of  $M$  vs  $V/V_B$  for two extreme cases,  $n = 1$  and  $n = \infty$ . These are shown in Fig. 4. From (5) it is evident that the ratio  $dQ_s/dQ$  is never greater than  $[1 - (1/M)]$ . For the case  $n = \infty$ , an infinitesimal change  $dV$  in voltage results in an  $M$  of unity and no charge is being stored. On the other hand, if  $n = 1$ , some charge will be stored even when  $V/V_B$  is much less than 1. Thus it is expected that  $C_{min}$  will increase as  $n$  is increased.

#### IV. THE SHAPE OF THE CURRENT PULSE

In addition to its use in the calculation of critical capacitances, the theory embodied in (1)-(4) provides a convenient basis for obtaining the detailed shape of the

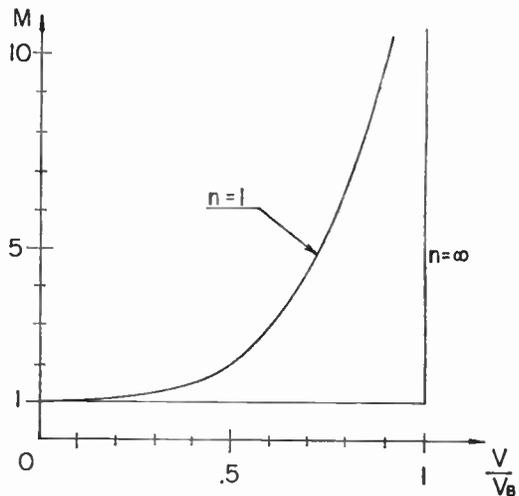


Fig. 4—Illustration of the dependence of  $M$  on  $n$ .

current pulse. It should be pointed out that significant changes occur in the pulse as  $C$  is varied. In particular, if  $C = C_{min}$ , the pulse is reasonably symmetrical about its maximum. If  $C > C_{min}$ , the pulse assumes the shape given in Fig. 1. For extremely large values of  $C$ , the pulse may differ considerably from that shown, and for this case, the theory does not apply. For simplicity in the calculations of peak current, rise time, and fall time which follow, the second of these conditions ( $C > C_{min}$ ) will be assumed.

#### Calculation of the Peak Current

A solution of the diffusion equation shows that if the base width does not change during the rise-to-peak current, the charge distribution in the base will be very nearly linear at peak current. On the other hand, if the base width changes appreciably (as in the case of a transistor operating near punch through) the peak current will occur before the charge reaches a linear distribution. For simplicity, the first case, that of a linear charge distribution, will be assumed.<sup>4</sup> The calculation of peak current, then, involves finding the voltage at peak current, calculating the stored charge; and calculating its gradient. The charge distribution is shown in Fig. 5.

With the assumption of a linear distribution, the incremental stored charge  $dQ_s$  is zero at peak current. If  $dQ_s$  were positive, an increase in stored charge would cause an increase in the gradient and, hence, in the current; if  $dQ_s$  were negative, a decrease in stored charge would cause a decrease in the gradient.

Setting  $dQ_s = 0$  in (7), the multiplication  $M$  (which will be almost unity) is found to be

$$M = \frac{C + C_c}{\gamma C}.$$

<sup>3</sup> J. R. A. Beale, W. L. Stephenson, and E. Wolfendale, "A study of high-speed avalanche transistors," *Proc. IEE*, pt. B, vol. 104, p. 398; July, 1957.

<sup>4</sup> For the second case, the approximate distribution can be found from a diffusion equation solution.

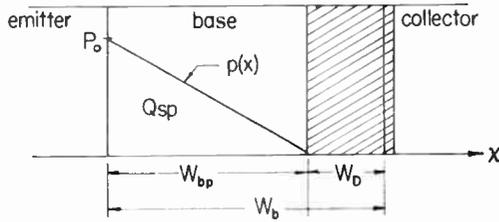


Fig. 5—Distribution of the stored charge at peak current.

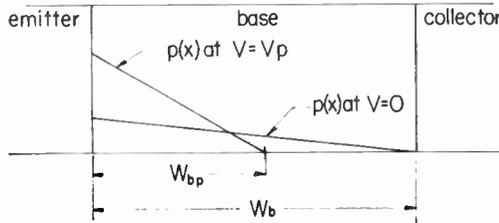


Fig. 6—Redistribution of stored charge after peak current.

If it is assumed that  $C \gg C_c$ , then  $M = 1/\gamma$ . The voltage  $V_p$  at peak current is found by using this result in the expression for  $M$ :

$$V_p = V_B(1 - \gamma)^{1/n} \tag{13}$$

Thus far it has been assumed that no recombination is present. If recombination exists, the steady state value of the transport factor at peak current is  $\beta_0$  and  $\gamma$  may be replaced by  $\gamma\beta_0 = \alpha_0$ .

As shown in Fig. 6, the base is assumed to be partially depleted, and the base width at peak current is denoted by  $W_{bp}$ . Since a step junction is postulated,  $W_{bp}$  is given by

$$W_{bp} = W_b - W_1\sqrt{-V_p}$$

where  $W_1$  is the width constant of the junction.

Assuming that  $C \gg C_c$ , the minority carrier charge  $Q_{sp}$  which has been stored when the current reaches its peak value is obtained from (11):

$$Q_{sp} \cong \frac{CV_B}{n+1} \left( \frac{V_0}{V_B} \right)^{n+1} \tag{14}$$

The peak current  $I_p$  is easily determined from the gradient of  $Q_{sp}$ :

$$I_p = \frac{2Q_{sp}D_e}{W_{bp}^2} \tag{15}$$

where  $D_e$  is the effective diffusion constant for minority carriers in the base.

(In many transistors sufficiently large currents exist when  $C \gg C_c$  to make  $p_0$ , the minority carrier density at the emitter junction, much greater than the impurity density  $N_D$  in the base. An appreciable field exists into the base region, and causes the same effect as an increase in the diffusion constant. Thus

$$D_e \cong D \left( 1 + \frac{p_0}{p_0 + N_D} \right) = 2D.^5$$

*The Rise to Peak Current*

During the initial stages of current build-up, if the multiplication remains essentially constant, the current and charge variables will exhibit a time dependence of the form  $\exp \lambda t$ . Using (2), (3) and (4) of Section II, the relation

$$M_e\beta\gamma = 1 \tag{16}$$

may be obtained. It is to be emphasized that  $\beta$  is the transport factor for transient conditions. It has the form<sup>6</sup>

$$1/\beta = \cosh W_1 \sqrt{\frac{\lambda}{D}}$$

Substitution of (16) in the expression for  $\beta$  yields

$$\lambda = \frac{D}{W_1^2} (\cosh^{-1} \gamma M_e)^2$$

When the multiplication varies during the current pulse (as it will for the relaxation oscillator), detailed analysis shows that the rise eventually becomes linear.<sup>2</sup> It is, in fact, observed that an appreciable portion of the current rise is linear. An estimate of a minimum value for  $di/dt$  during this linear portion can very simply be made by assuming a linear rise from zero to peak current. During the rise, the capacitor discharges from  $V_0$  to  $V_p$ , and the total charge removed from the capacitor must be

$$Q = C(V_0 - V_p) = \frac{I_p^2}{2 \frac{di}{dt}} \tag{17}$$

Substitution of (14) and (15) into (17) shows that  $di/dt$  is directly proportional to  $C$ . This has been experimentally observed over a fairly wide range of values of  $C$ .

*The Fall Time<sup>7</sup>*

When the peak current occurs, minority carrier charge  $Q_{sp}$  is stored in the base and is linearly distributed. As the voltage across the capacitor changes, the base width increases, causing a redistribution of the charge. A solution of the diffusion equation shows that the charge distribution remains essentially linear because  $Q_{sp}$

<sup>5</sup> See, for example, W. M. Webster, "On the variation of junction transistor current amplification with emitter current," *Proc. IRE*, vol. 42, pp. 914-920; June, 1954.

<sup>6</sup> W. Shockley and J. Gibbons, "Current build-up semiconductor devices," *Proc. IRE*, vol. 46, pp. 1947-1949; December, 1958.

<sup>7</sup> A similar approach has been used for punch-through transistors by H. Stutz and R. A. Pucel, "The spacistor, a new class of high frequency semiconductor devices," *Proc. IRE*, vol. 45, pp. 317-324; March, 1957.

reaches its equilibrium distribution much more rapidly than changes in base width occur. In addition to the distribution of  $Q_{sp}$  remaining linear, the following assumptions are made in calculating the fall time:

- 1)  $\gamma = 1$ ,
- 2)  $C \gg C_c$ . Thus the charge  $Q_D$  required by the depletion layer capacity is negligible in comparison to  $Q_{sp}$ , and  $Q_{sp}$  is conserved.

The change in voltage and the redistribution of charge are governed by

$$C \frac{dV}{dt} = I = \frac{2Q_{sp}D_c}{W(t)^2} \tag{18}$$

If the collector-base junction is a step  $p-n$  junction,

$$W = W_b - W_1\sqrt{-V} \tag{19}$$

where  $W_1 =$  width constant.

A simultaneous solution of (18) and (19) gives the voltage as a function of time after peak current:

$$\left[ W_b^2 V \left( 1 - \frac{4}{3} \frac{W_1}{W_b} \sqrt{-V} - \frac{W_1^2}{W_b^2} \frac{V}{2} \right) \right]_{V_p}^V = \frac{2Q_{sp}D_c}{C} t. \tag{20}$$

The fall time, defined here as the time elapsed while the voltage changes from  $V_p$  to 0, is found by using zero as the upper limit of integration in (20) and using (14) for  $Q_{sp}$ :

$$\tau = \tau_b(n + 1) \frac{V_p}{V_B} \left( 1 - \frac{4}{3} x + \frac{x^2}{2} \right) \left( \frac{V_B}{V_0} \right)^{n+1} \tag{21}$$

where

$$\tau_b = \frac{W_b^2}{2D_c}$$

and

$$x = \frac{W_1\sqrt{-V_p}}{W_b}.$$

Eq. (21) indicates that  $\tau$  is independent of  $C$ ; Fig. 7 shows the range of values of  $C$  for which this is true. Qualitatively the variation of  $\tau$  with  $C$  is explained as follows:

If the stored charge remains constant after peak current, the current fall will be determined by the variation of base width with time, and  $\tau$  will have its minimum value which is given by (21). However, if the stored charge decreases with time, an additional decrease of current results and a longer time is required to discharge the capacitor.

For  $C$  only slightly larger than  $C_{min}$ ,  $Q_D$  is not negli-

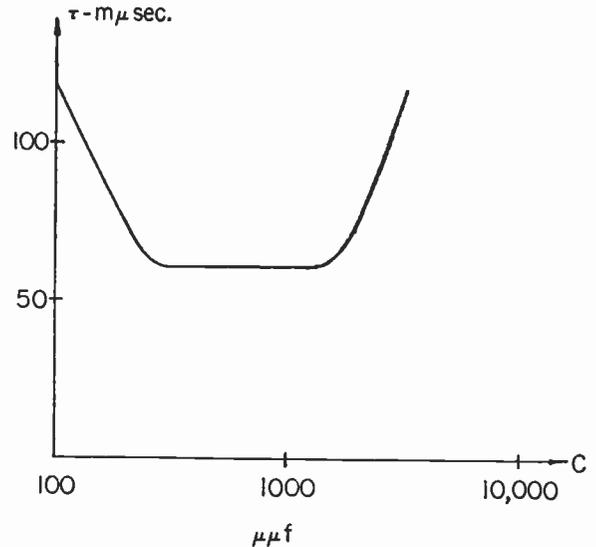


Fig. 7—Experimentally observed variation of fall time with external circuit capacitance.

ble in comparison to  $Q_{sp}$ , and the stored charge decreases causing, an increase in  $\tau$ . If  $C$  is made very large, the emitter efficiency  $\gamma$  decreases, resulting in a reduction of stored charge and an increase in  $\tau$ .

#### V. "THE PROOF OF THE PUDDING . . ."

The ultimate test of any physical theory is, of course, its agreement with observation. Before discussing the experimental results, however, it is worthwhile to comment briefly on the more important aspects of the theory.

It is significant that for circuit analysis, the complex physical details of the avalanche multiplication process can be simply represented as follows:

- 1) Multiplication occurs instantaneously and only at the metallurgical junction.
- 2) It is dependent only upon the junction voltage, in accordance with the empirical static multiplication relation.

Elimination of the details of the multiplication process thus permits the formulation of the problem in terms of charge variables, with multiplication serving only to determine boundary conditions. The result is the simple and straightforward theory just presented.

Table I gives the quantities  $V_B$ ,  $n$ ,  $C_c$ ,  $\alpha_0$ ,  $W_b$  and  $W_1$  which must be known for the transistor in order to make the calculations shown in Table II. These quantities were obtained as follows:

- 1)  $W_b$  and  $W_1$  were calculated from values of alpha cutoff frequency measured at several values of reverse collector voltage.
- 2)  $V_B$  was obtained from the voltage-current characteristic of the collector diode measured with a low duty cycle sawtooth voltage waveform.

TABLE I  
PARAMETER VALUES USED FOR CALCULATIONS

Transistor	$W_b \times 10^3$ cm	$W_c \times 10^3$ cm/ $V^{1/2}$	$V_B$ volts	$n$	$C_c \mu\mu f$ at $V_c = 1V$	$(1-\alpha_0)$ at $V_c = 1V$ $I_c = I_p$	$V_0$ volts
2N416-1	1.62	0.10	-76	4.7	19	0.091	-66
2N416-2	1.48	0.085	-80	4.6	21	0.043	-56
GT762-1	0.80	0.060	-30	4.75	30	0.0091	-29
GT762-2	0.75	0.057	-31	4.1	35	0.011	-28

TABLE II  
CALCULATED AND MEASURED VALUES FOR CIRCUIT OF FIG. 1,  $C = 200 \mu\mu f$

Transistor Type	$-V_p$		$I_p$ ma		$\tau$ $\mu\mu sec$		$C_{min}$ $\mu\mu f$		$1/\lambda$ $\mu\mu sec$		$di/dt$ ma/ $\mu\mu sec$	
	Calculated	Measured	Calculated	Measured	Calculated	Measured	Calculated	Measured	Calculated	Measured	Calculated	Measured
2N416-1	37	36	230	250	50	70	80	93	8	8	5.5	8
2N416-2	34	40	41	47	—	—	400	420	24	26	0.5	0.63
GT762-1	10	8	370	330	7.8	10	58	63	1	—	13	20
GT762-2	9.2	8	320	310	5.7	8	105	100	1.4	—	12.5	22

3)  $n$  was obtained by making pulse measurements, at a current of 10 ma, of the transistor  $\alpha$  and the voltage  $V_{\alpha M}$  at which  $M\alpha$  was unity.  $(1-\alpha)$  was assumed to be proportional to  $W^2$ , and the equation

$$M\alpha = \frac{\alpha}{1 - \left(\frac{V_{\alpha M}}{V_B}\right)^n} = 1$$

was solved for  $n$ .

- 4)  $C_c$  was measured for several values of reverse collector voltage to check its dependence upon  $(-V)^{-1}$ .
- 5)  $\alpha_0$  was measured at a collector voltage of  $-1$  volt and collector current equal to the measured value of  $I_p$ . The voltage dependence of  $\alpha_0$  was found by assuming that  $(1-\alpha_0)$  was proportional to  $W^2$ . (The value of  $\alpha_0$  at  $V_c = -V_p$ ,  $I = I_p$  was used for  $\gamma$  in the calculation of  $C_{min}$ .  $V_p$  was calculated by solving the equation  $M(V) \alpha_0(V) = 1$  for  $V$ ).

Table II compares observed and calculated values for the circuit of Fig. 1, with  $C = 200 \mu\mu f$ . The deviations which occur are considered to be attributable primarily to experimental inaccuracies, except in the cases of rate of rise and fall time.

Concerning rate of rise, calculated value is intended only to be a lower bound on the linear portion of  $di/dt$ . Since the entire current rise is not linear, calculated value is expected to be less than the measured value.

Furthermore, it is to be expected that the measured value of fall time  $\tau$  will be greater than the calculated value, for the following reasons:

- 1) At the high currents which are obtained, the emitter efficiency  $\gamma$  is less than unity and stored charge is being lost through the emitter. Thus the current is decreased and  $\tau$  is increased.
- 2) As the stored charge is redistributed, the field in the base may be reduced. The result is a decrease in  $D_e$ , hence a reduction of  $\tau$ .
- 3) The charge  $Q_D$  required by the depletion layer capacity is not negligible. This reduction in stored charge has the same effect as a non-unity emitter efficiency.

VI. ACKNOWLEDGMENT

The authors are grateful to Dr. J. G. Linvill for a critical reading of the paper and to Mrs. J. Weil, who assisted in the preparation of the manuscript.

# Pulse Amplification Using Impact Ionization in Germanium\*

M. C. STEELE†, L. PENSAK†, SENIOR MEMBER, IRE, AND R. D. GOLD†, MEMBER, IRE

**Summary**—Some aspects of the phenomena of impact ionization in an impurity doped semiconductor at 4.2°K are described. The ionization time is shown to be a strong inverse function of the electric field. Control of the breakdown process is used to obtain pulse amplification in the millimicrosecond range, using two- and three-terminal devices. These devices are inherently stable, very simply constructed (requiring only ohmic contacts), and have easily manageable dimensions.

In controlling the breakdown process, a small amount of control power or energy applied at an *early* time determines the course of an avalanche-type buildup so that at a *later* time a substantial change in power or energy level is affected. Although other rise time control amplifiers are known (*e.g.*, the super-regenerative vacuum tube amplifier), the present amplifier is believed to be the first solid state device which operates on this principle.

Peak power and current gains on the order of 20 have been obtained for 25 and 50  $\mu$ sec pulses. It also appears that voltage gain is possible under somewhat different operating conditions than have been examined to date. The lower limit for pulse length in these experiments was determined by measurement equipment limitations rather than by any device limitation.

## INTRODUCTION

IMPURITY-TYPE semiconductors can attain very high resistivities at sufficiently low temperatures.

Such behavior is due to the decrease in the number of mobile carriers available. Most of these carriers become reattached to the impurity atoms when the thermal energy,  $kT$ , becomes considerably less than the impurity activation energy,  $\epsilon_i$ . In general, the remaining carriers may have very high mobilities at such low temperatures. Therefore, the carriers (either holes or electrons) can acquire kinetic energy rapidly from rather small electric fields applied to the crystal. When they have attained energies  $>\epsilon_i$ , there exists the possibility of an ionizing inelastic collision with a neutral impurity atom. This results in more carriers becoming available for conduction. In fact, the current rises very steeply when the applied electric field is increased beyond the critical value needed to initiate the ionization. This effect is called impact ionization of impurities in the semiconductor.<sup>1</sup> It is analogous to a gaseous discharge in many respects. A typical current-voltage characteristic is shown in Fig. 1. The voltage  $V_B$  is designated as the breakdown voltage.

In germanium, the critical field is about 10 v/cm at 4.2°K. Previous experiments<sup>2</sup> with heavily doped Ge

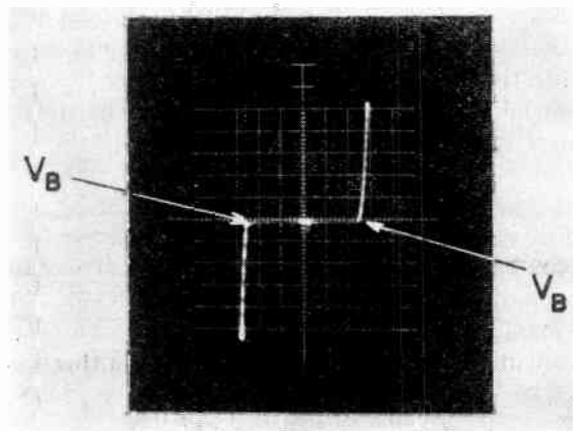


Fig. 1—Typical current-voltage characteristic;  $V_B$  is the breakdown voltage. Ordinate = 20 ma/cm; abscissa = 5 v/cm.

had revealed that the current rise time is in the  $10^{-9}$  sec range once the critical electric field is exceeded by about 50 per cent. The recombination process was also found to be in the  $10^{-9}$  sec range for the crystals examined. Two results in those experiments suggested the possibility of constructing a pulse amplifier based on impact ionization. They were:

- 1) The current rise time decreased rapidly as the applied electric field was increased above the critical value.
- 2) The current rise time decreased when the germanium was exposed to visible radiation.

The first of these results is readily understood in terms of the model<sup>1</sup> for explaining the impact ionization effect. The time needed for a carrier to gain the amount of energy required to produce ionization would certainly be an inverse function of the strength of the applied electric field. Precise measurements of the rise time as a function of electric field were not made in the previous<sup>2</sup> work. One of the objects of the present work was to study this effect in detail. In fact, the results of such measurements form the major basis of most of the material that follows in this report.

The second result is not easily understood in all its aspects. From the fact that light creates additional carriers which themselves can then contribute to the ionization process, it would follow that a given current level would be attained at an earlier time (all other factors being the same). However, experiments currently in progress indicate that the critical electric field, itself, decreases when germanium is exposed to visible light. Clearly such an effect would also be important in determining the rise time of the current.

\* Original manuscript received by the IRE, January 29, 1959; revised manuscript received, March 20, 1959.

† RCA Labs., Princeton, N. J.

<sup>1</sup> N. Sclar and E. Burstein, "Impact ionization of impurities in germanium," *J. Phys. Chem. Solids*, vol. 2, pp. 1-23; January, 1957.

<sup>2</sup> M. C. Steele and G. B. Herzog, unpublished report.

In light of these experiments, two types of pulse amplifiers were considered. The first is based on modulating the electric field over part of the germanium by means of a third electrode placed between the input and output terminals. All the contacts to the germanium would be *ohmic* in such a device. Another way of modulating the rise time in a two-terminal device is to alter the input pulse shape by means of a control signal. Such a diode can also give amplification under proper operating conditions. Both the two- and three-terminal devices were built and studied in detail. The results are discussed in this paper. It should be noted here that in the three-terminal device there is no isolation between the control signal source and the drive source. In this sense, the device differs from conventional amplifiers. The second type of pulse amplifier depends on being able to introduce light pulses on the germanium, thereby modulating its conductivity over a wide range. There are a number of interesting possibilities in such devices. Some of the possible uses for these and other impact ionization devices will be discussed in the conclusion of this report.

#### EXPERIMENTAL DETAILS

All the crystals used in these experiments were *n*-type germanium. The room temperature resistivities varied from 0.1 to 1 ohm-cm. A typical wafer, using ohmic line contacts, is shown in Fig. 2. The side contact was located either midway between the two end contacts, or near the end from which the output was taken.

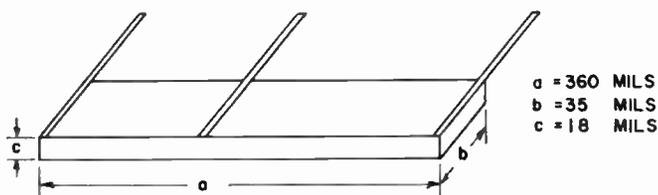


Fig. 2—Typical wafer dimensions.

Measurements were made with the wafer immersed in liquid helium at 4.2°K. Coaxial wires, supported within an inconel tube, were connected to the wafer as shown in Fig. 3. Connection was made to the side contact through a 7K resistor. Time delay effects in comparing signals at the three wafer contacts were eliminated by using "output" cables of equal length at the drive, control, and load terminals of the wafer, in addition to the cables for applying the drive and control signals (Fig. 3). The "output" cables were terminated in their characteristic impedance.

Amplification and current rise time measurements were made using pulses ranging from 10 to 100  $\mu\text{s}$  in length and from 12 mv to 200 v in amplitude. The rise and fall times of the applied pulses were about 1  $\mu\text{s}$ . Voltage measurements were made on an oscilloscope

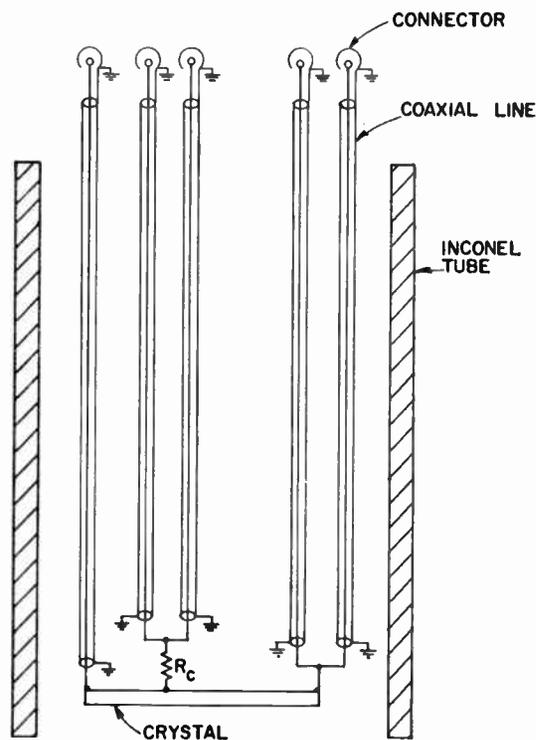


Fig. 3—Structure used to insert crystal in the liquid helium dewar, showing the coaxial line connections to the crystal.

having a 7- $\mu\text{s}$  rise time to square pulses and a maximum sensitivity of 0.1 v/cm.

The current-voltage characteristics at high current levels were obtained with the use of another pulse generator. This unit supplied a 40- $\mu\text{s}$  pulse at a pulse repetition frequency of 130 cps. The pulse amplitude was varied at a 60 cps rate. As viewed on an oscilloscope, the current-voltage characteristic appeared as a series of moving points which photograph as a smooth curve. A typical characteristic obtained in this way is shown in Fig. 1. For more accurate voltage-current measurements at low direct current levels (down to  $10^{-10}$  amperes), an electrometer type meter was used.

#### DISCUSSION OF OPERATION

##### Current Rise Time

The entire basis of the impact ionization amplifiers to be described below is that the current rise time can be strongly influenced by the magnitude of the applied electric field. Previous work<sup>2</sup> had only established qualitatively that the current rise time decreased as the electric field increased. In light of the fundamental importance of this functional relation in all the devices, it was studied in detail, experimentally. Before discussing the results of such experiments, a simple theoretical development of what might be expected will be given. The kinetic energy gained in a time  $t$  by a carrier moving in an electric field is:

$$\epsilon = e\mu E^2 t, \quad (1)$$

where  $\epsilon$ =energy,  $e$ =electronic charge,  $\mu$ =carrier mobility, and  $E$ =electric field. When  $\epsilon \geq \epsilon_i$ , the impurity activation energy, inelastic ionizing collisions can take place to produce more mobile carriers. It follows from (1) that  $\tau_0$ , the minimum time needed to produce carrier multiplication, is

$$\tau_0 = \frac{\epsilon_i}{e\mu E^2} \tag{2}$$

The simple derivation of (2) does not take into account any changes in mobility as the carrier gains energy. In actual crystals, the mobility will change with energy in a way determined by the dominant type of scattering. Notwithstanding this simplification, (2) gives a functional relation that is germane to the device operation. It points out the fundamental nature of the electric field effect on the current rise time.

Experiments to verify the general nature of (2) were performed at 4.2°K with an *n*-type germanium wafer having a resistivity of 0.6 ohm-cm at 300°K. Pulses of 100  $\mu$ s width and varying amplitudes were applied to the crystal. The output across a 75-ohm series resistor was applied directly to the plates of the oscilloscope. Under these conditions, the oscilloscope rise time was about 2  $\mu$ s and the vertical sensitivity was 11 v/cm. The time,  $\tau$ , required for the output current to reach about 15 ma (well below saturation) was noted for each voltage. Results of such measurements are given in Fig. 4, where  $\tau$  is plotted as a function of voltage on a log-log

basis. For low electric fields, it is found that  $\tau \propto 1/E^{3.2}$ , while at higher fields  $\tau \propto 1/E^{1.2}$ . It follows therefore that (2) is obeyed qualitatively. The deviation of the experimental results from (2) are most likely associated with, a) the simplification made in the mobility, and b) the variation of the cross section for an ionizing collision with the energy of the carrier. Since the purpose of the present work is to utilize the  $\tau$  vs  $E$  dependence in understanding device operation, a detailed comparison of Fig. 4 to theory was not pursued. It properly forms the basis of further fundamental studies in the phenomenon of impact ionization.

*Signal Amplification*

*Diode Operation:* The results shown in Fig. 4 suggest that properly chosen means for influencing the amplitude or phase of a voltage pulse by a control signal might result in amplification. From the viewpoint of simplicity of analysis the first case considered is a diode structure with the crystal itself as the load. Referring to Fig. 5, a drive voltage pulse of amplitude  $V_D$  and duration ( $t_D - t_c$ ) is applied to the crystal in the absence of any control signal. The control signal is a voltage pulse of amplitude  $V_c$  and precedes the start of the drive pulse by a time  $t_c$ . The duration of the control signal pulse is also  $t_c$ . The amplitudes of both  $V_c$  and  $V_D$  must exceed the breakdown voltage,  $V_B$ . During the time interval 0 to  $t_c$  it is assumed that the current is rising exponentially<sup>3</sup> with time according to the equation

$$I = I_0 \exp(t/\tau_2), \tag{3a}$$

where  $I_0$  is the initial current just before the onset of breakdown and  $\tau_2$  is the time constant characteristic of the voltage  $V_c$ . With no control signal present the current rise in the interval  $t_c$  to  $t_D$  is

$$I = I_0 \exp\left(\frac{t - t_c}{\tau_1}\right), \tag{3b}$$

where  $\tau_1$  is associated with  $V_D$ . When the control signal is present, the current rise in the interval  $t_c$  to  $t_D$  is

$$I = I_0 \exp\left(\frac{t_c}{\tau_2}\right) \exp\left(\frac{t - t_c}{\tau_1}\right). \tag{3c}$$

It then follows (see Appendix) that the energy gain produced by the control signal is

$$K = \frac{\Delta\epsilon}{\epsilon_{in}} = 1 + \left(\frac{V_D}{V}\right) \left(\frac{\tau_1}{\tau_2}\right) \left[ \exp\left(\frac{t_D - t_c}{\tau_1}\right) - 1 \right]. \tag{4}$$

In (4),  $\epsilon_{in}$  is the input energy of the control source, and  $\Delta\epsilon$  is the net change in the output energy during the time interval 0 to  $t_D$ . It is seen that for a fixed  $t_D$ , the energy gain is maximized for  $t_c = 0$ . When  $t_c = t_D$  the gain is unity.

<sup>3</sup> A good approximation at low currents. Exponential behavior is expected in the absence of strong three-body interactions.

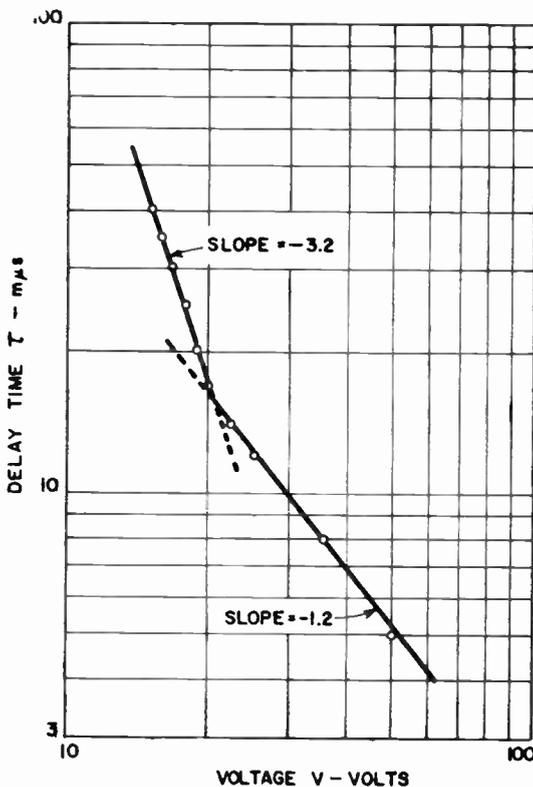


Fig. 4—Current delay time vs applied voltage, sample IIA-2.

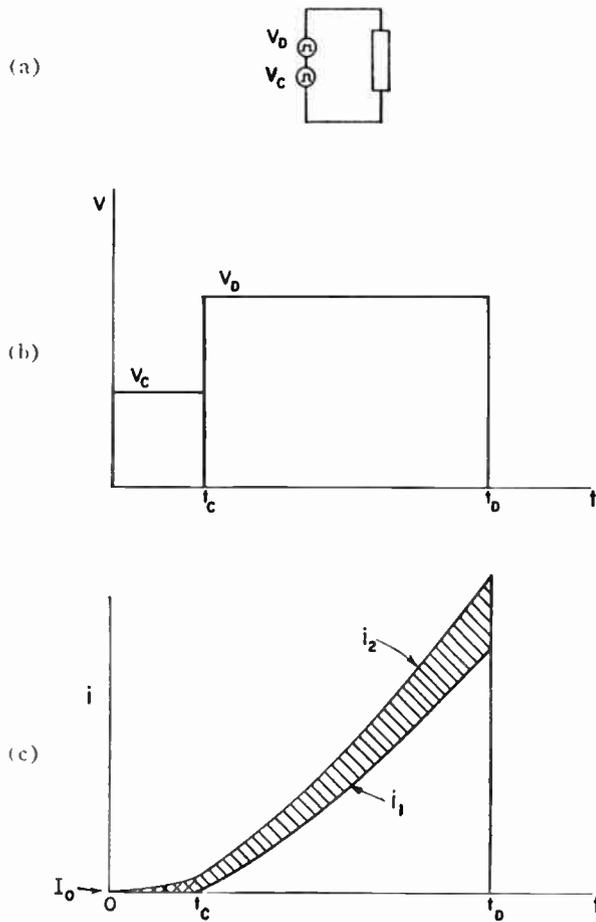


Fig. 5—Pulse time modulation method for energy gain in a diode. (a) Circuit. (b) Applied voltage vs time. (c) Diode current vs time;  $i_1$  is the current when no control voltage is applied;  $i_2$  is the current when control voltage is present.

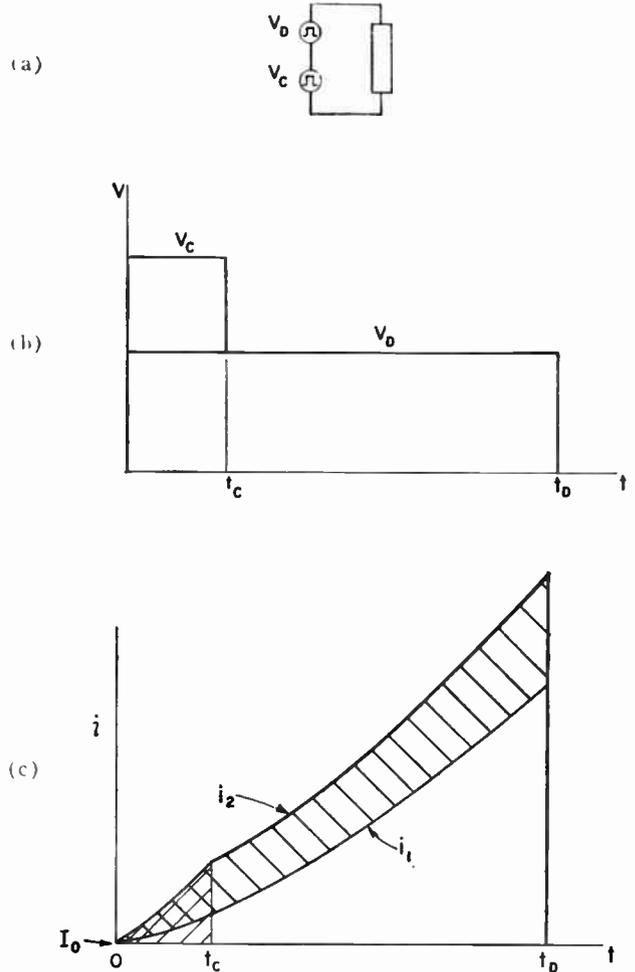


Fig. 6—Pulse amplitude modulation method for energy gain in a diode. (a) Circuit. (b) Applied voltage vs time. (c) Diode current vs time.  $i_1, i_2$  are as defined in Fig. 5.

Energy gain can also be obtained by superposing a short control signal of amplitude  $V_c - V_D$  for a time interval of  $t_c$  on a longer drive voltage of duration  $t_D$ . Such a situation is depicted in Fig. 6. In this case the energy gain is

$$K = 1 + \frac{\left(\frac{V_D}{V_c}\right)\left(\frac{\tau_1}{\tau_2}\right)\left[\exp\left(\frac{t_c}{\tau_2}\right)\left\{\exp\left(\frac{t_D - t_c}{\tau_1}\right) - 1\right\} - \left\{\exp\left(\frac{t_D}{\tau_1}\right) - 1\right\}\right]}{\left\{\exp\left(t_c/\tau_2\right) - 1\right\}} \quad (5)$$

If an external resistive load is put in series with the impact ionization diode shown in either Fig. 5 or Fig. 6, the analysis becomes complicated. The resulting expression for the current is a transcendental function whose argument is itself a function of the current. No attempt was made to seek a solution for this case.

**Triode Operation:** It is clear that any means of influencing the electric field in the impact ionization device might result in amplification. A triode configuration with a side contact (control) placed between the input (drive) and output (load) ends is shown in Fig. 7. Here, the drive amplitude is sufficiently large to produce an output across the load in the absence of a control volt-

age. The potential values along the device are depicted in Fig. 8 as the current in the unit builds up from zero to the final value in the load resistor. Fig. 8(a) shows the potential distribution at the instant the positive drive

pulse is applied. Since the crystal resistance is then very large compared to the control circuit resistance, the contact  $C$  is essentially at ground potential. As soon as some current (carrier multiplication) begins to flow from  $D$  through  $C$  to ground, the potential at  $C$  begins to increase with respect to point  $L$  (which is still at ground potential). The difference in potential between  $C$  and  $L$  finally reaches (or exceeds) the breakdown voltage for that section of the crystal. That situation is shown in Fig. 8(b). Now current flows through the external load resistor causing the potential at  $L$  to increase. This is shown in Fig. 8(c) at the instant the drive pulse is terminated.

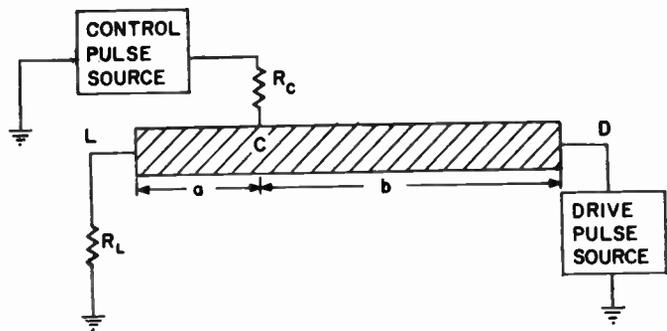


Fig. 7—Basic circuit of triode amplifier.

By introducing a control signal of opposite polarity to the drive, the breakdown in section *D-C* is accelerated since the electric field between these two points has been increased. This is depicted by the dotted line in Fig. 8(a). Because of this action, the section *C-L* also breaks down earlier causing a higher average load current to flow during the drive pulse. The control pulse is terminated before the end of the drive pulse. Details of the duration of the control pulse will be given later. If the drive pulse were on long enough to produce saturation of the load current, the difference in load current with and without the signal would be as shown in Fig. 9. No change would be made in the saturation current, but the average current would be greater when the control pulse was present. It follows similarly that if the drive and control pulses were of the same polarity, the average current would be smaller with the control present.

This qualitative description of the triode operation shows clearly that any possible amplification must occur during the *time of rise* of the current. In this sense, the device is a "rise-time controlled amplifier." The shaded area in Fig. 9 represents the integrated (current)  $\times$  (time) increase resulting from applying the control signal. For greater efficiency of operation, the drive pulse duration should be less than that required to produce saturation of the load current. Under these conditions, the peak power gain,  $K_p$ , can be defined as

$$K_p = \frac{\Delta P_0}{P_c} \tag{6}$$

where  $\Delta P_0$  is the change in peak power output resulting from applying the control signal of peak power  $P_c$ . For this circuit (Fig. 7)

$$\Delta P_0 = \frac{1}{R_L} [(V_{L2})^2 - (V_{L1})^2], \tag{7}$$

where  $V_{L2}$  and  $V_{L1}$  are the peak load voltages with and without a control signal, and  $R_L$  is the load resistance. Within the framework of the limits to be set forth below the instantaneous value of  $P_s$ , the input power, will be defined as

$$P_c = V_c I_c, \tag{8}$$

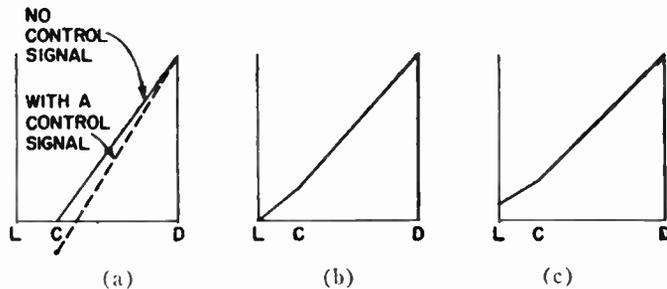


Fig. 8—Potential diagram of triode amplifier. (a) At instant when voltage is applied to the crystal. (b) When side contact reaches the breakdown voltage for region *C-L*. (c) When drive pulse is terminated.

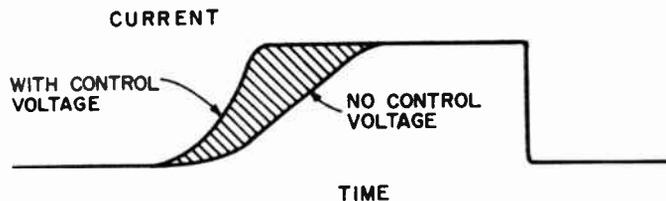


Fig. 9—Effect of control voltage on load current.

where  $V_c$  is the control voltage and  $I_c$  is the current in the control source. Clearly  $I_c$  will vary with time during the presence of the control signal. It must start from zero and increase monotonically (perhaps approaching a saturation value). Hence, in general, the input power is

$$P_c = \frac{V_c(V_x + V_c)}{R_c}, \tag{9}$$

where  $R_c$  is the resistance in series with the control circuit and  $V_x$  is the potential at point *C* of the crystal. In (9) it is assumed that the control voltage is  $-V_c$ . With the control and drive present, and before any current flows in the device,  $V_x = -V_c$ . The maximum value that  $V_x$  can attain is dictated by the position of the side contact relative to the drive and load ends. Referring to the dimensions shown on Fig. 7, that value is

$$(V_x)_{\max} = \left( \frac{a}{a+b} \right) (V_D - V_{L2}) + V_{L2}. \tag{10}$$

Note that the control pulse could be turned off before  $V_x$  reaches the value of (10) without affecting the output current. In fact, the control pulse does not contribute to the amplification after  $V_x$  exceeds the value of zero. From this criterion, the duration of the control pulse can be ascertained experimentally from the observed current rise time characteristics as a function of electric field. It follows that for optimum operation, the control pulse should be of shorter duration than the drive pulse. The maximum useful *peak* value of  $P_c$  is then given as

$$P_c = \frac{V_c^2}{R_c}. \tag{11}$$

Eq. (11) does not necessarily give the value of  $P_s$  needed to maximize the power gain. The calculation of that

value is complicated by the details of how the output varies with time for a given drive pulse.

In actual use, the energy gain per operation is of importance. This would be obtained by integrating (7) and (9) over the appropriate time intervals of drive and control duration respectively.

The power gain definition given by (7) and (8) is a very general one in that the net power output is considered to be the *change in output power* due to the input signal. It is important to distinguish that definition from the one conventionally used in the analysis of devices.<sup>4</sup> In the more conventional definition, the useful output power is defined in terms of the *change in output voltage* (or current) due to the input signal. By this criterion, the power gain would be defined by

$$K_p' = \frac{\frac{1}{R_L} (V_{L2} - V_{L1})^2}{V_{L1} I_c} \quad (12)$$

Eq. (12) implies that information is carried by the voltage in excess of a residual or carrier voltage rather than by the power in excess of a residual power. It is seen, however, that the values of  $K_p$  and  $K_p'$  become identical when  $V_{L1}=0$ . For all other values of  $V_{L1}$ ,  $K_p > K_p'$ .

There are several ways to make  $V_{L1}$  negligibly small compared with  $V_{L2}$  in the impact ionization amplifier. The first one would be to make the drive pulse duration short enough to insure that  $V_{L1} \ll V_{L2}$ . Such a direction is also the most desirable for fastest possible operation. The second way would be to have a dc bias present at the control terminal of such a polarity as to cut off the pulse output when no control signal is applied. This would be analogous to biasing the grid of a tube to cut-off. In the present situation, the bias potential would be the *same sign* as the drive pulse potential, while the control pulse would be of opposite polarity.

Although the power gain definition given by (12) is the conventional one, the use of (7) and (8) to define power gain presents a more fundamental description of the impact ionization amplifier.

## RESULTS

### Direct Current Control Signals

The general ideas of triode operation were first tested with direct current control signals and pulsed drive. Fig. 7 shows the circuit used for this purpose except that the control pulse source was replaced with a dc source. A 25- $\mu\text{s}$  positive pulse of fixed voltage was applied to the crystal. The output current into the 75- $\Omega$  load was measured. Then dc was passed from point *C* through the 75  $\Omega$  load and the effect on the output pulse was noted. Fig. 10 shows a sketch of the results. Trace *a* is the pulse current when no dc was applied. Trace *b* is the pulse current when +0.2  $\mu\text{a}$  of dc flowed from point *C* through the output load to ground. Trace *c* is the pulse current

for -0.2  $\mu\text{a}$  of dc. It is seen that a very small amount of dc was able to affect a large change in the output pulse current. If a current gain,  $K_I$ , is defined as

$$K_I = \frac{\Delta I_L}{I_{dc}}, \quad (13)$$

where  $\Delta I_L$  is the net change in the *pulse* output current and  $I_{dc}$  is the dc current, it is found that  $K_I \approx 10^4$  for the results shown in Fig. 10.

The observed changes in output pulse current are brought about by changing the electric field between points *D* and *C* when the dc is applied. In fact, these observations conform with the explanation offered above for the triode operation. The polarity of the applied voltage on point *C* can cause an increase or decrease of pulse output since the rate of rise of the current is a function of the electric field. The large values of current gain merely emphasize that the control is *field* dependent rather than current dependent. In this respect an arrangement, such as shown in Fig. 10, may be very useful for measuring small currents in high impedance circuits.

### Pulsed Control Signals

If the dc source is replaced by a signal pulse source the operation allows one pulse to control the other (see Fig. 7). The mechanism would then be exactly as depicted in Fig. 8 and Fig. 9 for the triode amplifier. For a positive drive pulse, a positive control pulse should decrease the output current, while a negative control pulse should increase the output. These anticipated results were verified experimentally. A typical result for a negative drive and a positive control pulse is shown in Fig. 11. Trace *a* is the output current pulse for a *negative* drive pulse with *no control* pulse applied. Trace *b* is the output current pulse for a *positive* control pulse with no drive pulse applied. Trace *c* is the output current pulse when the same applied pulses for traces *a* and *b* are *simultaneously* applied to their respective terminals. From Fig. 11 it is seen that a current input of 0.67 ma has produced a change in output current of 5.3 ma.

### Triode Power Gain with Pulsed Control Signals

Power gain into a load resistor was demonstrated using the circuit shown in Fig. 12. The transformer in the control circuit inverts the polarity of the control pulse relative to the drive pulse. For these tests, both the drive and control pulses were 50  $\mu\text{s}$  long. With the definition of  $P_c$  given in (9) the results are shown in Fig. 13. For these data,  $V_x$  was taken as the value given by (10). Actually  $V_x$  could have been smaller than the values calculated in this way, but the resulting maximum differences in  $P_c$  would only be about 20 per cent less than the values shown in Fig. 13. It is seen that peak power gain was exhibited. In particular, the power gain becomes greater as the drive pulse amplitude is increased and the control pulse is decreased. Such a result

<sup>4</sup> The authors wish to thank T. O. Stanley for bringing this matter to their attention.



*Diode Power Gain with Pulsed Control Signal*

The circuit shown in Fig. 6 was altered to include a series load resistor (75 ohms). With this circuit energy gain into an external load was demonstrated. The actual input and output pulse forms are shown in Fig. 15. From these data, the input and output energies were computed. The resulting input energy was  $1.1 \times 10^{-10}$  watt-sec, while the net output energy was  $2.4 \times 10^{-9}$  watt-sec; this corresponds to an energy gain,  $K_E$ , of 22. Such a result confirms, qualitatively, the analysis leading up to (4) and (5), and points to the potentially important use of two-terminal impact ionization devices.

DISCUSSION

*Use of Amplifier*

In general, the experimental results confirm the models described for the operation of the two terminal and three terminal impact ionization amplifiers. There are specific comments that can now be made about the use of the three-terminal devices. To begin with, it can be used as a fast current amplifier. This suggests the possibility of counting nuclear events very rapidly. Such use would be most efficiently realized if the low currents were associated with a high impedance source.

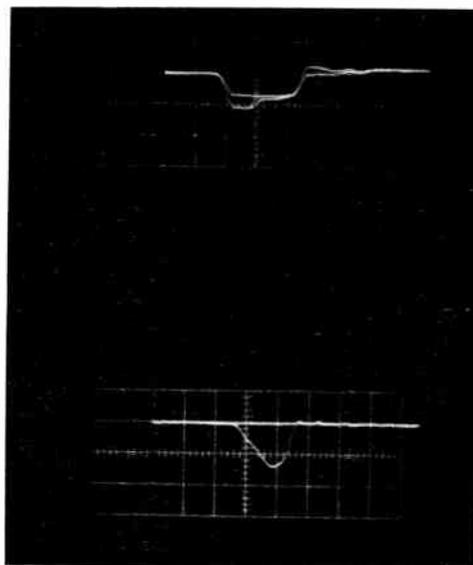
Because current gain is so readily attained, the device could be used to drive low impedance current sensitive loads. Superconducting memory elements fall into this class. The concept of controlling a transient condition forms the basis of the work described in this report. Such a concept may be extended to other transient phenomena that involve sharp nonlinearities in a current-voltage relation.

*Voltage Gain*

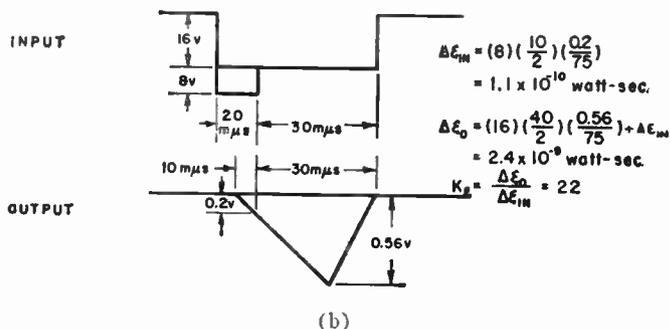
The possibility of driving one impact ionization amplifier with another leads to the question of voltage gain. For most of the experiments performed thus far, the measured voltage gain has ranged from 0.01 to 0.1, with the larger value corresponding to a higher drive voltage. The voltage gain was relatively insensitive to changes in control voltage. Since impact ionization devices are field sensitive rather than current sensitive, voltage amplification is necessary for one unit to drive another. Several possibilities present themselves for controlling one unit from the output voltage of another unit.

The most obvious solution to the multistage drive problem (in the absence of voltage gain) is to use pulse coupling transformers to step up the output voltage. However, since the transformer must be located at the germanium crystal (to avoid loading effects of the connecting transmission lines), this raises the problem of finding a core material that would be suitable at liquid helium temperatures. At present, no such material is available.

A second method is to connect a large resistor directly to the output terminal of the crystal, and couple this point to the control terminal of the next crystal. This avoids the low impedance limitation imposed on the



(a)



(b)

Fig. 15—Energy gain in a diode amplifier. (a) Typical input and output waveforms. Upper photo shows input voltages; ordinate = 20 v/cm, abscissa = 20 mμs/cm. Lower photo shows output voltage across 75 Ω load; ordinate = 0.4 v/cm, abscissa = 20 mμs/cm. (b) Simplified representation of input and output waveforms, and calculation of energy gain.

previous voltage gain measurements by the transmission line characteristic impedance (75 ohms). Some preliminary measurements have been made using this technique. In one test, a voltage gain on the order of 1.8 was obtained, but equipment limitations resulted in an uncertainty of about 50 per cent. Further, this limitation also prevented testing under the most optimum conditions. In a second preliminary test, one unit was made to control the output of another. There appears to be no inherent reason why voltage gain should not be attained. One approach for further work is to use a large load resistor at the output terminal of the crystal, and to use higher drive voltages (which would also result in faster operation).

*Comparison with Super-Regenerative Amplification*

The impact ionization amplifier is characterized as a rise time controlled device. This bears some resemblance to the operation of the super-regenerative amplifier familiar to radio engineers. In this type of amplifier, an oscillator is periodically turned off by means of a "quench" voltage. The time required for oscillations to

build up when the quench voltage is removed depends on whether the oscillations are initiated by the inherent tube noise or by an externally applied signal. The oscillations reach saturation earlier when a signal is applied, thus changing the average plate current amplitude. The duration of the ON time (no quench voltage) determines whether the oscillations will reach saturation (logarithmic mode) or be turned OFF by the quench voltage before saturation (linear mode).

The mode of operation of the impact ionization amplifier described here may be better understood by comparison with the linear mode super-regenerative amplifier. However, there is an important difference in the rate of rise; *i.e.*, the time constant in the exponential buildup expression. For the super-regenerative receiver, this is a constant determined by the lumped constant parameters of the equivalent RLC circuit, and is independent of the signal voltage. For the impact ionization amplifier, on the other hand, the rate of rise is a strong inverse function of signal voltage, as shown by (2) and Fig. 4. Despite this, it is felt that this comparison serves to relate the operation of a new device to a familiar one.

#### APPENDIX

##### Diode Analysis

Referring to Fig. 5, the input signal energy is

$$\epsilon_{in} = V_c \int_0^{t_c} I dt, \quad (14)$$

where  $I$  is given by (3a) in the text. Carrying out the integration gives

$$\epsilon_{in} = V_c I_0 \tau_2 \left[ \exp\left(\frac{t_c}{\tau_2}\right) - 1 \right]. \quad (15)$$

With no control signal present, the output energy is

$$\epsilon_0 = V_D \int_{t_c}^{t_D} I dt, \quad (16)$$

where  $I$  is now given by (3b). This integration yields

$$\epsilon_0 = V_D I_0 \tau_1 \left[ \exp\left(\frac{t_D - t_c}{\tau_1}\right) - 1 \right]. \quad (17)$$

With control signal present, the output energy is

$$\epsilon_1 = \epsilon_{in} + V_D \int_{t_c}^{t_D} I dt. \quad (18)$$

where  $I$  is now given by (3c). The integration of (18) gives

$$\epsilon_1 = \epsilon_{in} + V_D I_0 \tau_1 \left[ \exp\left(\frac{t_D - t_c}{\tau_1}\right) - 1 \right] \cdot \left[ \exp\left(\frac{t_c}{\tau_2}\right) \right]. \quad (19)$$

From (14), (16), and (18) it follows that

$$K = \frac{\Delta \epsilon}{\epsilon_{in}} = \frac{\epsilon_1 - \epsilon_0}{\epsilon_{in}} = 1 + \left(\frac{V_D}{V_c}\right) \left(\frac{\tau_1}{\tau_2}\right) \left[ \exp\left(\frac{t_D - t_c}{\tau_1}\right) - 1 \right]. \quad (20)$$

Eq. (20) is the value of  $K$  given by (4) in the text.

#### ACKNOWLEDGMENT

The authors wish to thank G. B. Herzog and E. O. Johnson for profitable suggestions throughout the course of this work.

## The Transpolarizer: An Electrostatically Controlled Circuit Impedance with Stored Setting\*

C. F. PULVARI†, SENIOR MEMBER, IRE

**Summary**—This new device operates by the controlled transfer of polarization through two or more ferroelectric dielectric sections in series and therefore is named "transpolarizer."

\* Original manuscript received by the IRE, November 24, 1958; revised manuscript received, February 16, 1959. This work was supported by the U. S. Air Force, Office of Air Research, Contract No. AF 33(616)-2934.

† Dept. of Elec. Engrg., The Catholic University of America, Washington, D. C.

It represents a new basic means for storing and gating electrical signals and, in general, means to control circuit impedance in any predetermined manner according to a stored setting.

The operation of a two section transpolarizer is described.

The unique storage, switching, and control properties of the transpolarizer open a large field of new applications and permit production of new devices and systems such as: recording and reproducing of intelligence in general and, more particularly, switching with a permanent setting, small and large scale storage devices with nondestructive readout, decoders, function generators, etc.

THE transpolarizer,<sup>1,2</sup> analogous to its magnetic counterpart, the transfuxor,<sup>3</sup> consists of at least two capacitors with a preferable crystalline ferroelectric dielectric and a nearly rectangular hysteresis loop. Control of polarization transfer through two or more ferroelectric dielectric sections in series represents a new basic means for storing and gating electrical signals and, in general, a means for controlling circuit impedance in any predetermined manner according to a stored setting.

Non-linear "impedance" of a circuit at a given frequency can be set to any predetermined level by a single pulse, and it will remain for a long period of time. Therefore transmission of an ac power can be controlled or set to a desired level by a single setting pulse. This level may assume any value between zero and the maximum limits of the device, and an on-off or continuous stored control operation is possible.

Although various multisection transpolarizer arrangements have been operated and tested only the basic two section transpolarizer is described below. This basic circuit element shows good efficiency of power transmission, short setting time, negligible coupling between setting and output circuits and sharp setting thresholds.

#### INTRODUCTION

Recent material developments in ferroelectrics<sup>4,5</sup> with nearly rectangular hysteresis loops result in new storing and switching devices. Ferroelectric amplifier-choppers, gates, logical devices, bistable and astable multivibrators shift registers and memory matrices became practical. The use of ferroelectrics until now was limited because of fatigue effects and lack of a sharp threshold field.<sup>6</sup> These problems have been resolved. The new materials such as tri-glycine sulfate (TGS) and tri-glycine fluoberyllate (TGFB) do not have fatigue, and the threshold problems have been solved by a combination of semiconductors and ferroelectrics.<sup>6</sup> The basic element in all of these devices is a condenser with ferroelectric dielectric. When two or more ferroelectric capacitors form a common section or sections and polarization of the condensers can be individually controlled, interesting new switching and storing functions can be achieved.

The controlled transfer of polarization through ferro-

electric sections of dielectrics suggested its name, "Transpolarizer." The transpolarizer can be regarded as a capacitor with a dielectric of variable dielectric constant to permit any setting of impedance in a continuous way between the blocked and unblocked states with a single setting pulse or bias, depending on circuit requirements. Any state once set will maintain its nonlinear impedance until another state is set. The transpolarizer is a voltage device with a short setting time, and when energized by an ac voltage, the current through it will correspond to its set impedance level. When operated in its blocked or unblocked states, it represents a storage element with "non-destructive" readout which permits construction of storage devices with permanent memory.

The unique high-impedance input feature of the transpolarizer renders it readily adaptable for control with an electron beam or photoelectric cell, and, as a result, such new devices as large display screens,<sup>7</sup> new visual pickup devices, combined ferroelectric and magnetic devices, etc., become feasible.

This paper presents the basic operation of the transpolarizer and will discuss its features as a circuit element.

#### BASIC OPERATION

If two ferroelectric elements are connected in series and if the polarization vectors of both elements are identical, as shown in Fig. 1(a), the two series-connected elements exhibit electrical behavior identical to a single ferroelectric element.

If both elements have identical properties, a low-frequency hysteresis loop will be observed which has a coercive voltage of  $2V_c$  and a total switched charge of  $Q_s$ .

If the polarization vectors of the two elements are in opposite directions, as shown in Fig. 1(b), the two-terminal device has strikingly different electrical properties from that of a single ferroelectric element. When a positive voltage—say a step function of finite rise time—is applied between the two external terminals, no net switching of charge results and the device resembles a normal linear dielectric capacitor. This can be explained with the aid of the low-frequency hysteresis loops of Fig. 1(b). As the applied voltage rises, it is at first equally divided between both elements. At a certain value of applied voltage, the upper element starts to switch, *i.e.*, the instantaneous value of charge starts to move around the lower knee of the hysteresis loop. At this point, the dielectric constant increases and therefore the capacitance of the upper element takes on a large value. Consequently, the applied voltage is no longer equally distributed across both elements but is divided in inverse proportion to its capacitances.

As applied voltage increases, voltage across the lower

<sup>1</sup> C. F. Pulvari, USAF Prog. Rep. No. 11, Contract No. AF 33(616)-2934; January, 1957.

<sup>2</sup> J. R. Anderson, "Ferroelectric Storage Device," U. S. Patent No. 2,695,396; November 23, 1954.

<sup>3</sup> J. A. Rajchman and A. W. Lo, "The transfuxor—a magnetic gate with stored variable setting," *RCA Rev.*, vol. 16, pp. 303-311; June, 1955.

<sup>4</sup> R. Pepinsky, Y. Okayo, and F. Jona, "Ferroelectricity and structure of tri-glycine fluoberyllate and its isomorphs," *Bull. Am. Phys. Soc.*, ser. 11, p. 220; 1957.

<sup>5</sup> C. R. Pulvari and W. Kuebler, "Polarization reversal in tri-glycine fluoberyllate and tri-glycine sulfate single crystals," *J. Appl. Phys.*, vol. 29, pp. 1742-1746; December, 1958.

<sup>6</sup> C. F. Pulvari, USAF Tech. Rep. No. 58-657, Contract No. AF 33(616)-2934.

<sup>7</sup> E. A. Sack, "A new electroluminescent display," 1958 IRE NATIONAL CONVENTION RECORD, pt. 3, pp. 31-39.

element increases and the instantaneous value of charge moves in the saturation region of the hysteresis loop. For this reason, the upper element cannot be switched and the voltage on it remains constant at a value slightly less than its coercive voltage. Thus, no switching results. If polarity of applied voltage is reversed, the role of upper and lower element interchanges. Since no switching takes place, a straight line, which is characteristic of linear capacitors, will be observed in place of the normal hysteresis loop.

If no switching occurs, the two elements are said to be in a "blocked" condition; if switching can take place, the two elements are said to be "unblocked."

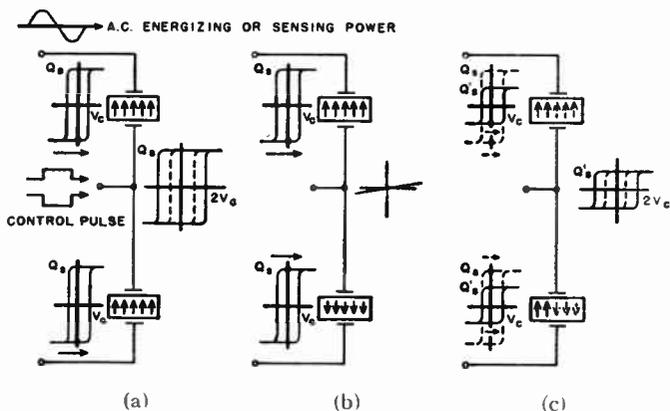


Fig. 1—Various states of a transpolarized: (a) unblocked state; (b) blocked state; (c) partially blocked state.

Fig. 1(c) shows a partially blocked state. This is obtained when the polarization of one of the capacitors is partially switched from state *a* into state *b*, thereby leaving part of the polarization vectors in both elements identical. As a result, part of the polarization behaves blocked (as indicated by the dotted arrows) and only that part of the polarization contributes to the loop which is unblocked and can be freely switched. Coercive voltage will be again  $2V_c$  but the switched charge  $Q_s'$  is less than  $Q_s$  in the completely unblocked state. It is clear that the partially blocked state *c* represents a lower capacitance than the unblocked state *a*, and that any level of partial blocking can be obtained depending on the control signal applied as indicated in Fig. 1 which can be either a single pulse or a dc bias depending on whether a static or dynamic operation is required.

The most interesting feature of the transpolarizer is that it can be changed from a ferroelectric capacitor into a linear-capacitor and can assume any intermediate polarization level between these two limits; furthermore, it is capable of controlling a flow of ac electric power according to its setting.

These novel switching and storage features spurred interest to determine the behavior of the transpolarizer in its blocked, partially blocked, and unblocked states as well as its features as a circuit element.

THE BLOCKED STATE

In order to investigate this state, the circuit shown in Fig. 2 was used. Again a two section transpolarizer comprising two identical ferroelectric storage cells connected in series was used. Actually, a twin storage cell was evaporated on a small single crystal of TGFB. The energizing ac or sensing signal will or will not produce a signal output in form of ac or transient as shown in Fig. 2 depending upon the setting pulse which may be a block or unblock pulse. The setting pulses are supplied through double anode saturation diodes as shown in Fig. 2 in order to isolate the setting pulse source from the circuit itself.

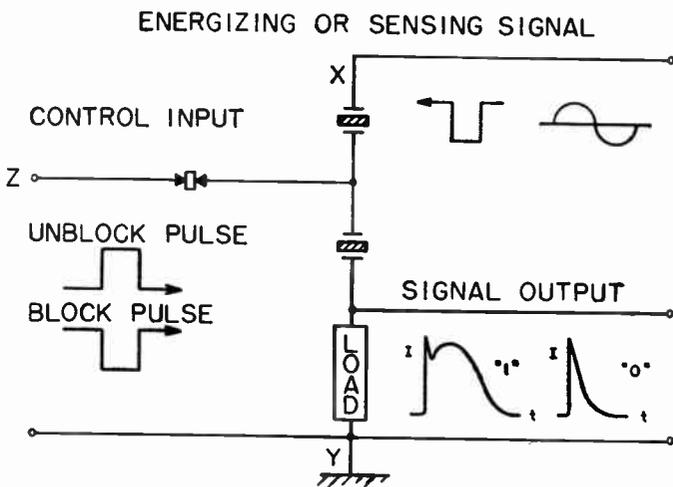


Fig. 2—Basic circuit of a transpolarizer.

Assuming a blocked state it was important to see how stable this state is when a) the energizing ac or sensing signal voltage increases, and b) a certain energizing voltage acts for a longer period of time.

Fig. 3 presents the variation of polarization,  $\Delta P/P_s$ , as a function of applied energizing voltage for various degrees of blocking.  $P_s$  is the saturation polarization. It is shown that a transpolarizer with a coercive voltage of  $V_c = 17$  v does not change the blocked states, *i.e.*,  $\Delta P$  is practically zero up to about 10 times its coercive voltage (170 v) and that to introduce a slight change starts at 200 v which increases to a value of  $\Delta P/P_s = 1.2$  per cent at 400 v. When the energization potential is reduced to zero, a remanent change of 0.2 per cent was observed. This process could be repeated many times, and with each repetition, the remanent change increased by an additional 0.2 per cent. It is noted, however, that when the energizing voltage was only 170 volts and this was reduced, no change could be measured.

Fig. 4 shows  $\Delta P/P_s$  as a function of time, for energizing voltages of 50, 150 and 400. It is again apparent that about one and one-half hours were needed for ef-

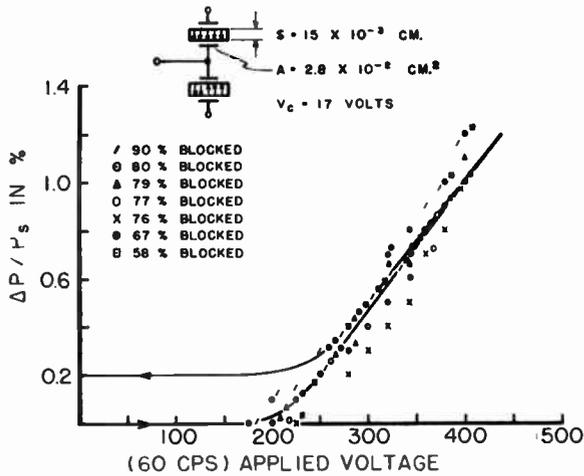


Fig. 3—The effect of energizing voltage on various blocked states.

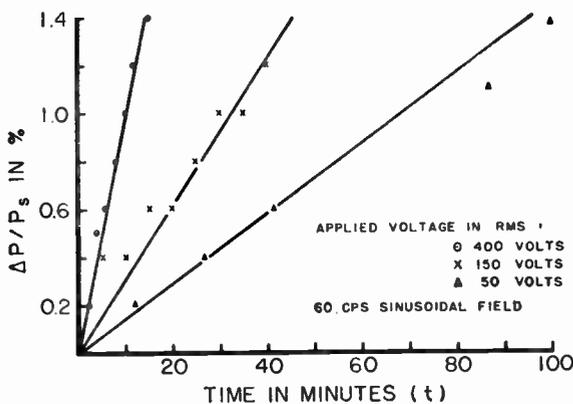


Fig. 4—Stability of the blocked state as a function of time.

fecting a remanent change of 1.4 per cent with a voltage three times the coercive voltage. As an illustration, it is noted that the number of switching cycles during this time was 325,000 with a 60 cps energizing voltage.

Fig. 5 gives the rate of change of polarization in function of applied energizing voltage.

These measurements show that the blocked state is quite stable and that it is difficult to force a change in this state.

### THE UNBLOCKED STATE

This state has been discussed in connection with the illustrations in Figs. 1(a) and 1(c). The polarization of the unblocked state can be reversed. It then behaves like a normal ferroelectric condenser.<sup>8</sup> For determining the behavior of the unblocked state, pulse measurements were used. The plot given by circles in Fig. 6 is a plot of reciprocal switching time  $1/t_{max}$ , which is the switching time pertaining to the maximum value of the

<sup>8</sup> C. F. Pulvari and W. Kuebler, "A phenomenological theory of polarization reversal in BaTiO<sub>3</sub> single crystals," *J. Appl. Phys.*, vol. 29, pp. 1315-1321; September, 1958.

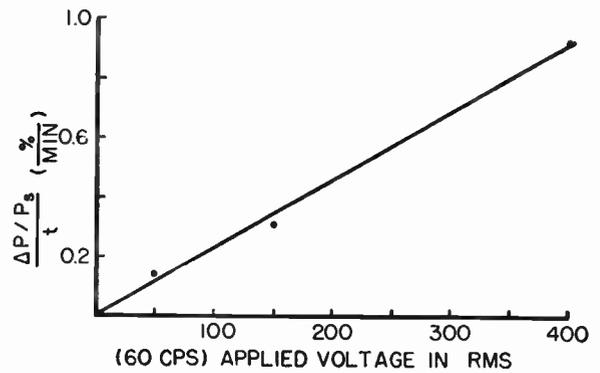


Fig. 5—Rate of change of polarization  $\Delta P/P_s$  as a function of applied energizing voltage for the blocked state.

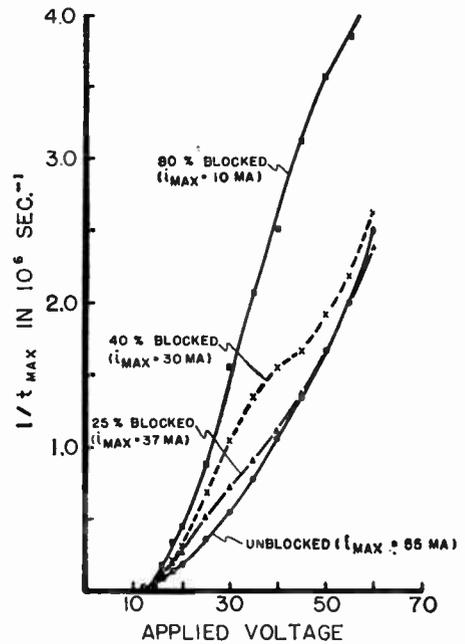


Fig. 6—Switching time characteristic of the unblocked state.

current transient  $i_{max}$ , vs applied pulse amplitude (in this case  $i_{max} = 55$  ma). The coercive voltage  $V_c$  was 18 v for the unblocked state. The plots given by the measured points in triangles, crosses and squares pertain to 25 per cent, 40 per cent and 80 per cent blocked states, which leave unblocked 75 per cent, 60 per cent and 20 per cent, respectively. The maximum values of the current transients  $i_{max}$  were decreasing as 37, 30, and 10 ma corresponding to the decreasing fractions of unblocked polarization. These plots revealed an interesting change in the switching behavior, namely, the slope of the plots increase with increase of blocked-, and decrease of unblocked-polarization. This means that, in the partially blocked state, switching proceeds faster than in the unblocked state, and the mobility of domains increases. While the mobility increases, the coercive voltage  $V_c$  decreases. Fig. 7 shows the variation of  $V_c/V_{cs}$  (where  $V_{cs}$  corresponds to the completely un-

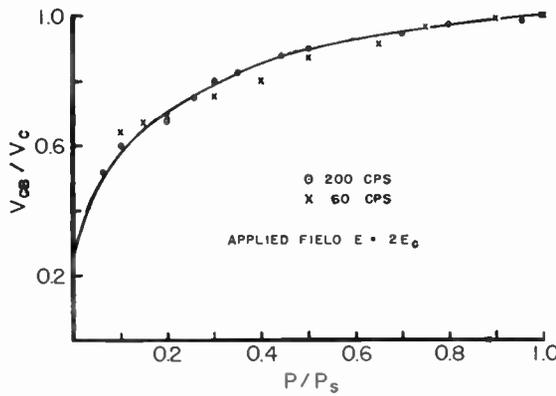


Fig. 7—Variation of coercive voltage as a function of unblocked fraction of  $P_s$ .

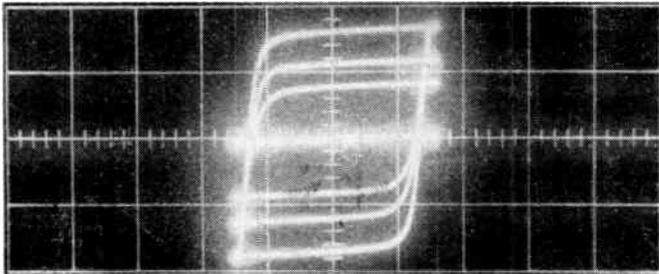


Fig. 8—Hysteresis loops of the transpolarizer for various settings.

blocked or saturated polarization) in function of unblocked fractions of  $P_s$ . It can be seen that, as blocked polarization increases and unblocked polarization decreases, the coercive voltage decreases.

A further important feature of the partially blocked state is that any level of unblocked polarization can be set. Fig. 8 is a photograph showing a family of hysteresis loops, each pertaining to a different setting including the completely blocked and unblocked maximum settings, and showing that, when a low polarization level is set, repetition rate can be increased up to a megacycle or higher (see Fig. 9). Since any appreciable hysteresis is sufficient for a transpolarizer action, it is apparent that very little power is required for control purposes. As an illustration, a transpolarizer using as a dielectric TGS or TGFB whose polarization is about  $3 \times 10^{-6}$  coulomb/cm<sup>2</sup> with an area of  $A = 5.25 \times 10^{-6}$  cm<sup>2</sup> and a coercive voltage of 5 volts needs only about  $10^{-10}$  watt-second, or less, for setting. If the area  $A = 25 \times 10^{-4}$  cm and BaTiO<sub>3</sub> is used, this increases to about  $3 \times 10^{-7}$  watt-second. Small transpolarizers have been controlled with electron beams<sup>9</sup> in the microsecond region and it is visualized that new, large-scale storage, display<sup>7</sup> and pick-up, etc., devices would now be practical.

<sup>9</sup> F. Weekes, "The effect of secondary electron emission on the switching transient of ferroelectrics," unpubl. Master's thesis, Dept. Elec. Engrg., The Catholic University of America, Washington, D.C.; January, 1959.

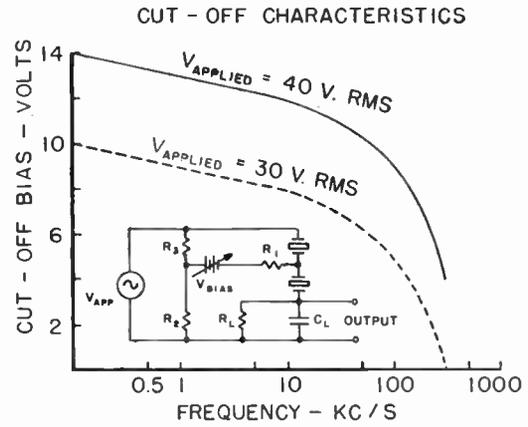


Fig. 9—Cutoff characteristic of a transpolarizer as a function of frequency.

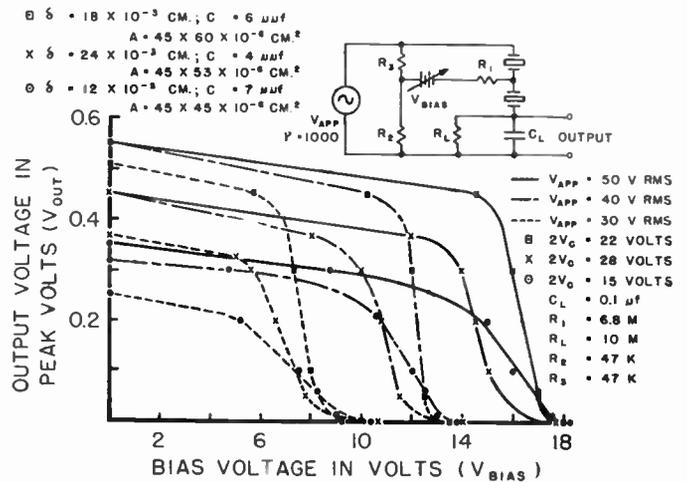


Fig. 10—Cutoff characteristic of a transpolarizer with capacitive load.

### BLOCKING CHARACTERISTICS

When the blocking characteristic of a two-path transpolarizer is investigated, the control function of an applied bias is of interest. Fig. 10 presents a BaTiO<sub>3</sub> transpolarizer with a capacitive load. A low measuring frequency of 1000 cps was applied in this case; this permits observation of a hysteresis loop.  $R_L$  was a large value of 10 megohms which served only to secure a path for the bias. The variable bias was applied through  $R_1 = 6.8$  megohms. Its value was so chosen that its shunt effect on the storage elements should be negligible.  $R_2 = R_3$  was 47 Kohm.  $P_r$  could be calculated from the output voltage on the load capacity  $C_L$  and area of condenser. The output voltage is given as the half-loop height, which corresponds to a charge transfer through the transpolarizer,

$$Q = P_r A = VC_L \tag{1}$$

Because the transpolarizer transfers  $Q$  with an applied

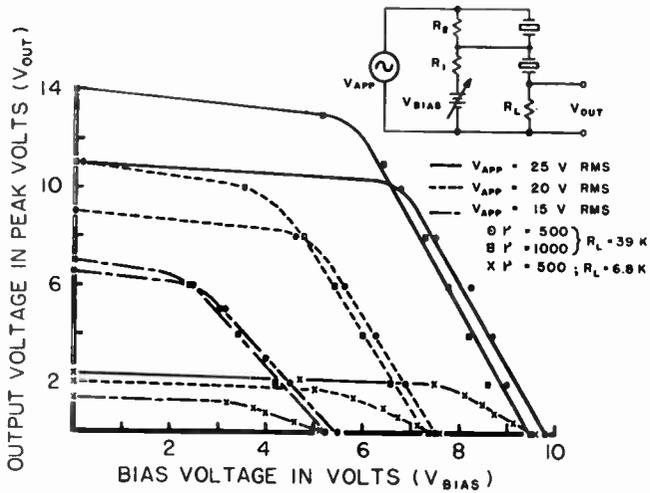


Fig. 11—Cutoff characteristic of a transpolarizer with resistance load.

voltage of  $V_{app}$ , the nonlinear capacitance  $C_T$ , during switching of the unblocked polarization becomes,

$$C_T = \frac{Q}{(V_{app} - V_{out})} \quad (2)$$

Using the data given in Fig. 10, the maximum value of  $C_T$  was approximately 1000  $\mu\mu\text{f}$  and the minimum value of  $C_T$  when blocked was approximately 3-4  $\mu\mu\text{f}$  for the two condensers in series. This gives an impedance ratio of approximately 300:1, which agrees well with data obtained by resonance measurements. A similar measurement was made when TGS and TGFB were used as dielectrics; the impedance ratio for these cases was found to be about 1000:1.

Fig. 10 gives variation of output voltage as the dc bias was varied. It is apparent from this measurement that a rather sharp threshold or cutoff bias value exists which depends only slightly on crystal thickness, and mainly on applied voltage. As energizing voltage increases, blocking potential required as a bias also increases. These characteristic plots present the blocking function of the bias.

Fig. 11 gives a similar plot with a resistive load. In this case the output was not sinusoidal and  $i_{max}$  was measured, and since no integration occurs, the blocking potential gave a sharper threshold value. The application of dc bias was slightly different. In this case, the frequency of applied voltage, as well as load resistance, was varied to see what effect this would have on the cutoff bias. For the frequency used,  $R_L$  is negligible compared to the capacitive reactances and would not have much effect on total current. Therefore, cutoff will not be affected significantly by  $R_L$  but the output voltage will be proportional to  $R_L$ . This circuit can be used as a power amplifier, chopper, etc. Using the same basic circuit as shown in Fig. 10, the variation of cutoff bias was measured when the frequency of applied energizing voltage was varied. In Fig. 9, the cutoff values are

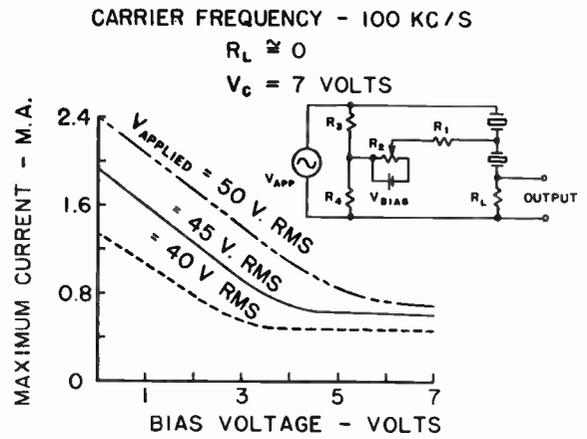


Fig. 12—Short-circuited transfer characteristic at 100 kc/s as a function of bias voltage and applied voltage.

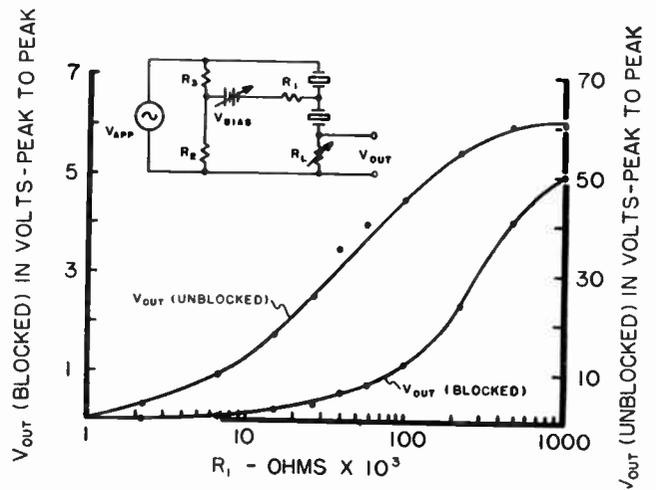


Fig. 13—Signal-to-noise ratio as a function of load resistance.

given for  $V_{app} = 30$  and 40 v. This measurement indicates that, when the frequency of  $V_{app}$  increases, a lower cutoff bias is required to set the blocked state. This plot also shows that, in a partially blocked state, repetition rate can be increased up to the megacycle region.

Fig. 12 gives the short-circuited transfer characteristic (because  $R_L$  was nearly 0) at 100 kc in function of bias voltage and applied field. Regarding the transpolarizer as a four-terminal network, the maximum power gain attainable was 2100. The input impedance of the device was 30 megohms while the output impedance was found to be 1 Kohm, at this power gain. A TGS transpolarizer was used for these measurements.

Although this paper discusses the two-section transpolarizer with two elements of nearly equal electrical and physical characteristics, the foregoing applies also to a multisection transpolarizer composed of asymmetrical elements.

For the measurements presented in Fig. 9 and Figs. 11-13,  $\text{BaTiO}_3$  transpolarizers were used having a coercive voltage  $V_a = 7$  volts, an area  $A = 45 \times 45 \times 10^{-6} \text{ cm}^2$  and a thickness,  $\delta = 12 \times 10^{-3} \text{ cm}$ .

### CONTROL RATIO

The device signal-to-noise ratio can be defined as the signal ratio of the unblocked and blocked transpolarizer. Fig. 13 gives output voltages for the unblocked- and blocked-state when the load resistance  $R_L$  is varied and the frequency of applied voltage was constant at 1000 cps. As load resistances decrease, signal-to-noise ratios increase and its maximum value is about 100; as the load resistance increases, this value drops to about 12. Although the impedance ratio of a transpolarizer is much higher (300:1), circuit capacitances increase the small blocked capacity and this, in turn, decreases the actual obtainable signal-to-noise ratio.

A similar investigation can be carried out when the frequency of applied voltage is varied. These measurements were not made because special instrumentation would have been needed to avoid the heating effect involved at the higher frequencies.<sup>10</sup> Briefly, it is noted that calculation shows a similar change for increasing frequency, *i.e.*, the signal-to-noise ratio decreases with increasing frequency.

### CONCLUSION

The important basic features of the transpolarizer are summarized as follows:

a) Polarization conditions, once set in the two cells, stay permanently and represent a stored control; the transpolarizer, in an "on-off" operating mode, is a basic one-bit storage element whose state can be sensed a large number of times without causing an alteration in its state; consequently, it affords a "nondestructive" read-out.

b) Besides being an "on-off" device, an intermediate setting is possible by using a single appropriate amplitude block- or unblock-pulse according to which an output of any desired level in a continuous range

between the blocked- (almost zero) and unblocked-levels can be produced (Fig. 8). This property permits use of only a fraction of the polarization to be switched and a high repetition rate becomes possible.

c) It has also the ability to control a long flow of alternating electric power according to the setting established by a single electric pulse. The energizing ac sensing signal will or will not produce an ac output depending on the sign and amplitude of the last setting, block- or unblock-pulse, respectively.

d) It is a high-impedance input device and a low-impedance output device and needs very little control power; therefore, it is readily controlled by such high-impedance means as an electron beam or photoelectric device.

e) The transpolarizer combines in effect the function of a dielectric amplifier and a ferroelectric storage cell.

The very special switching and storage properties, as well as the stability, of the transpolarizer open a large field of new applications and permit production of new devices and systems employing switching with a permanent setting, and, more particularly, permit production of small- and large-scale storage devices with non-destructive readout, large screen display devices with storage properties, high impedance input switches, and in general, photoelectric switches, decoders, function generators, etc.

The transpolarizer is in effect a solid state relay, *i.e.*, a switching device with storage ability and simple construction. It is rugged and stable in operation; no permanent deterioration can occur by overdriving any of the associated circuits.

### ACKNOWLEDGMENT

The author wishes to acknowledge the assistance of George E. McDuffie, Jr., Wolf Kuebler, and E. Joseph Thompson who helped in many experimental phases of the work and greatly contributed to the functional understanding of the transpolarizer.

<sup>10</sup> J. Hou-Ying Tai, "Pulsed sine-wave hysteresis loop of BaTiO<sub>3</sub>," unpubl. Master's thesis, Dept. Elec. Engrg., The Catholic University of America, Washington, D. C.; May, 1957.

# Geometric-Analytic Theory of Transition in Electrical Engineering\*

E. FOLKE BOLINDER†, MEMBER, IRE

**Summary**—A geometric-analytic theory of transition is presented and applied to circuit theory. A transition from one state to another is represented in a complex plane by two points which, by variation of a parameter, approach each other, coalesce, and then separate along trajectories perpendicular to the original trajectories.

Three analogous cases are treated, namely

- 1) Movements of fixed points in the complex impedance plane and the complex reflection coefficient plane (Smith chart),
- 2) Movements of poles in the complex frequency plane, and
- 3) Movements of saddle points in the complex frequency plane.

In the analytic treatment, the linear fractional transformation (Möbius transformation) is used, which makes conformal graphical methods applicable in the geometric treatment. Such a method is, for example, the isometric circle method.

By mapping stereographically the complex plane on the Riemann unit sphere, we see that a transition can be represented in three dimensions by the movements of two straight lines, each being the polar of the other with respect to the sphere. The transition takes place when both lines are perpendicular and tangent to the sphere at a point corresponding to the transition point.

## I. INTRODUCTION

A COMPLEX quantity  $w = u + jv$  ( $j^2 = -1$ ) can be represented by a point with the coordinates  $u$  and  $v$  in a plane. The use of such a "complex plane" seems to have been suggested independently by Wessel (1745–1818), Argand (1768–1822), and Gauss (1777–1855). Ever since Steinmetz, in 1895, introduced the complex plane in electrical engineering, this simple mathematical tool has obtained extensive applications, especially in circuit theory.

In the present paper, the complex plane will be used in a simple geometric-analytic theory of transition in electrical engineering. The geometric part of the theory utilizes two "isometric" circles which are usually used in connection with a graphical method of transforming a complex quantity by means of the linear fractional transformation (Möbius transformation), called the "isometric circle method."<sup>1</sup> The analytic part consists of the use of quadratic equations. After a brief study of these tools, we are going to investigate some simple examples of fixed point trajectories in the complex impedance and reflection coefficient planes, and pole and saddlepoint trajectories in the complex frequency plane.

\* Original manuscript received by the IRE, December 10, 1958; revised manuscript received, March 2, 1959. Presented at the URSI-IRE Fall Meeting, Pennsylvania State College, State College, Pa., October 21–23, 1958.

† Electromagnetic Radiation Lab., Air Force Cambridge Res. Center, Bedford, Mass.

<sup>1</sup> E. F. Bolinder, "Impedance and polarization-ratio transformations by a graphical method using the isometric circles," IRE TRANS. ON MICROWAVE THEORY AND TECHNIQUES, vol. MTT-4, pp. 176–180; July, 1956.

## II. ISOMETRIC CIRCLES

The isometric circle is defined as the circle that is the complete locus of points in the neighborhood of which lengths are unaltered in magnitude by the linear fractional transformation

$$w' = \frac{aw + b}{cw + d}, \quad ad - bc = 1 \quad (1)$$

where  $w = u + jv$ ,  $w' = u' + jv'$ , and  $a$ ,  $b$ ,  $c$ , and  $d$  are complex constants obeying the condition  $ad - bc = 1$ .

From (1) we obtain

$$\frac{dw'}{dw} = \frac{1}{(cw + d)^2}, \quad ad - bc = 1, \quad (2)$$

so that the isometric circle of the direct transformation is

$$|cw + d| = 1, \quad c \neq 0. \quad (3)$$

Similarly, the inverse transformation

$$w = \frac{-dw' + b}{cw' - a}, \quad ad - bc = 1 \quad (4)$$

has the isometric circle

$$|cw' - a| = 1, \quad c \neq 0. \quad (5)$$

Thus the isometric circle of the first direct transformation,  $C_d$ , has its center at  $O_d = -d/c$  and radius  $R_c = 1/|c|$ ; the isometric circle of the inverse transformation,  $C_i$ , has its center at  $O_i = a/c$  and the same radius. See Fig. 1.

Mathematically, (1) is divided into two classes of transformation: the loxodromic transformation, characterized by  $a+d = \text{complex}$ , and the nonloxodromic transformation, characterized by  $a+d = \text{real}$ . The second class is further divided into hyperbolic ( $|a+d| > 2$ ), parabolic ( $a+d = \pm 2$ ), and elliptic ( $|a+d| < 2$ ) transformations.

In Fig. 1 the isometric circle method is outlined. It consists in the loxodromic case of 1) an inversion in the isometric circle of direct transformation  $C_d$ , ( $w \rightarrow w_1$ ); 2) a reflection in the symmetry line  $L$  to the two circles, ( $w_1 \rightarrow w_2$ ); and 3) a rotation around the center  $O_i$  of the isometric circle of the inverse transformation through an angle  $-2 \arg(a+d)$ , ( $w_2 \rightarrow w'$ ). In the nonloxodromic case  $a+d = \text{real}$ , so that the third operation is eliminated ( $w_2 = w'$ ). In this work only the isometric circles and not the isometric circle method itself will be used.

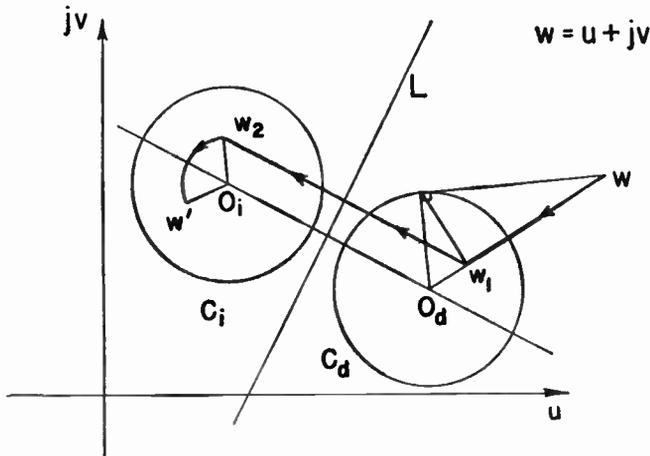


Fig. 1—The isometric circle method.

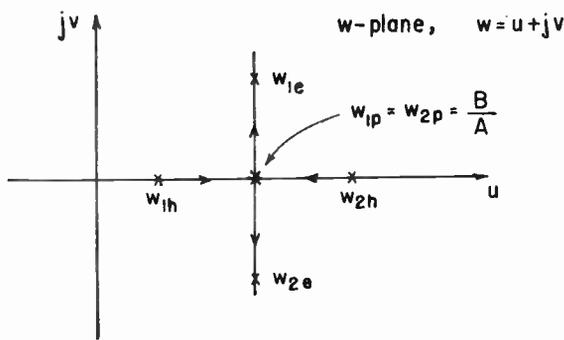


Fig. 2—The roots of a quadratic equation plotted in a complex plane.

III. QUADRATIC EQUATIONS

If we solve an arbitrary complex quadratic equation

$$Aw^2 - 2Bw + C = 0 \tag{6}$$

where  $A$ ,  $B$ , and  $C$  are real constants, we obtain

$$\left. \begin{matrix} w_1 \\ w_2 \end{matrix} \right\} = \frac{B}{A} \pm \frac{\sqrt{B^2 - AC}}{A} \tag{7}$$

In (7)  $B^2 - AC$  is called the discriminant. If the discriminant is positive, two real roots  $w_{1h}$  and  $w_{2h}$  are obtained which we can plot in the complex  $w$  plane. See Fig. 2. If the discriminant is negative, two conjugate complex roots  $w_{1e}$  and  $w_{2e}$  are obtained:

$$\left. \begin{matrix} w_{1e} \\ w_{2e} \end{matrix} \right\} = \frac{B}{A} \pm j \frac{\sqrt{AC - B^2}}{A} \tag{8}$$

See Fig. 2. Finally, if the discriminant equals zero the two roots coalesce and we obtain

$$w_{1p} = w_{2p} = \frac{B}{A} \tag{9}$$

If we assume that the constants  $A$ ,  $B$ , and  $C$  are functions of a certain parameter, then, by varying the parameter,  $w_1$  and  $w_2$  may approach each other, coalesce, and separate along trajectories perpendicular to the original trajectories. A "transition" between two

"states" is obtained. We are now going to study simple examples of how such a transition may occur in connection with fixed point trajectories in the impedance and reflection coefficient planes, and pole and saddle-point trajectories in the complex frequency plane.

IV. FIXED POINT TRAJECTORIES

Impedance and Reflection Coefficient Transformations Through Bilateral Two-Port Networks

The input voltage  $V'$  and the input current  $I'$  of a bilateral two-port network can be expressed in the output voltage  $V$  and output current  $I$  by the following linear equations:

$$\left. \begin{matrix} V' = aV + bI \\ I' = cV + dI \end{matrix} \right\} \tag{10}$$

where, at a fixed frequency, the four complex constants satisfy the reciprocity relations

$$ad - bc = 1. \tag{11}$$

If we put

$$\left. \begin{matrix} \frac{V'}{I'} = Z' \\ \frac{V}{I} = Z \end{matrix} \right\} \tag{12}$$

we obtain

$$Z' = \frac{aZ + b}{cZ + d}; \quad ad - bc = 1. \tag{13}$$

This is a linear fractional transformation. By using the property that the imaginary axis of the  $Z$ -plane,  $Z = R + jX$ , is invariant under a lossless transformation, we obtain for such a transformation

$$Z' = \frac{a'Z + jb''}{jc'Z + d'}, \quad a'd' + b''c'' = 1 \tag{14}$$

where

$$\left. \begin{matrix} a = a' + ja'' \\ b = b' + jb'' \\ c = c' + jc'' \\ d = d' + jd'' \end{matrix} \right\} \tag{15}$$

The fixed points of (14) are:

$$\left. \begin{matrix} Z_{f1} \\ Z_{f2} \end{matrix} \right\} = \frac{a' - d' \pm \sqrt{(a' + d')^2 - 4}}{2jc''} \tag{16}$$

By using (14) and the well known expression for the reflection coefficient

$$\Gamma = \frac{\frac{Z}{\sqrt{2}} - \frac{1}{\sqrt{2}}}{\frac{Z}{\sqrt{2}} + \frac{1}{\sqrt{2}}}; \quad \Gamma' = \frac{\frac{Z'}{\sqrt{2}} - \frac{1}{\sqrt{2}}}{\frac{Z'}{\sqrt{2}} + \frac{1}{\sqrt{2}}} \tag{17}$$

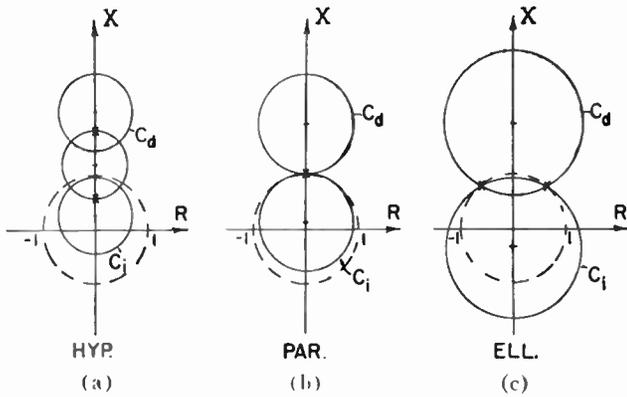


Fig. 3—Fixed point trajectories in the complex impedance plane.

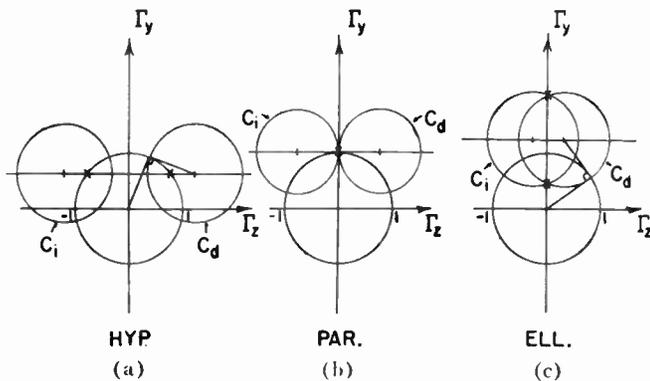


Fig. 4—Fixed point trajectories in the complex reflection coefficient-plane (Smith chart).

we obtain

$$\Gamma' = \frac{A\Gamma + C^*}{C\Gamma + A^*} \tag{18}$$

where an asterisk indicates a conjugate complex quantity, and

$$\begin{aligned} A &= \frac{1}{2} [(a' + d') - j(b'' + c'')] \\ C &= \frac{1}{2} [(a' - d') - j(b'' - c'')] \end{aligned} \tag{19}$$

*Fixed Point Trajectories in the Z Plane and the Smith Chart, and on the Riemann Sphere*

Since the distance between the centers of the two isometric circles in the nonloxodromic case is  $(a' + d')/c''$ , while the sum of the two radii is  $2/c''$ , it follows that the hyperbolic case is obtained, if the two circles are external; the parabolic case, if they are tangent; and the elliptic case, if they intersect. In the Z plane the isometric circles have their centers on the imaginary axis in the lossless case. A simple example indicating the isometric circle positions of a lossless exponentially tapered transmission line is shown in Fig. 3. A geometric interpretation of (16) shows that in the hyperbolic or below cutoff case, the fixed points are obtained where a circle, orthogonal to the two isometric circles, cuts the imaginary axis. In the parabolic or cutoff case,

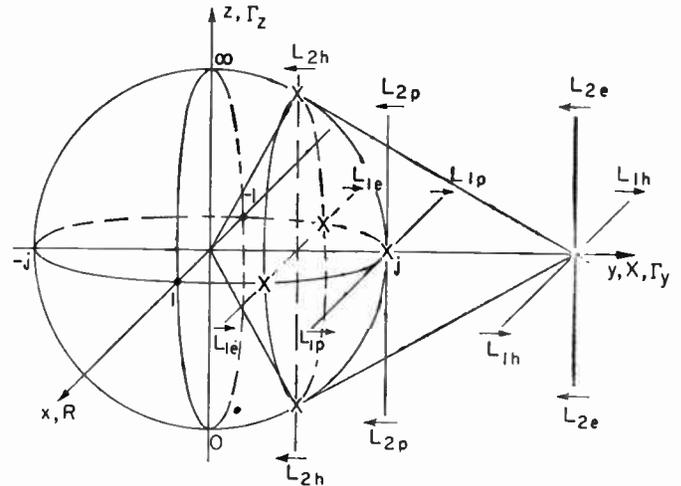


Fig. 5—Fixed point trajectories of the Riemann unit sphere.

the fixed points coalesce at the point of tangency of the two isometric circles. Finally, in the elliptic or above cutoff case, the fixed points constitute the crossover points of the two isometric circles. We find, that if we vary a parameter, for example, the length of the line or the frequency, the fixed points move on the imaginary axis, coalesce, and then separate perpendicularly to the original trajectories following the unit circle.

The corresponding trajectories in the Smith chart are shown in Fig. 4. Here the fixed points move on the unit circle in the hyperbolic case, coalesce, and then separate perpendicularly following the imaginary axis of the  $\Gamma$ -plane ( $\Gamma = \Gamma_z + j\Gamma_p$ ) in the elliptic case.

If we map the Z plane stereographically on the unit Riemann sphere with the top of the sphere,  $(x, y, z) = (0, 0, 1)$ , as projection center, the Smith chart is obtained in the  $yz$  plane by an additional stereographic projection from the point  $(-1, 0, 0)$ .<sup>2-4</sup> See Fig. 5. By plotting the fixed point trajectories of the lossless exponentially tapered transmission line studied above on the sphere, we find that by using three dimensions the discontinuous movements of the fixed points in two dimensions are replaced by a continuous movement. In the elliptic case, the fixed points move on the unit circle in the  $xy$  plane. We connect the points by a straight line  $L_{1e}$  parallel to the  $x$  axis. This line has a polar line  $L_{2e}$  parallel to the  $z$  axis. When the line  $L_{1e}$  moves towards the point  $(0, 1, 0)$ ,  $L_{2e}$  moves towards the same point. In the parabolic case both lines, now called  $L_{1p}$  and  $L_{2p}$ , are tangent to the sphere. If the lines continue their movements the hyperbolic case is obtained. Now the

<sup>2</sup> F. Steiner, "Die Anwendung der Riemannschen Zahlenkugel und ihrer Projektionen in der Wechselstromtechnik," *Radiowelt*, vol. 1, pp. 23-26; October, 1946.

<sup>3</sup> H. A. Wheeler, "Geometric relations in circle diagrams of transmission-line impedance," *Wheeler Monographs*, Wheeler Labs., Great Neck, N. Y., vol. 1, no. 4; 1948.

<sup>4</sup> G. A. Deschamps, "Geometric viewpoints in the representation of waveguides and waveguide junctions," *Proc. Symp. on Modern Network Synthesis*, Polytechnic Inst. of Brooklyn, Bklyn., N. Y., vol. 1, pp. 277-295; April, 1952.

first line, called  $L_{1h}$ , falls outside the sphere, while the second line  $L_{2h}$  cuts the sphere in the fixed points situated on the unit circle in the  $yz$  plane. Thus a smooth transition between the elliptic and hyperbolic states is obtained.

V. POLE TRAJECTORIES

The Resonant Circuit

The input impedance of a simple parallel resonance circuit consisting of an inductance  $L$  with a series resistance  $r$ , and a capacitance  $C$  with a conductance  $G$  (see Fig. 6) is

$$Z(s) = \frac{1}{C} \frac{s + r/L}{s^2 + (r/L + G/C)s + (1 + rG)/LC} \quad (20)$$

where  $s$  is the complex frequency,  $s = \sigma + j\omega$ . Eq. (20) can be written

$$Z(s) = \frac{1}{C} \frac{1}{s + \frac{sG/C + (1 + rG)/LC}{s + r/L}} \quad (21)$$

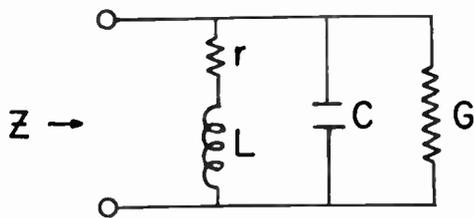


Fig. 6—Resonant circuit.

The positions of the poles,  $s_{p1}$  and  $s_{p2}$ , are obtained by making the denominator equal to zero

$$s_p = \frac{-s_p G/C - (1 + rG)/LC}{s_p + r/L} \quad (22)$$

Eq. (22) is analogous to the equation used in the preceding section in obtaining the fixed points  $Z_{f1}$  and  $Z_{f2}$  of the linear fractional transformation (13). To obtain an exact analogy, the coefficients in (22) have to obey the condition  $ad - bc = 1$ . Eq. (22) then transforms into

$$s_p = \frac{-s_p G\sqrt{L/C} - (1 + rG)/\sqrt{LC}}{s_p\sqrt{LC} + r\sqrt{C/L}} \quad (23)$$

Pole Trajectories in the Complex Frequency Plane and on the Riemann Sphere

In the  $Z$  plane, the positions of the fixed points were obtained from the positions of the isometric circles. Analogous conditions yield two circles in the  $s$  plane with centers at  $-d/c = -r/L$  and  $a/c = -G/C$ , both having the radius  $1/|c| = 1/\sqrt{LC} = \omega_0$ , the resonance (radian) frequency. The positions of these circles immediately specify the pole positions. See Fig. 7. If the

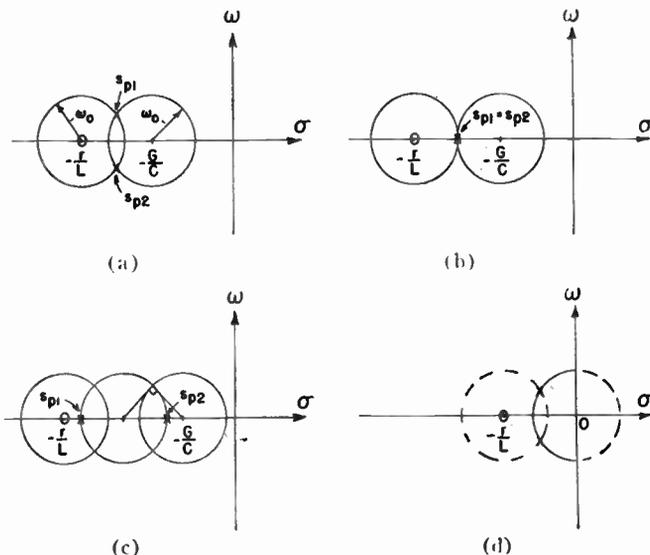


Fig. 7—Pole trajectories in the complex frequency plane.

two circles intersect [Fig. 7(a)], two complex conjugate poles are obtained, corresponding to the damped oscillatory case. If the two circles are tangent [Fig. 7(b)], two coalescing real poles are obtained, corresponding to the critically damped case. If, finally, the two circles are external [Fig. 7(c)], two real poles are obtained as the crossover points of the real axis and a circle that is orthogonal to the two given circles, corresponding to the aperiodic case. In all cases one zero of (20) is situated at infinity,  $s_{01} = \infty$ , and another at  $s_{02} = -r/L$ .

Thus, we find that by varying the conductance  $G$ , for example, the poles follow the fixed circle with its center at  $-r/L$ , coalesce, and then separate perpendicularly following the real axis. In various textbooks, we find that the pole trajectories shown in Fig. 7(d) are obtained for an  $rLC$  circuit by varying the resistance  $r$ . This figure is immediately explained by the graphical method that has been described. With  $G = 0$ , one circle is fixed with its center at the origin of the  $s$  plane, and the other moves along the negative real axis as  $r$  is varied. The trajectories are therefore the real axis and the fixed circle.

If we map the  $s$  plane stereographically on the unit Riemann sphere, the example chosen can be treated by the polar theory of the preceding section. See Fig. 8. The pole trajectories on the sphere are indicated by heavy lines.

VI. SADDLEPOINT TRAJECTORIES

The Inverse Laplace Transform Treated by the Saddlepoint Method

Saddlepoints of the second order play an important role in methods of approximate integration of the inverse Laplace transform

$$f(t) = \frac{1}{2\pi j} \int_{\gamma} F(s)e^{st} ds; \quad s = \sigma + j\omega. \quad (24)$$

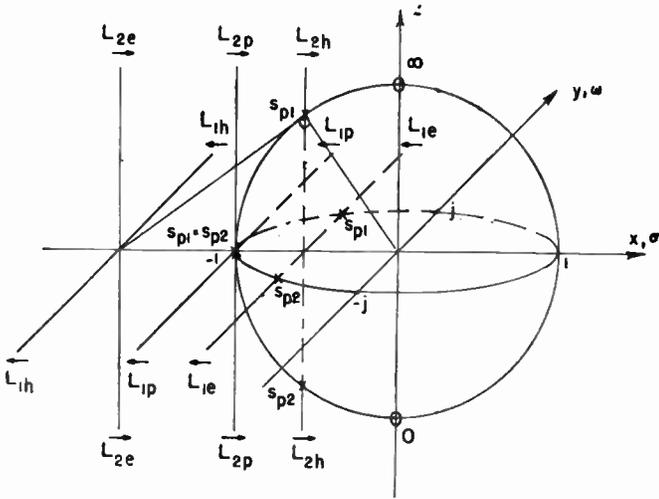


Fig. 8—Pole trajectories on the Riemann unit sphere.

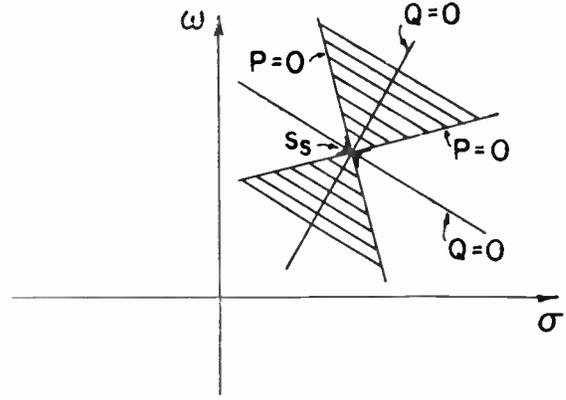


Fig. 10— $P=0$  and  $Q=0$  lines through a saddlepoint of second order.

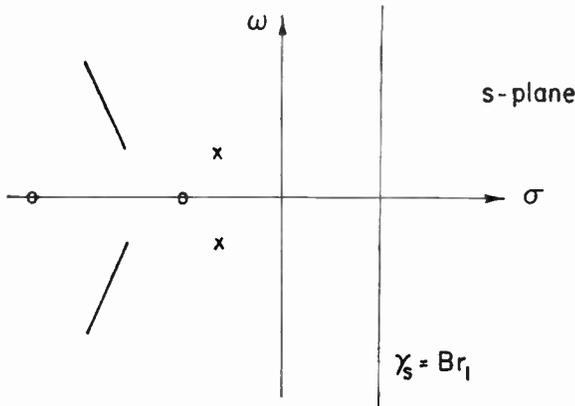


Fig. 9—Example of poles, zeros, branch-cuts, and the Bromwich-t contour of integration in the complex frequency plane.

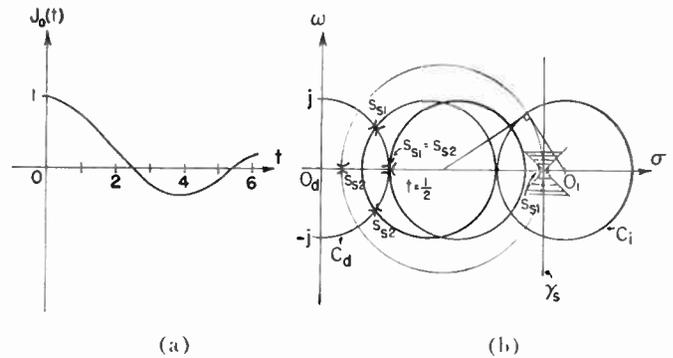


Fig. 11—(a) The Bessel function of the first kind and of order zero (b) Saddlepoint trajectories in the complex frequency plane.

In (24),  $t$  indicates time, and  $\gamma_s$  the contour of integration in the  $s$  plane. This contour of integration may be a straight line parallel to the  $\omega$  axis (“Bromwich—1” contour) (see Fig. 9) or the topologically equivalent contour folded around the singularities in the left half of the  $s$  plane (“Bromwich—2” contour).

We can write (24) in the following form:

$$f(t) = \frac{1}{2\pi j} \int_{\gamma_s} e^{ln F(s) + st} ds = \frac{1}{2\pi j} \int_{\gamma_s} e^{W(s,t)} ds. \quad (25)$$

If we put

$$\frac{\partial W(s, t)}{\partial s} = 0, \quad (26)$$

the roots of (26) correspond to points, called saddlepoints, in the  $s$  plane. These points, which are used in a theory of approximate integration—“the saddlepoint method” (first invented by B. Riemann in 1863)—follow certain trajectories in the  $s$  plane when the time  $t$  varies. The contour of integration is changed until it passes through the saddlepoints along the paths of

steepest descent. These paths are found by expanding  $W(s, t)$  in a Taylor series in the vicinity of the saddlepoint  $s_s$ .<sup>5</sup> We put

$$W(s, t) - W(s_s, t) = P + jQ \quad (27)$$

and determine the line  $Q=0$  for which the real part of  $W(s, t) - W(s_s, t)$  is maximum negative. In Fig. 10 the region of  $P < 0$  around a saddlepoint of the second order is shadowed. The line  $Q=0$  in that region constitutes the line of steepest descent.

Let us study a simple example.<sup>5</sup> The Bessel function of the first kind and of order zero,  $J_0(t)$ , plotted in Fig. 11(a), has the Laplace transform  $1/\sqrt{s^2+1}$ , so that

$$f(t) = J_0(t) = \frac{1}{2\pi j} \int_{\gamma_s} \frac{1}{\sqrt{s^2+1}} e^{st} ds. \quad (28)$$

Here

$$\frac{\partial W(s, t)}{\partial s} = -\frac{s}{s^2+1} + t = 0 \quad (29)$$

<sup>5</sup> M. V. Cerrillo, “On the evaluation of integrals of the type

$$f(\tau_1, \tau_2, \dots, \tau_n) = \frac{1}{2\pi j} \int F(s) e^{W(s, \tau_1, \tau_2, \dots, \tau_n)} ds$$

and the mechanism of formation of transient phenomena,” in “An Elementary Introduction to the Theory of the Saddlepoint Method of Integration,” Res. Lab. of Electronics, M.I.T., Cambridge, Mass., Tech. Rep. No. 55:2a; May 3, 1950.



# A Phenomenological Theory of the Reggia-Spencer Phase Shifter\*

JERALD A. WEISS†

**Summary**—The novel ferrite-actuated variable phase shifter reported by F. Reggia and E. G. Spencer operates on a principle which, while resembling that of the familiar Faraday rotation devices, differs from them in a way which leads to quite different and unexpected behavior. The observed properties of the device depend on the interaction between the radiation and the ferrite taking place under several special conditions: namely, elliptic waveguide symmetry, a variation in the permeability associated with the magnetic disorder of the unsaturated ferrite, and a dielectric-waveguide effect.

The essential properties of the device can be explained with the aid of a simplified model in which the radiation is represented by plane waves. The contribution of the dielectric-waveguide effect can be estimated on the basis of known solutions of a simpler problem, namely that of a dielectric rod partially filling circular guide.

The model furnishes insight into the characters of the modes of propagation under conditions resembling those of the Faraday rotator but subject to elliptic symmetry and other complications.

## INTRODUCTION

A VARIABLE microwave phase shifter, employing a longitudinally magnetized ferrite rod in rectangular waveguide has been described by Reggia and Spencer.<sup>1</sup> The structure has also been mentioned by Brown, Cole, and Honeyman.<sup>2</sup> The behavior of the device differs radically from that of the well-known Faraday rotator which it resembles, suggesting that the two effects are special cases of a more general principle, or combination of principles, which may lead to a variety of novel devices. The phase shifter itself offers a fresh approach to the problem of producing phase variations under conditions appropriate to the various applications, such as scanning antennas, bridge circuits, etc.

The following discussion presents a composite phenomenological model which exhibits the essential properties of the Reggia-Spencer phase shifter and incorporates the minimum number of contributing phenomena required to produce the over-all effect. The aim of the model is to reduce the problem to such a degree of simplicity that it can be solved without recourse to numerical methods, thereby providing an advantageous starting point for such calculations, as well as a qualitative guide to further experimental development.

\* Original manuscript received by the IRE, October, 3, 1958; revised manuscript received, February 27, 1959.

† Bell Telephone Labs., Inc., Murray Hill, N. J.

<sup>1</sup> F. Reggia and E. G. Spencer, "A new technique in ferrite phase shifting for beam scanning of microwave antennas," *Proc. IRE*, vol. 45, pp. 1510-1517; November, 1957.

<sup>2</sup> A. C. Brown, R. S. Cole, and W. N. Honeyman, "Some applications of ferrites to microwave switches, phasers, and isolators," *Proc. IRE*, vol. 46, pp. 722-727; April, 1958.

## STRUCTURE AND BEHAVIOR OF THE PHASE SHIFTER

The device consists simply of a ferrite rod mounted on the axis of a section of standard rectangular waveguide and magnetized longitudinally by means of a solenoid. See Fig. 1. Large variations in the phase of the transmitted radiation occur in the range of applied magnetic fields in which the ferrite is partially magnetized. The effect is reciprocal: on application of the dc field, the change in phase shift (usually increasing at first) is the same for both directions of propagation. Typical data are shown by Reggia and Spencer<sup>1</sup> and in Figs. 2 and 3. After reaching a maximum at a field corresponding approximately to the saturation magnetization of the ferrite, the phase decreases slowly with further increase in the field. Associated with the phase variation is an undulation in insertion loss which is, however, too small to be accounted for by rotation effects. It can be minimized by means of familiar matching techniques such as those employed by Reggia and Spencer;<sup>1</sup> namely, tapering of the ferrite rod or addition of dielectric matching elements. Variations in insertion loss between a minimum of 0.5 db and a maximum of 1.5 db are typical when no special provision is made for matching. The amount of phase shift is approximately proportional to the length of the rod and is frequency sensitive to a degree; see Fig. 3. It is extremely sensitive to the diameter of the rod, undergoing a spectacular increase over a narrow range whose significance will be discussed below. Finally, it is roughly proportional to the saturation magnetization of the ferrite.

## THE PHENOMENA CONTRIBUTING TO THE BEHAVIOR OF THE PHASE SHIFTER

A theoretical representation of this device is required to show at least the following details: 1) The origin of the large phase shift; 2) Why large fluctuations in transmission with changes in frequency and applied field, such as would be associated with Faraday rotation,<sup>3,4</sup> do not occur; 3) What parameters determine the operating range of the device; 4) What role is played by the magnetic and dielectric properties of the ferrite; 5) Some

<sup>3</sup> C. L. Hogan, "The ferromagnetic Faraday effect at microwave frequencies and its applications," *Rev. Mod. Phys.*, vol. 25, pp. 253-263; January, 1953.

<sup>4</sup> J. A. Weiss, "An interference effect associated with Faraday rotation, and its application to microwave switching," *Proc. AIEE Conf. on Magnetism and Magnetic Materials*, Boston, Mass., pp. 580-585; October, 1956.

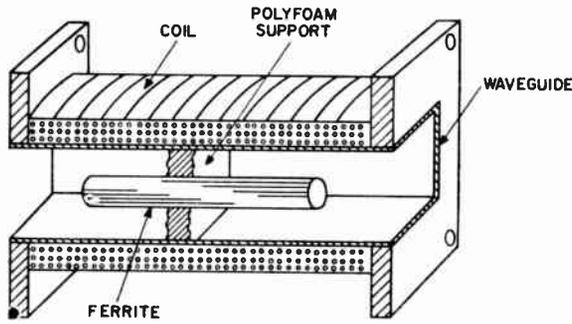


Fig. 1—Cutaway view of a Reggia-Spencer phase shifter.

general features of the modes of propagation; 6) What considerations govern the scattering characteristics of the ferrite-loaded section.

The development presented below provides more or less schematic information on these points. It will be shown there that the phase shift arises essentially from variations in the diagonal component  $\mu$  of the permeability tensor<sup>5</sup> which characterizes the ferrite, the off-diagonal component  $\kappa$  playing a secondary role. As shown in Fig. 5, the quantity  $\mu$  starts from a value somewhat less than one in the unmagnetized ferrite and increases abruptly on application of the longitudinal dc field, nearly reaching one as the material saturates and then decreasing again. This phenomenon originates in the presence of internal demagnetizing fields<sup>6</sup> associated with the magnetic disorder of the partially magnetized polycrystalline ferrite. The increase in  $\mu$  causes a redistribution of the radiation fields with the wave concentrating progressively more heavily in the rod; that is, the tendency of the rod to behave as a dielectric waveguide is steadily enhanced. Associated with this redistribution is a decrease in guide wavelength which results in the observed phase variations.

Although the high dielectric constant of the ferrite (typically about ten times that of air) plays an important part in the effect, the immediate cause of the variations in phase is magnetic in origin. Thus the phenomenon might properly be called a "variable-dielectric-magnetic-waveguide effect." Since the field configurations involved are, however, of the same general character as those encountered in the familiar case of a purely dielectric rod, we shall for brevity use the term "dielectric waveguide."

In a range of rod diameters in which this variable-

<sup>5</sup> D. Polder, "On the theory of ferromagnetic resonance," *Phil. Mag.*, vol. 40, pp. 99-115; January, 1949. The components  $\mu$  and  $\kappa$  of the permeability tensor are given in terms of the magnetization  $M$ , dc magnetic field  $H$ , angular frequency  $\omega$ , and the gyromagnetic ratio  $\gamma \approx 2.8$  mc per oersted, as

$$\mu = 1 + \frac{4\pi\gamma^2 M H}{\gamma^2 H^2 - \omega^2}, \quad \kappa = \frac{4\pi\gamma M \omega}{\gamma^2 H^2 - \omega^2}$$

<sup>6</sup> See R. C. LeCraw and E. G. Spencer, "Domain structure effects in an anomalous ferrimagnetic resonance of ferrites," *J. Appl. Phys.*, vol. 28, pp. 399-405; April, 1957.

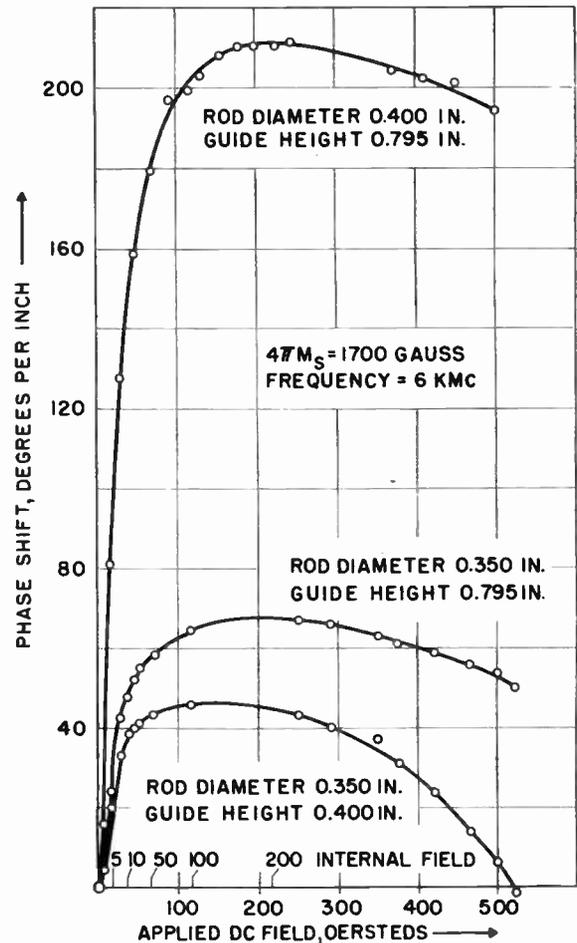


Fig. 2—The Reggia-Spencer phase shifter: observed phase shift in three structures as functions of applied dc field. The internal field values are estimated from the magnetization curve of the ferrite; the rods are 5 inches long.

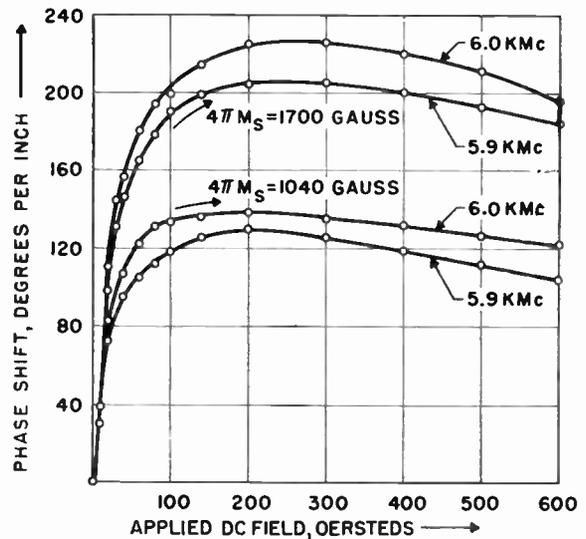


Fig. 3—The Reggia-Spencer phase shifter: observed phase shift as a function of applied dc field, showing dependence on frequency and saturation magnetization.

dielectric-waveguide effect is strong, the composite ferrite-air-metal structure can support only a single, elliptically polarized mode. Beyond a certain critical rod diameter, a second mode, also elliptically polarized, breaks into propagation, and the device exhibits the properties of a more or less elliptical Faraday rotator, becoming useless as a phase shifter due to the large fluctuations in transmission (although the phase shift is still present). Thus the operating range of the device can be specified (by an appropriate combination of frequency, ferrite diameter, and waveguide dimensions) as extending from the onset of the dielectric waveguide effect to the onset of rotation, the total amount of available phase shift becoming progressively larger from the former to the latter limit. In that range the conditions for coupling from the dominant mode in empty guide into the propagating mode on the rod appear to be peculiarly favorable. A full explanation of this will have to wait for more detailed analysis of the modes, but it is undoubtedly related to the curious fact that the two normal modes in the loaded section, one propagating and the other cut off, have the *same* sense of elliptic polarization.

THEORY

To analyze the structure, we begin by neglecting entirely the geometrical details of the waveguide cross section and represent the radiation by plane waves. The anisotropic character of the rectangular waveguide is represented by a phenomenological tensor "dielectric" constant whose components may be adjusted to represent the role of the guide geometry in establishing the phase velocities in two mutually perpendicular polarizations. This model indicates the forms of the normal modes and their dependence on the magnetic parameters  $\mu$  and  $\kappa$  as well as on the ellipticity of the structure. Following this, we perform a semiempirical comparison with the known solutions of a problem involving geometry which is similar to that of the phase shifter: namely, a circular dielectric rod on the axis of circular waveguide (the shielded dielectric waveguide)<sup>7</sup> in order to verify that the contribution of the variable-dielectric-waveguide effect is capable of accounting for the observed phase variations. Finally, we formulate some scattering relations for the structure under the assumed conditions, for comparison with the observed fluctuations in transmission.

In Maxwell's equations,

$$\begin{aligned} \text{curl } \mathbf{E} &= -\overset{\leftrightarrow}{\mu} \frac{\partial}{\partial t} \mathbf{H} \\ \text{curl } \mathbf{H} &= \overset{\leftrightarrow}{\epsilon} \frac{\partial}{\partial t} \mathbf{E} \end{aligned} \quad (1)$$

assume for  $\overset{\leftrightarrow}{\mu}$  Polder's form

$$\overset{\leftrightarrow}{\mu} = \begin{bmatrix} \mu & -i\kappa & 0 \\ i\kappa & \mu & 0 \\ 0 & 0 & \mu_0 \end{bmatrix} \quad (2)$$

and for  $\overset{\leftrightarrow}{\epsilon}$  the diagonal form

$$\overset{\leftrightarrow}{\epsilon} = \begin{bmatrix} \epsilon_x & & \\ & \epsilon_y & \\ & & \epsilon_0 \end{bmatrix}. \quad (3)$$

Also assume

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial y} = 0, \quad \frac{\partial}{\partial t} = i\omega \quad (4)$$

for plane-wave propagation in the  $\pm z$  directions at angular frequency  $\omega$ . Combining the field equations (1) we obtain the system

$$\frac{d^2}{dz^2} \begin{bmatrix} E_x \\ E_y \end{bmatrix} = -\omega^2 \begin{bmatrix} \mu\epsilon_x & -i\kappa\epsilon_y \\ i\kappa\epsilon_x & \mu\epsilon_y \end{bmatrix} \begin{bmatrix} E_x \\ E_y \end{bmatrix} \quad (5)$$

to be solved for the transverse components of the electric field  $\mathbf{E}$ . It is convenient to write (5) in the more general form

$$\frac{d^2}{dz^2} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (6)$$

The solutions of the characteristic equation

$$\lambda^2 - (a + d)\lambda + (ad - bc) = 0 \quad (7)$$

are the eigenvalues

$$\lambda_{\pm} = \frac{1}{2}(a + d) \pm \sqrt{\frac{1}{4}(a - d)^2 + bc}. \quad (8)$$

Let

$$\sigma_{\pm} = \frac{1}{2}(a \pm d)$$

and

$$R = (\text{sign of } \sigma_+) \sqrt{\sigma_-^2 + bc}; \quad (9)$$

then

$$\lambda_{\pm} = \sigma_+ \pm R. \quad (10)$$

The transformation  $S$  which diagonalizes the matrix of (6) is

$$S = \begin{bmatrix} \sigma_- + R & -b \\ c & \sigma_- + R \end{bmatrix}. \quad (11)$$

Let

$$\gamma_{\pm} = \pm \sqrt{\lambda_{\pm}} \quad (12)$$

denote the propagation constants. Then the normal modes  $p(z)$  and  $n(z)$  of the system (6) are

$$p = e^{\pm\gamma_+z}, \quad n = e^{\pm\gamma_-z}. \quad (13)$$

<sup>7</sup> See, for example, S. A. Schelkunoff, "Electromagnetic Waves," D. van Nostrand Co., Inc., New York, N. Y., sec. 10.20; 1943. Also, R. E. Beam and H. M. Wachowski, "Shielded dielectric-rod waveguides," *Trans. AIEE*, vol. 70, pp. 874-880; May, 1951. The data appropriate to the present problem are from Seidel and Brannon (see footnote 9).

Transforming back again by means of  $S$ , we obtain  $x(z)$  and  $y(z)$ , which may be conveniently expressed in the following form:

$$\begin{bmatrix} x \\ y \end{bmatrix} = U_+ \begin{bmatrix} X_+^0 \\ Y_+^0 \end{bmatrix} + U_- \begin{bmatrix} X_-^0 \\ Y_-^0 \end{bmatrix} \quad (14)$$

where  $X_{\pm}^0$  and  $Y_{\pm}^0$  are constants of integration and

$$U_{\pm} = \begin{bmatrix} \tau_+ e^{\pm\gamma+z} + \tau_- e^{\pm\gamma-z} & \frac{b}{2R} (e^{\pm\gamma+z} - e^{\pm\gamma-z}) \\ \frac{c}{2R} (e^{\pm\gamma+z} - e^{\pm\gamma-z}) & \tau_- e^{\pm\gamma+z} + \tau_+ e^{\pm\gamma-z} \end{bmatrix} \quad (15)$$

in which

$$\tau_{\pm} = \frac{1}{2} \left( 1 \pm \frac{\sigma_-}{R} \right). \quad (16)$$

Returning to the original system (5), we have for the eigenvalues,

$$\lambda_{\pm} = -\omega^2 \left[ \frac{1}{2} \mu (\epsilon_x + \epsilon_y) \pm \sqrt{\frac{1}{4} \mu^2 (\epsilon_x - \epsilon_y)^2 + \epsilon_x \epsilon_y \kappa^2} \right]. \quad (17)$$

It is convenient to introduce the combinations

$$\epsilon = \frac{1}{2} (\epsilon_x + \epsilon_y), \quad \delta = \frac{1}{2} (\epsilon_x - \epsilon_y). \quad (18)$$

Then

$$\lambda_{\pm} = -\omega^2 \left[ \mu \epsilon \pm \sqrt{\mu^2 \delta^2 + \kappa^2 (\epsilon^2 - \delta^2)} \right]. \quad (19)$$

With the aid of (19) we distinguish the various ranges of operation, which may be expressed in terms of the ellipticity parameter  $\delta$ . The case  $\delta=0$  corresponds to circular symmetry. For  $\delta>0$  but sufficiently small that both  $\lambda_+$  and  $\lambda_-$  are negative, two elliptically polarized modes can propagate, and the structure exhibits the properties of an elliptical Faraday rotator. At  $\delta=\epsilon$  one of the modes, namely  $n(z)$  in the notation of (13), goes into cutoff; the range of  $\delta$  appropriate to the phase shifter begins at this point. The device is observed to produce phase shift most copiously under the condition that  $\delta$  exceeds  $\epsilon$  only slightly; hence the term  $(\epsilon^2 - \delta^2)\kappa^2$  in (19) is ordinarily small and negative. This means that the off-diagonal component  $\kappa$  of the Polder tensor has lost the central role it plays, for example, in the Faraday rotator: the forms of the  $x$ - and  $y$ -polarized components of the radiation have become so dissimilar because of the ellipticity of the structure that coupling between them has become a secondary effect. Neglecting the term in  $\kappa$  for the moment, we obtain as the eigenvalue for the propagating mode  $p(z)$  the simple form

$$\lambda_{\pm} \cong -\omega^2 \mu \epsilon_x. \quad (20)$$

The observed properties of the phase shifter are to be explained in terms of the magnetic parameter  $\mu$  and the combined dielectric and geometrical parameter  $\epsilon_x$ .

The forms of the elliptically polarized normal modes  $p(z)$  and  $n(z)$  may be deduced from  $S$ , (11). They are given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = S \begin{bmatrix} p \\ n \end{bmatrix} = \begin{bmatrix} (\sigma_- + R)p - bn \\ c p + (\sigma_- + R)n \end{bmatrix} \quad (21)$$

where

$$\begin{aligned} \sigma_+ &= -\omega^2 \mu \epsilon \\ \sigma_- &= -\omega^2 \mu \delta \\ R &= (\text{sign of } \sigma_+) \omega^2 \sqrt{\mu^2 \delta^2 + \kappa^2 (\epsilon^2 - \delta^2)} \\ b &= i \omega^2 \kappa \epsilon_y = i \omega^2 \kappa (\epsilon - \delta) \\ c &= -i \omega^2 \kappa \epsilon_x = -i \omega^2 \kappa (\epsilon + \delta). \end{aligned}$$

In Fig. 4 the modes are shown schematically in four special cases:  $\delta/\epsilon=0, 1/2, 1,$  and  $3/2$ . The values  $\mu=1$  and  $\kappa=-1/2$  have been assumed. The rectangles in the figure represent the relative magnitudes of the major and minor axes of the guide cross section. The dashed ellipse in Fig. 4(d) indicates that the mode  $n$  is in cutoff.

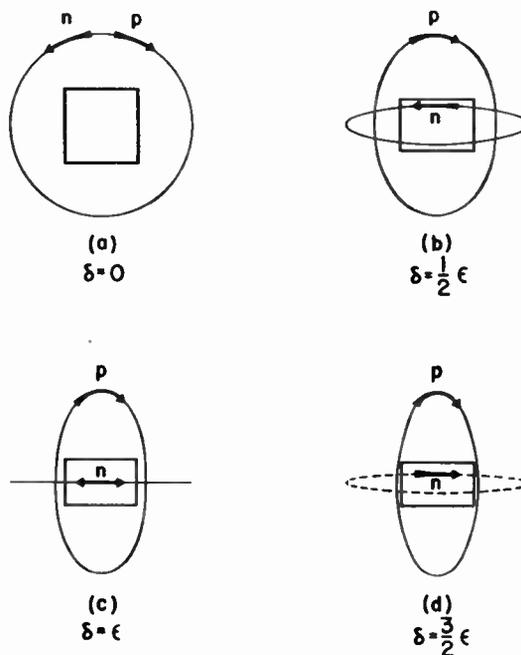


Fig. 4—The normal modes  $p(z)$  and  $n(z)$  in four special cases. The ellipses represent the electric field amplitudes, and the arrows indicate the senses of elliptic polarization. The rectangles represent schematically the degree of ellipticity. The dashed ellipse in (d) indicates that the mode  $n$  is in cutoff.

### THE ANOMALOUS VARIATION OF $\mu$ IN THE UNSATURATED FERRITE

Fig. 5 illustrates<sup>8</sup> the dependence of  $\mu$  on (internal) dc field, showing an abrupt increase at very low fields. No such effect is predicted by Polder's form<sup>5</sup> for  $\mu$  if the magnetization  $M$  and internal field  $H$  are assigned their mean values over the body of the partially magnetized

<sup>8</sup> The data of Fig. 5 were furnished by E. Kankowski, using a system designed by J. H. Rowen for measurement of the tensor permeability. See W. von Aulock and J. H. Rowen, "Measurement of dielectric and magnetic properties of ferromagnetic materials at microwave frequencies," *Bell Sys. Tech. J.*, vol. 36, pp. 427-448; March, 1957.

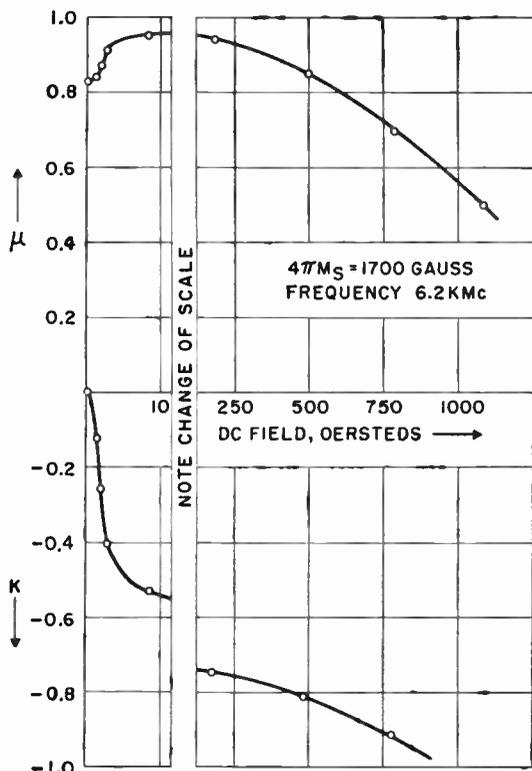


Fig. 5—The components  $\mu$  and  $\kappa$  of the permeability tensor as functions of internal dc field, at 6.2 kmc on polycrystalline MgMn ferrite of saturation magnetization  $4\pi M_s = 1700$  gauss.

ferrite. A more appropriate averaging process is to form the mean value of the product  $MH$ , thus incorporating the strong correlation<sup>6</sup> between the two factors. The result depends on the detailed crystallographic structure of the material, but in any case for the unmagnetized material  $\mu$  takes the form

$$\mu \cong 1 - \frac{\gamma^2}{\omega^2} N_0 M_s^2 \tag{22}$$

where  $\gamma$  is the gyromagnetic ratio ( $\sim 2.8$  mc per oersted for electrons),  $M_s$  is the saturation magnetization, and  $N_0$  is a mean effective demagnetizing factor which can be evaluated empirically. The value of  $N_0$  deduced from the data of Fig. 5 is 3.7, corresponding to a mean internal field of about 500 oersteds. When the field is applied, the magnetization becomes more orderly and the internal field declines at first, leading to the observed form of  $\mu$ .

THE VARIABLE-DIELECTRIC-WAVEGUIDE EFFECT

The sensitivity of  $\mu$  to the detailed structure of the ferrite leads to a variation which is in the right direction to contribute to an increase in phase on application of the dc field. The other contribution, which is represented in our schematic model by an increase in  $\epsilon_r$ , is what we have termed the variable-dielectric-waveguide effect. As  $\mu$  increases, the effectiveness of the ferrite rod as a

dielectric waveguide improves, the radiation crowds into the rod, and the wavelength decreases. Fortunately a means for estimating the magnitude of this effect exists in the form of calculations on rods of dielectric constant  $\epsilon = 10$  in circular waveguide.<sup>9</sup> The data of particular interest are the curves shown in Fig. 6 giving the guide wavelength and its rate of variation with increase in dielectric constant as functions of the rod diameter. The curves refer to the dominant mode ( $HE_{11}$ ). The amount of dielectric loading of the guide is represented by the ratio of radii of the rod and waveguide, denoted by  $c$ . The numerical calculation referred to<sup>9</sup> includes an assumed value of 0.4 for the ratio of guide radius to free-space wavelength. The solid curve gives the ratio of free-space wavelength  $\lambda_0$  to guide wavelength  $\lambda_g$  as a function of  $c$ . The dashed curve represents the rate of variation of guide wavelength with dielectric constant in the neighborhood of  $\epsilon = 10$ ; the quantity plotted, denoted by  $L$ , is the log-log derivative  $d(\log \lambda_0/\lambda_g)/d(\log \epsilon) = (\epsilon \lambda_g/\lambda_0) d(\lambda_0/\lambda_g)/d\epsilon$ .

An empirical correlation of these theoretical results with the phase shift data of Fig. 2 can be obtained by observing the wavelength on the unmagnetized rod in the phase-shifter structure and assigning an effective rod diameter which yields the same wavelength in the circular case (reading from the solid curve of Fig. 6). We find the following from some typical observations:

frequency	6000 mc	
free-space wavelength $\lambda_0$	5 cm (1.97 inches)	
waveguide dimensions	1.590 x 0.795 inch	
material	MgMn ferrite, $4\pi M_s = 1700$ gauss	
rod diameter	rod A 0.350 inch	rod B 0.400 inch
wavelength $\lambda_g$ at $H=0$	1.89 inches	1.52 inches
$\lambda_0/\lambda_g$	1.04	1.29
loading ratio $c$	0.25	0.27

Note that the amount of ferrite filling is sufficient to reduce the guide wavelength below that for free space; that is,  $\lambda_0/\lambda_g > 1$ . The condition  $\lambda_0/\lambda_g = 1$  is a critical point in the theory of shielded dielectric waveguide at which the dielectric rod first takes over from the shield the dominating influence in confining the radiation. Now with the values of  $c$  thus deduced, we read in Fig. 6 the corresponding values of the derivative  $L$ :

rod	rod A	rod B
$L$ , predicted	0.85	1.06

These values of  $L$  constitute a predicted value of the rate of increase in phase of the structure with increase in the effective dielectric constant of the rod, deduced from the theory with the aid of the observed wavelength in the unmagnetized state. In order to compare the prediction with an observed rate of phase shift, we must make the following, somewhat questionable, assumption: that the product  $\mu\epsilon$  in the present problem

<sup>9</sup> H. Seidel, unpublished work on shielded dielectric-rod waveguide, with computations by M. J. Brannon; 1954.

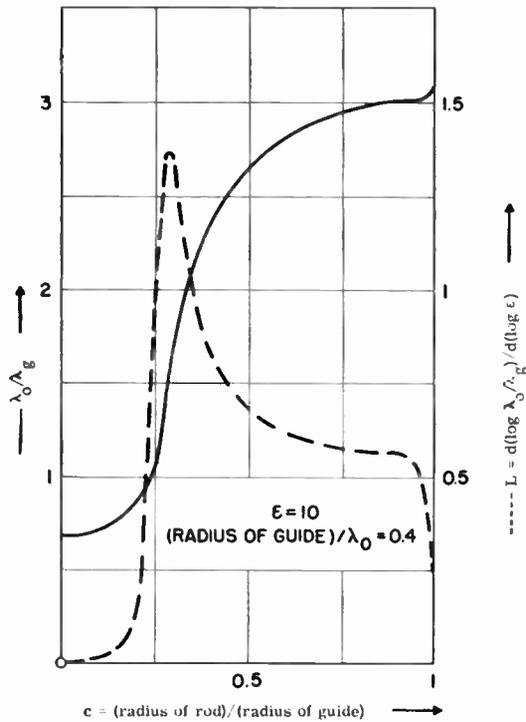


Fig. 6—Guide wavelength and its variation with dielectric constant in shielded dielectric-rod waveguide.<sup>9</sup>

can be identified with the parameter  $\epsilon$  in the dielectric-waveguide theory. Although the assumption is strictly justifiable only in the case of plane waves, to which the dielectric-rod considerations are obviously inapplicable, nevertheless, it is no worse than consistent with the approximate nature of our development; no worse, certainly, than the correlation established above between rectangular and circular guide. The forms of the (exact) characteristic equations for shielded and unshielded dielectric waveguide<sup>7</sup> suggest that the approximation is reasonable, especially for the large value of  $\epsilon$  with which we are concerned.

The observed variation of phase (Fig. 2) and of  $\mu$  (Fig. 5), as functions of the internal dc field (that is, adjusted for demagnetizing effects) have more or less uniform slopes over the range of internal fields in which the most rapid phase change occurs. From these slopes, we deduce representative values of the rate of change of guide wavelength with  $\mu$  and thence empirical values for the logarithmic derivative  $L$ , to be compared with the predicted values noted in the previous paragraph. Denoting phase shift per unit length by  $\phi$  and internal dc field by  $H$ , we obtain

$d\mu/dH$		0.107 oersted <sup>-1</sup>
rod	rod A	rod B
$d\phi/dH$	14	24° inches <sup>-1</sup> oersted <sup>-1</sup>
$L$ , empirical	0.69	0.95
$L$ , predicted	0.85	1.06

Comparing these empirical values of  $L$  with those predicted above, we note that the agreement is reasonable,

considering the many approximations involved. It is illuminating to observe that the above comparison of values of  $L$  shows the amount of phase shift available in the device (which appeared at the outset to be unaccountably large) to be actually somewhat less than predicted. We may not be justified in putting a physical interpretation on these small differences at this rudimentary stage in the theory, but it is a fact that the term  $\kappa^2(\epsilon^2 - \delta^2)$  in the propagation constant  $\gamma_+$ , which we have neglected [see (19) and (20)], does tend to reduce the phase shift. The two lower curves in Fig. 2 show the reduction which results when the difference  $|\epsilon^2 - \delta^2|$  is increased by reducing the height of the guide.

### SCATTERING IN PHASE SHIFTER AND ROTATOR STRUCTURES

Simple qualitative expressions for the scattering by the structure in the two cases  $\delta < \epsilon$  (elliptical Faraday rotator) and  $\delta > \epsilon$  (phase shifter) may be derived in the following way. We describe the input and output planes of the loaded section by a phenomenological scattering matrix of appropriate form and solve for the steady state under excitation by a wave of unit amplitude incident from one direction. The calculation is a straightforward generalization of the analysis of the interference effect in the Faraday rotator.<sup>4</sup> In the phase shifter case, where the structure supports only a single propagating mode, associated with the propagation constant  $\gamma_+$ , we are justified in neglecting the evanescent waves associated with  $\gamma_-$  provided they are attenuated appreciably in a distance which is short compared to the length of the structure. Observations of the frequency dependence of transmission show that this condition is fulfilled everywhere except very close to the limiting case  $\delta = \epsilon$ . With this simplification, the matrices  $U_{\pm}$  of (15) reduce to

$$U_{\pm} = e^{\pm\gamma_+ z} \begin{bmatrix} \tau_+ & \frac{b}{2R} \\ \frac{c}{2R} & \tau_- \end{bmatrix} \quad (23)$$

Then the scattering calculation leads to the following expression for the amplitude  $E$  of the transmitted wave:

$$E = \left[ \frac{s' s \tau_+}{1 - (r' \tau_+ + t \tau_-) e^{-2\gamma_+ d}} \right] e^{-\gamma_+ d} \quad (24)$$

where  $d$  is the length of the structure and  $s, s', r', t$  are scattering coefficients referring to the terminal planes. Of these,  $s, s', r'$  define respectively inward transmission, outward transmission, and internal reflection of radiation polarized in the  $x$  direction (the polarization for propagation in empty guide), and  $t$  defines internal reflection of radiation polarized in the  $y$  direction. The resonance described by the denominator in (24) accounts for the undulations in transmission discussed earlier

when reasonable values are assigned to the coefficients. Of the two scattering coefficients,  $r'$  and  $t$ , appearing in the denominator,  $r'$  can be made arbitrarily small by the addition of matching elements, as mentioned earlier, but the term containing  $t$  represents a residual mismatch which cannot be removed by this means, since it describes the full reflection of cross-polarized radiation at the ends of the loaded section.

In the rotator case, the presence of two propagating modes complicates the problem; the most conspicuous feature is the appearance of the interference effect<sup>4</sup> which produces violent, frequency-dependent fluctuations in transmission. The matrices  $U_{\pm}$  become

$$U_{\pm} = e^{\pm i\delta} \begin{bmatrix} \cos \rho z \mp i \frac{\sigma_{-}}{R} \sin \rho z & \pm i \frac{b}{R} \sin \rho z \\ \pm i \frac{c}{R} \sin \rho z & \cos \rho z \pm i \frac{\sigma_{-}}{R} \sin \rho z \end{bmatrix} \quad (25)$$

where

$$\gamma_{\pm} = i\beta_{\pm},$$

$$\beta = \frac{1}{2}(\beta_{-} + \beta_{+}), \quad \rho = \frac{1}{2}(\beta_{-} - \beta_{+}).$$

The transmission  $E$  now takes the form

$$E = \frac{1}{\Delta} s' s e^{-i\delta d} \left[ (1 - t^2 e^{-2i\delta d}) \cos \rho d + \frac{i\sigma_{-}}{R} (1 + t^2 e^{-2i\delta d}) \sin \rho d \right] \quad (26)$$

where the denominator  $\Delta$  is a long expression describing coupled resonances in the  $x$  and  $y$  polarizations. If the structure is well matched at zero applied field, the effect of  $\Delta$  is secondary. The factors  $(1 \pm t^2 e^{-2i\delta d})$  describe the interference effect.

An illuminating picture of the two cases (24) and (26) emerges when the structure is viewed over a range of frequencies including the point where  $\delta$  and  $\epsilon$  are equal; that is, where the second mode first breaks into propagation. Examples of such a picture as presented by a swept-frequency test circuit are shown in Fig. 7. The quantity displayed is  $|E|^2$  over the frequency range 5730 to 6310 mc. The structure is composed of a ferrite rod of saturation magnetization  $4\pi M_s = 1700$  gauss, 0.400 inch diameter  $\times$  5 inches long, mounted in standard WR-159 guide.

Fig. 7(a) shows the transmission at zero magnetization. Only the broad undulations in transmission are visible. The insertion loss varies between 0.8 and 1.3 db over the band. On application of the dc field, a number of additional features abruptly appear, as shown in Fig. 7(b), where the applied field is 14 oersteds. The transitional point  $T$  at which  $\delta = \epsilon$  occurs at 6122 mc. On the low frequency (phase shifter) side of  $T$  the undulations persist, with some increase in amplitude and shortening of "wavelength." [For clarity, the undulations have been intentionally made rather large by

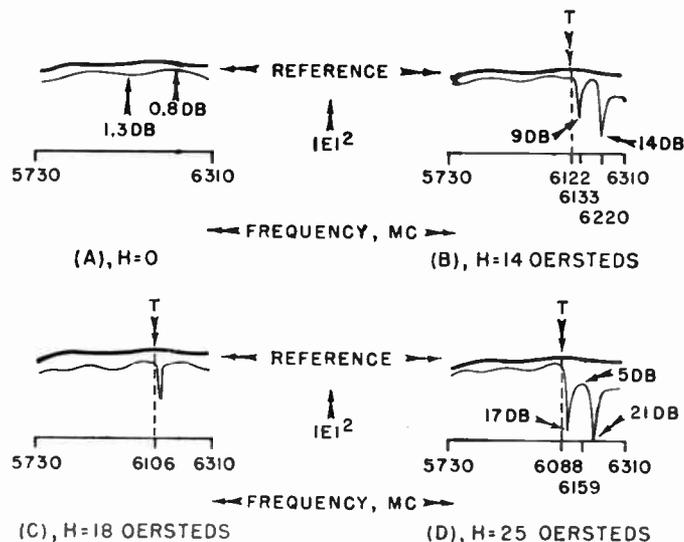


Fig. 7—Swept-frequency presentation of the transmission  $|E|^2$  of the Reggia-Spencer phase shifter, showing the transition from the phase shift region at frequencies less than  $T$  to the rotator region at frequencies greater than  $T$ . The insertion loss values, in decibels, are those referred to in the text.

using a somewhat oversized rod with flat ends. The decrease in "wavelength" on increase of the applied field is, of course, a measure of the increase in phase, as indicated by (24).] The interference "spikes" are visible on the high-frequency (rotator) side of  $T$ : one very close to  $T$ , at 6133 mc, and the next at 6220 mc. With further increase in applied field, the spikes alternately disappear [Fig. 7(c), 18 oersteds] and reappear [Fig. 7(d), 25 oersteds] as the Faraday rotation [measured by  $\rho d$  in the notation of (25)] passes through successive multiples of  $90^\circ$ . The motion of the point  $T$  on the frequency scale reflects the variable-dielectric-waveguide effect: as the radiation concentrates in the ferrite rod, the constraint imposed by the waveguide walls is gradually relieved, and cutoff for the mode  $n(z)$  moves to lower frequencies. In Fig. 7(b), 7(c), and 7(d), the point  $T$  is at 6122, 6106, and 6088 mc, respectively. On account of the ellipticity of the structure, the insertion loss in the rotation regions between the interference spikes is limited to about 5 db, as for example at 6159 mc in Fig. 7(d). The spikes themselves reach successively higher insertion loss values at each reappearance, thereby reflecting the anticipated diminution of "low field loss" effects;<sup>6</sup> the two spikes in Fig. 7(b) reach 9 and 14 db, but those in Fig. 7(d) reach 17 and 21 db.

### CONCLUSIONS

The correspondence between the highly simplified phenomenological model and the behavior of the Reggia-Spencer phase shifter serves to explain the *modus operandi* of the device. It also accounts for a number of the finer details in the observations. Further systematic study of the structure will enable us to establish numerical correlations between the phenomenological parameters, such as  $\delta$  and  $\epsilon$ , and the geometrical details, so

that in subsequent development work, we shall be in an advantageous position to interpret the data, distinguish between significant and secondary effects, etc. Ultimately, if necessary, the composite cross section of the structure may be subjected to analysis of a higher degree of approximation or to exact numerical treatment in the practical cases of greatest interest. The present analysis of the propagating and nonpropagating modes sheds some light on the question of Faraday rotation subject to elliptic waveguide symmetry. It also suggests ways of coupling into and out of the modes for the purpose of producing other novel effects.

#### ACKNOWLEDGMENT

The author wishes to acknowledge the assistance of several colleagues. R. W. Judkins and C. E. Barnes performed the measurements on the phase shifter, and E. Kankowski furnished the permeability data shown in Fig. 5. H. Seidel offered some valuable comments on the work and also made available his notes on the dielectric-waveguide problem together with the results of Miss Brannon's computations. W. von Aulock contributed the results of his experience in the form of several stimulating discussions, as well as his own experimental data.

## A New Look at the Phase-Locked Oscillator\*

HAROLD T. McALEER†, ASSOCIATE MEMBER, IRE

**Summary**—The uses of phase-locked oscillators are briefly reviewed. A simple automatic-phase-control (APC) system is analyzed as a servomechanism analog. Three major characteristics of the system are considered: the lock range, the capture range, and the filter bandwidth. The lock range is the total drift in the unlocked oscillator frequency which can be exactly compensated by the locked system. The capture range is the largest unlocked frequency difference at which synchronization, or lock-in, will occur. The filter bandwidth of the system expresses the performance of the system as a low-pass filter with respect to FM noise components existing in the input to the system and as a high-pass filter with respect to FM noise components generated within the output oscillator.

The mutual interdependence of these characteristics and the various quantities affecting each one are discussed. The conditions for stable operation of the system are established. The unlocked, locking-in, and locked conditions of operation and the effects of the low-pass filter are discussed. Simple design criteria are established.

#### INTRODUCTION

THE phase-locked oscillator system has been analyzed many times in the past.<sup>1-7</sup> Most of these analyses have been in classical mathematical

terms. In this paper the author is attempting a new approach to the problem, utilizing concepts which are well understood by people familiar with servomechanism and feedback system design. It is hoped that this approach will produce a clearer understanding of the system, both qualitative and quantitative.

#### USES OF PHASE-LOCKED OSCILLATORS

There are many varied uses for a phase-locked oscillator or automatic-phase-control (APC) system as it is sometimes called. An APC system can be used in a receiver to increase the power level and attenuate the noise of a weak FM signal. A similar system, involving somewhat different design parameters, can be used to reduce the jitter or frequency noise of a high-powered oscillator. In the field of frequency measurement and synthesis, in particular, the APC system is extremely useful.

Many frequency measurement and synthesis systems involve the generation and selection of a single frequency signal. Because of the generation process, a poor signal-to-noise ratio (SNR) results—the noise in this case taking the form of adjacent frequency components (or sidebands) and small deviation FM or phase jitter of the desired component. A phase-locked oscillator can serve as a filter of arbitrarily narrow bandwidth (arbitrarily high  $Q$ ) for the selection of the desired signal, attenuating the unwanted components and reducing the phase jitter.

#### DESCRIPTION OF A PHASE-LOCKED OSCILLATOR SYSTEM

Fig. 1 shows the basic components of an elementary APC system. There are many ramifications of this basic system, some using frequency dividers or multipliers,

\* Original manuscript received by the IRE, December 12, 1958; revised manuscript received, March 2, 1959.

† General Radio Co., West Concord, Mass.

<sup>1</sup> R. Ley, "Phase synchronization of an oscillator, application to a continuously variable oscillator of high stability," *Ann. Radioelec. (Paris)*, vol. 13, pp. 212-233; July, 1958.

<sup>2</sup> T. J. Rey, "Effects of the Filter in Oscillator Synchronization," Lincoln Lab., M.I.T., Lexington, Mass., Tech. Rep. No. 181 (ASTIA 133854); May, 1958.

<sup>3</sup> R. Leek, "Phase-lock A.F.C. loop, tracking signals of changing frequency," *Electronic and Radio Engr.*, vol. 34, pp. 114 and 177; April and May, 1957.

<sup>4</sup> D. Richman, "Color-carrier reference phase synchronization accuracy in NTSC color television," *Proc. IRE*, vol. 42, pp. 106 and 288; January, 1954.

<sup>5</sup> W. J. Gruen, "Theory of A.F.C. synchronization," *Proc. IRE*, vol. 41, pp. 1043-1049; August, 1953.

<sup>6</sup> G. W. Preston and J. C. Tellier, "The lock-in performance of an A.F.C. circuit," *Proc. IRE*, vol. 41, pp. 249-251; February, 1953.

<sup>7</sup> E. Labin, "Theory of synchronization by control of phase," *Philips Res. Rep.* vol. 4, pp. 291-315; August, 1949.

and some using offset comparison frequencies; but we shall restrict our attention to the simple system shown here. No great generality will be lost by this procedure. The system contains: an oscillator with a nominal frequency equal to the desired output frequency; some form of reactance modulator or other means for voltage control of the oscillator frequency; a phase detector which compares the outputs of the oscillator and the reference source; and a low-pass filter which filters the output voltage of the phase detector before it is applied to the reactance modulator.

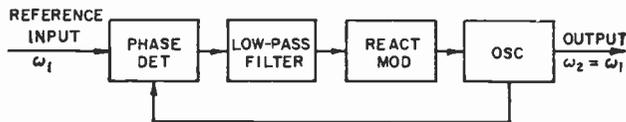


Fig. 1—Block diagram: automatic phase-control system.

The operation of the system can be understood qualitatively by assuming that the oscillator frequency is equal to that of the reference. The phase-detector output is then a dc voltage dependent on the phase difference between the output signals of the oscillator and the reference. This voltage is applied through the low-pass filter to the reactance modulator and thereby governs the oscillator frequency. If the oscillator frequency tends to change, this attempted change is first felt as a phase-difference change in the phase detector. This produces a change in phase-detector output voltage which acts to hold the oscillator frequency constant. As the oscillator drifts, its output phase, relative to that of the reference, will drift, but its average frequency will remain fixed. The system operates exactly like a positional servomechanism wherein, for constant input position, the output position is exactly equal to the input, with zero steady-state error.<sup>8</sup> To understand the lock-in performance of the system and its quantitative behavior with respect to noise and drift requires a closer look.

SERVO ANALYSIS OF A PHASE-LOCKED OSCILLATOR

A block diagram of the simple phase-locked oscillator system, which lends itself more readily to analysis, is shown in Fig. 2. This block diagram, incidentally, is identical to that for a positional servo. The circled  $\Sigma$ 's represent summation points where various signal variables (voltages or frequencies) are combined. Frequency noise (incidental FM or jitter) contributed by the reference source is represented by  $N_1$ . The  $1/S$  term represents the integration of frequency difference to phase difference which occurs in the phase detector. It is this integration which gives the system its unique properties and distinguishes it from the more familiar automatic-

frequency-control (AFC) system. In an AFC system, the frequency of the oscillator is compared to a reference frequency. For example, the resonant frequency of a passive circuit, and the frequency difference—not phase difference—is used to generate a signal which tends to reduce the frequency difference. Such a system requires a small, but finite, error of the controlled variable (the output frequency) in order to operate. The phase-lock system, on the other hand, requires no steady-state error of the controlled variable, but instead utilizes an error in the *integral* of the controlled variable, *i.e.*, an error in phase difference. The gain of the phase detector in volts per radian is represented by  $K_1$ . The frequency characteristic of the low-pass filter is indicated by  $F(S)$ , and  $K_2$  represents the gain of the reactance-modulator-oscillator combination in radians per second, per volt.  $N_2$  represents noise voltage at the input to the reactance modulator, and  $\Omega$  includes both the detuning of the oscillator and frequency noise in its unlocked output.

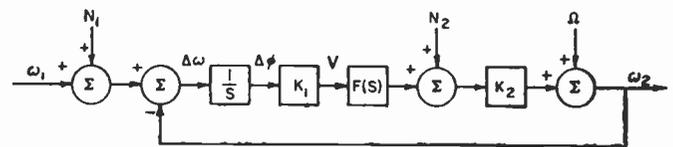


Fig. 2—Servo block diagram.

Lock Range

There are several quantities of interest for this system. Among the most important are the lock range, the capture range, and the filter bandwidth. From a knowledge of these quantities, much of the behavior of the system can be calculated, such as the static phase error, the response to an input transient, etc. The lock range is the total drift in unlocked output frequency which can be exactly compensated by the system. For a multiplier or divider system, the lock range is usually chosen so that all expected drifts in the oscillator frequency due to component changes and ambient temperature variations can be compensated as the phase-detector output traverses its entire range. From Fig. 2 it can be seen that the lock range is equal to the range of the reactance-modulator-oscillator combination caused by the maximum voltage range of the phase detector. For a linear reactance modulator and a linear phase detector operating over a range of  $\pi$  radians, the lock range would equal  $\pi K_1 K_2$  radians per second. That is, a phase-difference change of  $\pi$  radians produces a voltage change of  $\pi K_1$  volts out of the phase detector, which in turn produces a change of  $\pi K_1 K_2$  radians per second in the frequency of the reactance-modulator-oscillator combination. For a linear reactance modulator and a typical nonlinear balanced phase detector, however, which has a maximum output swing of  $2K_1$  volts, the lock range would be  $2K_1 K_2$  radians per second.

A given total lock range can be achieved with a large  $K_1$  and a small  $K_2$  or vice versa. Whichever gain should

<sup>8</sup> H. Chestnut and R. Mayer, "Servomechanisms and Regulating System Design," John Wiley and Sons, Inc., New York, N. Y.; 1950 and 1955.

be made larger depends on the noise susceptibilities of the various components. Since, in a practical system, the input to the reactance modulator is usually the most noise-sensitive point in the system, a large  $K_1$  and a small  $K_2$  are desirable. This is achieved with a high-output phase detector and a "stiff" reactance-modulator-oscillator combination. Indeed, a "stiff" oscillator offers many advantages in a system for removing phase modulation from the reference signal. An oscillator whose frequency is hard to pull usually has good long and short-term stability and thus requires a smaller lock range and produces a cleaner output than a "loose" oscillator. For these reasons, a quartz crystal oscillator is sometimes used at the output of a phase-locked multiplier.

*Capture Range*

The capture range is the largest unlocked frequency difference at which the system will lock-in. This capture range cannot be larger than the lock range, but it can be smaller. To gain a clear understanding of the operation of the system, let us ignore for a time the low-pass filter. That is, let us assume that the cutoff frequency,  $\omega_f$ , of this filter is greater than certain other frequencies of interest. The reason for this will be more apparent later. *For this condition the capture range is equal to the lock range.*

Fig. 3 is a diagram which aids in obtaining a physical picture of the operation of the system—both for the locked and the unlocked conditions. Oscillator frequency is plotted along the vertical axis, and phase difference is plotted along the horizontal axis. The curves represent the frequency of the oscillator as a function of oscillator-reference phase difference and repeat for every integral multiple of  $2\pi$ . There exists a family of these curves, each one corresponding to a different rough tuning of the oscillator.

of these intersections is stable, the other intersection corresponding to a condition of positive rather than negative feedback. The stable intersection is indicated in Fig. 3. This intersection may be called the locked operating point and indicates the phase difference required to maintain lock (the static phase error). If the rough tuning is varied (or drifts) through  $\Omega_3$  to  $\Omega_2$ , the oscillator will remain locked and the operating point will move toward one end of the stable region. If the rough tuning is varied still further, say to  $\Omega_1$ , the oscillator will unlock and the operating point will move along the  $\Omega_1$  curve in the direction indicated by the arrows, producing continuous frequency modulation of the oscillator.

Now let us reverse the procedure. Assume that the oscillator tuning has been brought to  $\Omega_1$  from some far-distant value. Since the  $\Omega_1$  curve does not intersect the  $\omega_1$  line, locking is impossible and, as mentioned before, the operating point moves along the  $\Omega_1$  line. The nature of this motion can be understood qualitatively by the following physical argument. Since the frequencies of the oscillator and reference are not equal, the output of the phase detector is a beat voltage. The "frequency" of this beat will equal the difference between the oscillator and reference frequencies. When the operating point on the  $\Omega_1$  curve is in the region furthest removed from  $\omega_1$ , the beat frequency is high and the operating point moves rapidly. When the operating point is in the region near  $\omega_1$ , the beat frequency is low and the operating point moves slowly. The oscillator frequency "hesitates" for a time as it nears the reference frequency. Then it sweeps rapidly away and back again. Appendix 11 shows that, under these conditions, the output voltage of the phase detector is a series of back-to-back exponentials, with flat regions as the operating point hesitates near  $\omega_1$  and sharp cusps as the operating point moves rapidly down one side of the phase-detector characteristic and up the other. As the tuning is changed from  $\Omega_1$  toward  $\Omega_2$ , the hesitation lasts longer and longer until, at  $\Omega_2$ , the operating point stops, right at the edge of the lock range, and the oscillator is locked.

For the case considered, *i.e.*, neglecting the low-pass filter, it can be shown that the capture range is equal to the lock range.<sup>7</sup> That is, if the oscillator frequency vs phase-difference line intersects the desired frequency, the oscillator will lock at the stable point of intersection.

*Filter Bandwidth*

The filter bandwidth of the system represents its behavior as a "frequency noise" or "jitter" filter. This can be best expressed in terms of a cutoff frequency  $\omega_c$ . It can be seen in Fig. 2 that the system behaves as a low-pass filter with respect to input noise  $N_1$ . That is, components of  $N_1$  with rates below  $\omega_c$ , *i.e.*, slowly varying changes in the reference frequency, appear directly in the output, whereas components with rates above  $\omega_c$  are attenuated. Conversely, the system behaves as a high-pass filter with respect to internal noise  $N_2$ . That

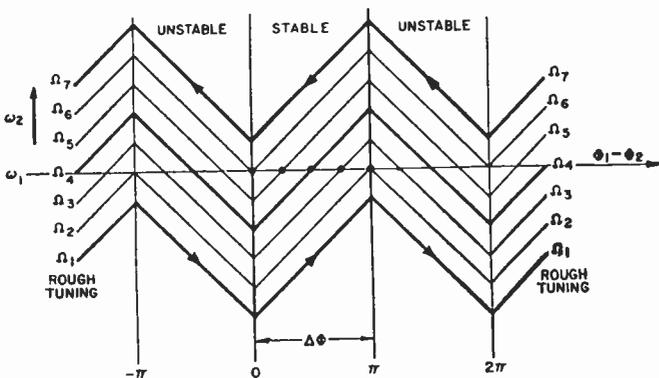


Fig. 3—Composite locking diagram.

In order for locking to occur, the rough tuning must be such that the oscillator frequency curve intersects the  $\omega_1$  line (*i.e.*,  $\omega_2 = \omega_1$ ). Assume that  $\Omega_4$  in Fig. 3 represents the oscillator rough tuning. There are two possible intersections of the  $\Omega_4$  curve with the  $\omega_1$  line. Only one

is, components of  $N_2$  with rates below  $\omega_c$  are attenuated in the output, whereas components with rates above  $\omega_c$  appear directly. This topic is discussed further in Appendix I.

In a narrow-band system, the output oscillator does not follow rapid excursions of the reference frequency such as rapid FM or jitter, but any internally generated rapid excursions appear directly. In a wide-band system, on the other hand, the output oscillator follows rapid excursions of the reference frequency, but internally generated rapid excursions are attenuated. Depending on the purposes of the system, whichever noise source is most objectionable determines whether  $\omega_c$  should be made small (narrow band) or large (wide band). For example, if the purpose of the system is to select the long-term average of the input frequency and remove incidental FM or jitter existing in the reference, an inherently clean oscillator (such as a crystal oscillator) should be chosen and  $\omega_c$  should be made as small as possible. On the other hand, if the purpose of the system is to reproduce the frequency excursions of the input<sup>3</sup> or to remove incidental FM in the output oscillator (such as a klystron oscillator),  $\omega_c$  should be made as high as possible. *However, one is not free to adjust the bandwidth without, at the same time, affecting the lock range or the capture range.*

is equal to  $\pi K_1 K_2$ , the capture range is equal to the lock range, and the cutoff frequency  $\omega_c$  is equal to  $K_1 K_2$ . The static phase-error constant is  $1/K_1 K_2$  radians phase difference for each radian per second attempted frequency difference. Lock range, capture range, and bandwidth all increase in proportion to  $K_1 K_2$ . Fortunately, this is often advantageous. A system for locking a klystron oscillator, for example, should have a large  $K_1 K_2$  to accommodate wide drift in the klystron frequency. A large  $K_1 K_2$  product automatically produces a desirable wide capture range and the wide bandwidth necessary for noise attenuation. Similarly, designing a system with small  $K_1 K_2$  to accommodate a crystal oscillator automatically produces a small capture range and a narrow bandwidth.

*Effect of the Low-Pass Filter*

Although the system bandwidth cannot be made greater than  $1/\pi$  times the lock range, one might think that the bandwidth could be reduced by making the cutoff frequency,  $\omega_f$ , of the low-pass filter lower than  $K_1 K_2$  but one has to be a little careful about this. Several changes occur as the filter cutoff frequency,  $\omega_f$ , is lowered. The phase shift at the frequency at which the loop gain is unity gets closer to  $\pi(180^\circ)$ , and the system transient response to a step input develops underdamped ringing.<sup>9</sup> Another important happening is a reduction of the capture range. The unlocked oscillator frequency must be brought nearer to the reference frequency than indicated by the lock range, for locking to occur. A hysteresis effect exists. For this case it can be shown that the reduction in capture range is approximately equal to the reduction in system bandwidth (unity-gain frequency) caused by a lowering of  $\omega_f$ .<sup>1,2</sup>

A better method of narrowing the bandwidth involves the use of a lag (integral-compensation) network as the low-pass filter.<sup>8</sup> The effect of this network is shown in the dashed curve of Fig. 4. Such a network also reduces the capture range, but not as drastically as the simple RC filter discussed above. It can be shown that the use of a lag network with constants chosen to insure a minimum noise bandwidth (for the chosen system bandwidth) will produce a reduction in capture range approximately equal to the square root of the reduction in system bandwidth.<sup>2,5</sup> (See Appendix I.)

A reduction in capture range can possibly be tolerated if the operator of the system is required to lock the oscillator before use or if a nonlinear lag network<sup>1</sup> is used. However, such a system may be only conditionally stable. If the oscillator drifts outside the capture range—still remaining within the lock range—a small perturbation may cause it to unlock. The system is also conditionally stable in the sense that any nonlinearity or shift in  $K_1$  or  $K_2$  can cause the open-loop unity-gain frequency to move to an unstable region.

It is usually desirable to have the capture range as

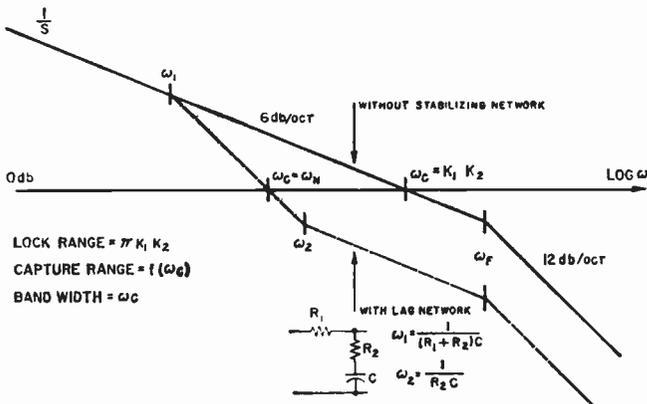


Fig. 4—Open loop gain.

Fig. 4 shows a plot of the open-loop gain of the system in decibels vs logarithmic frequency. Because of the  $1/S$  integration term, the loop gain decreases with frequency at a rate of 6 db per octave, intersecting the zero-decibel (unity gain) line at a frequency  $\omega_c = K_1 K_2$ . *This is the cutoff frequency of the closed-loop system.* That is, the transfer characteristic of the system equals unity ( $\omega_2 = \omega_1$ ) for input variations with rates below  $\omega_c$  and falls off for rates above  $\omega_c$ . The low-pass filter is assumed to be a simple RC section with a cutoff frequency of  $\omega_f$  (shown greater than  $\omega_c$ ). This filter causes an increase in the rate of open-loop and closed-loop attenuation to 12 db per octave above  $\omega_f$ .

This open-loop gain-frequency diagram completely defines the characteristics of the system. The lock range

<sup>9</sup> See Appendix I.

nearly equal to the lock range as possible, to insure locking whenever the oscillator tuning is within the lock range. This requires a low-pass filter with a cutoff frequency greater than  $K_1K_2$ .

### Summary

To summarize the foregoing results: The lock range of the system in radians per second is equal to  $\pi K_1K_2$ ,  $\pi$  times the product of the gains of the phase detector and reactance-modulator-oscillator combination. The bandwidth of the system is determined by the cutoff frequency,  $\omega_c$ , the frequency at which the open-loop gain falls to unity. This is equal to  $K_1K_2$  in the absence of a low-pass filter. The capture range of the system is a function of the cutoff frequency,  $\omega_c$ . With no low-pass filter, the capture range is equal to the lock range. With a simple RC filter reducing  $\omega_c$ , the capture range is reduced by approximately the same factor. With an optimized lag network reducing  $\omega_c$ , the capture range is reduced by approximately the square root of the bandwidth reduction.

### DESIGN PROCEDURE

The usual procedure is to design the best oscillator economically feasible to produce the desired output frequency. The reactance modulator may be a reactance tube, a voltage-variable inductor or capacitor, or the like. The phase detector may be a balanced type to reduce amplitude sensitivity or a simple tuned-circuit-diode rectifier if the amplitudes involved are constant. Use of a limiter is desirable since amplitude variations in the phase detector can result in frequency modulation of the output oscillator. The product  $K_1K_2$ , preferably with large  $K_1$  and small  $K_2$ , should be chosen to insure locking for all expected drifts in oscillator or reference frequency. For a maximum capture range, the low-pass filter should have a cutoff frequency slightly greater than  $K_1K_2$ . If a wider bandwidth is desired to attenuate internal noise, the lock range may be made greater than necessary. This will also reduce the phase error required to correct a given frequency error.

If a small static phase error and a narrow bandwidth are desired and a reduction in capture range can be tolerated, a lag network should be used.<sup>9</sup>

It should be emphasized that in actual practice the parameters,  $K_1$  and  $K_2$ , may be variable. The common balanced phase detector, for example, has a sensitivity ( $K_1$ ) which depends on the phase-difference operating point. Similarly, most reactance-tube controlled oscillators have sensitivities ( $K_2$ ) which depend on the reactance-tube bias, the *O*-bias (high gm) region being the more sensitive. Therefore, the bandwidth and stability may vary with the lock operating point. All possible values of  $K_1$  and  $K_2$  should be considered in the design.

Ordinarily  $K_1$ ,  $K_2$ , and the low-pass filter characteristic are chosen for the best compromise with respect to lock range, capture range, static phase error, transient response to input signal variations, input noise rejection,

internal noise rejection, and economy, for the particular application. Often a great deal of subjective experimenting is required. The criteria mentioned above should serve as a useful guide, however.

### CONCLUSION

In conclusion, the phase-locked oscillator is a versatile device. It can be used to increase the power level and SNR of a weak, jittery signal or to reduce short-term frequency excursions of either the reference source or the oscillator itself. The techniques described have been used successfully to multiply standard frequencies from 100 kc to 1 mc, 5 mc, 10 mc, and 100 mc with locked crystal oscillators, and to 1000 mc with a locked klystron oscillator.<sup>10,11</sup>

### APPENDIX I

#### Analysis of the Low-Pass Filter Effect

*System Transfer Function:* In Fig. 2, let the transmission of the low-pass filter be:

$$F(s) = \frac{1 + t_2s}{1 + t_1s} \quad (1)$$

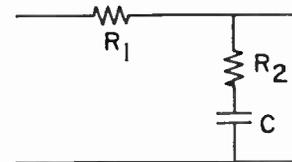


Fig. 5—Lag-network.

This corresponds to the use of a lag network shown in Fig. 5, where

$$t_1 = (R_1 + R_2)C = \text{the time constant of the lag break}$$

$$t_2 = R_2C = \text{the time constant of the lead break.}$$

The transfer function of the system can be calculated to be:

$$\frac{\omega_2(s)}{\omega_1(s)} = \frac{K(1 + t_2s)}{t_1[s^2 + s(1/t_1 + Kt_2/t_1) + (K/t_1)]} \quad (2)$$

where  $K = K_1K_2$ .

The denominator is of the general form

$$s^2 + 2\zeta\omega_n s + \omega_n^2,$$

where

$$\omega_n = \text{the natural resonant frequency} = \sqrt{\frac{K}{t_1}} \quad (3)$$

$$\zeta = \text{the damping ratio} = \frac{1 + Kt_2}{2\omega_n t_1} \quad (4)$$

<sup>10</sup> J. K. Clapp and F. D. Lewis, "A unique standard-frequency multiplier," 1957 IRE NATIONAL CONVENTION RECORD, pt. 5, pp. 131-136.

<sup>11</sup> F. D. Lewis, "New standard frequency multipliers," *General Radio Experimenter*, vol. 32, pp. 3-9; July, 1958.

It will be shown later that  $\omega_n$  equals  $\omega_c$ , the open-loop crossover frequency.

*Noise Bandwidth:* The response of the system to random disturbances on the input can be expressed in terms of its noise bandwidth:<sup>1,2,5</sup>

$$F_n = \int_0^\infty \left| \frac{\omega_2}{\omega_1} (j\omega) \right|^2 d\omega. \quad (5)$$

This is the area under the square of the closed-loop frequency response curve.

Eq. (2) can be substituted in (5) and the integration carried out to yield<sup>1,2,5</sup>

$$F_n = \frac{\pi\omega_n}{4\zeta} \left[ 1 + \left( 2\zeta - \frac{\omega_n}{K} \right)^2 \right], \quad (6)$$

or, in terms of  $\alpha = t_1/t_2 =$  the network ratio,

$$F_n = \frac{\pi}{2} K \frac{\alpha^2 + Kt_1}{\alpha^2 + \alpha Kt_1}. \quad (7)$$

For the minimum noise bandwidth, it can be shown that the following conditions should be satisfied:<sup>2</sup>

$$\zeta_0 = \frac{1}{2} \sqrt{1 + \left( \frac{\omega_n}{K} \right)^2}, \quad (8)$$

or, in terms of  $\alpha$ ,

$$\alpha_0 = 1 + \sqrt{1 + Kt_1} = 1 + \sqrt{1 + \left( \frac{K}{\omega_n} \right)^2}. \quad (9)$$

These equations can be manipulated to yield the following relationship between the several frequencies of interest:

$$\left( \frac{1}{t_2\omega_n} \right)^2 = 1 + \frac{2}{Kt_1} [1 + \sqrt{1 + Kt_1}]. \quad (10)$$

The left side of this equation is the square of the ratio of the network lead frequency to the resonant frequency. According to the right side, this ratio can never be less than one. This indicates that, in a system with minimum noise bandwidth and for the usual case  $K > (1/t_1)$ , the resonant frequency  $\omega_n = \sqrt{K/t_1}$  equals  $\omega_c$ , the open-loop crossover frequency, and the crossover frequency occurs in the 12-dB-per-octave region of the open-loop gain.

*Limiting Cases:* With no low-pass filter at all, the noise bandwidth becomes:

$$F_n = \frac{\pi}{2} K \quad (11)$$

This can be seen from (5) by letting  $\alpha = 1$ . Similarly, with a single-section RC filter

$$F_n = \frac{\pi}{2} K. \quad (12)$$

This can be seen by letting  $\alpha$  approach infinity in (5).

This last result is interesting since it shows that as the crossover frequency,  $\omega_c = \sqrt{K/t_1}$ , is lowered by increasing  $t_1$ , the damping ratio decreases so as to maintain the noise bandwidth constant at  $\pi/2K$ . In other words, as the crossover frequency is lowered, the peaking of the frequency response curve is increased, maintaining the area under the square of this curve constant.

With the use of the optimized lag network mentioned, the noise bandwidth varies from  $\pi/2\omega_c$  for  $\omega_c = K$  to  $\pi\omega_c$  for  $\omega_c \ll K$ . The optimum value of  $\zeta$  varies from 0.707 for  $\omega_c = K$  to 0.5 for  $\omega_c \ll K$ . Similarly, the optimum value of  $\alpha$  ranges from 2.4 for  $\omega_c = K$ , to  $K/\omega_c$  for  $\omega_c \ll K$ . This indicates that for values of  $\omega_c$  much smaller than  $K$  the lead break occurs immediately after the crossover frequency. The values are repeated in the following table:

$\frac{\omega_c}{K}$	1	$\ll 1$
$\zeta_0$	0.7	0.5
$\alpha_0$	2.4	$\frac{K}{\omega_c}$
$F_n$	$\frac{\pi}{2} \omega_c$	$\pi\omega_c$

*Capture Range:* Gruen<sup>5</sup> has shown that the capture range is

$$|\Delta\omega| \simeq \sqrt{2\zeta\omega_n K}, \quad \text{for } \omega_n \ll K \quad (13)$$

or, expressed as a "capture ratio,"

$$\left| \frac{\Delta\omega}{K} \right| = \frac{\text{capture range}}{\text{lock range}} \simeq \sqrt{2\zeta \frac{\omega_n}{K}}. \quad (14)$$

Substituting  $\zeta_0 = 0.5$  for  $\omega_n \ll K$ , this ratio becomes

$$\frac{\text{capture range}}{\text{lock range}} \simeq \sqrt{\frac{\omega_n}{K}} \quad (15)$$

for a system with an optimized lag network.

For a system with a simple RC filter ( $t_2 = 0$ ),

$$\zeta = \frac{1}{2} \frac{\omega_n}{K}. \quad (16)$$

Substituting (16) in (14) yields

$$\frac{\text{capture range}}{\text{lock range}} \simeq \frac{\omega_n}{K}. \quad (17)$$

*Capture Time:* Richman<sup>4</sup> states that the capture time depends on the initial frequency offset and the noise bandwidth in the following manner:

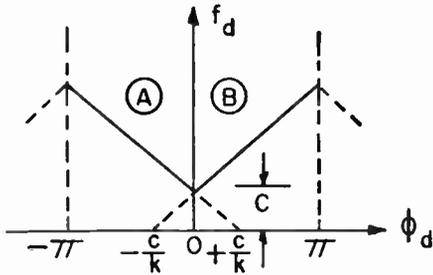


Fig. 6—Composite locking diagram.

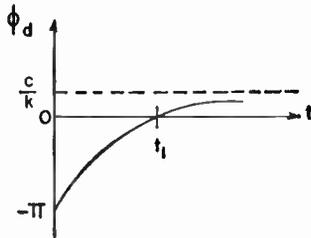


Fig. 7—Phase behavior in Region A.

$$t \simeq 4 \frac{\Delta f^2}{F_n^3}, \tag{18}$$

except near the limit of the capture range, where the capture time approaches infinity.

APPENDIX II

Pulling Behavior of an APC Loop

Fig. 6 illustrates what occurs if the rough tuning of the oscillator is such that the range of the phase-detector-reactance-modulator combination is insufficient to allow locking.

In this case,

$f_d = f - f_0 =$  instantaneous frequency difference between oscillator and reference.

$\phi_d = \phi - \phi_0 =$  instantaneous phase difference.

$c =$  minimum frequency difference allowed by device.

$k =$  slope of  $f_d$  vs  $\phi_d$  characteristic.

*Initial Conditions:* Assume that at  $t = 0$ ,  $\phi_d = -\pi$ , i.e., the operating point starts at the left extremity of Region A.

In Region A:

$$f_d = -k\phi_d + c,$$

$$f_d = (d/dt)\phi_d.$$

Solving:

$$(d/dt)\phi_d = -k\phi_d + c,$$

$$(d/dt)\phi_d + k\phi_d = c,$$

$$\phi_d = (c/k) + Ae^{-kt}.$$

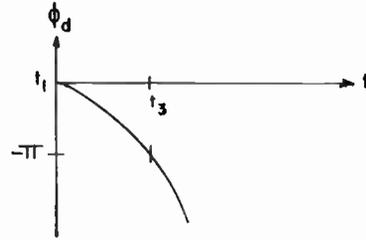


Fig. 8—Phase behavior in Region B.

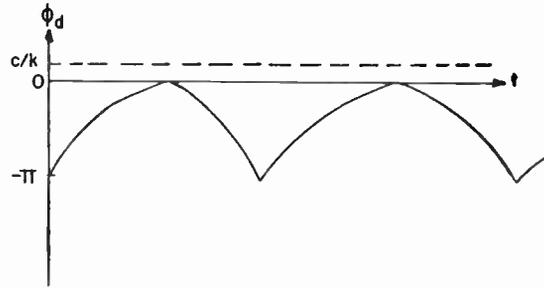


Fig. 9—Steady-state phase behavior.

Substituting initial condition:

$$-\pi = (c/k) + A,$$

$$A = -\pi - (c/k),$$

$$\phi_d = (c/k) - (\pi + c/k)e^{-kt}.$$

The time behavior of  $\phi_d$  is shown in Fig. 7.

However, at  $t = t_1$ , the operating point enters Region B with initial conditions

$$t = t_1 = t_0'$$

$$\phi_d = 0.$$

In Region B

$$f_d = k\phi_d + c,$$

$$f_d = (d/dt)\phi_d,$$

$$(d/dt)\phi_d = k\phi_d + c,$$

$$(d/dt)\phi_d - k\phi_d = c,$$

$$\phi_d = - (c/k) + Ae^{kt}.$$

Initially,

$$0 = - (c/k) + A,$$

$$A = + (c/k),$$

$$\phi_d = + (c/k)(1 - e^{kt}).$$

The time behavior of  $\phi_d$  is shown in Fig. 8.

However, at  $t_3$ , operating point reaches  $\pi$ , re-entering Region A at the proper initial conditions and the process repeats as shown in Fig. 9.

ACKNOWLEDGMENT

The author wishes to express appreciation to M. J. Fitzmorris of the General Radio Company for his assistance in the system analysis.

# Correspondence

## The Antipodal Reception of Sputnik III\*

The permanent recording of satellite signals on 20 mc, carried out at Stanford, Calif., gave some important results. Garriott and Villard, Jr.<sup>1</sup> found that during a period of several weeks, the signal was often detected at a time midway between two direct afternoon passes, while between the morning passes of the same period, practically no antipodal signal was detected. During the next period of about two months no antipodal reception could be noted at all. Practically all detected antipodal signals arrived from the southeast. The authors deduced from the constancy of the Doppler shift and from the direction of arrival that the arriving energy must have been confined to a relatively narrow cone in some unchanging direction. That was in a good agreement with the theoretical conclusions drawn by Woyk.<sup>2-4</sup>

It was inferred<sup>2</sup> that in an ionized layer surrounding the earth there must exist two critical levels along which a radio wave can perpetually propagate. Their somewhat different character is shown in Fig. 1. The

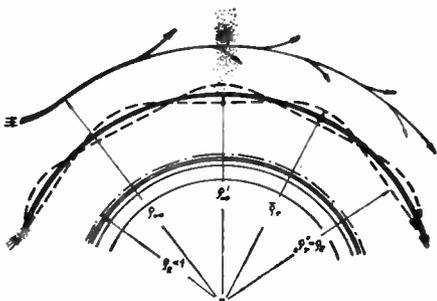


Fig. 1.

propagation along the well-known upper level,  $\rho_{\omega}$ , lying close below the maximum of electron density, is rapidly attenuated by a considerable spread of the ray pencil. On the other hand, the new kind of around-the-earth propagation along the lower critical level  $\rho_{\omega}'$  is so stable that any ray—once horizontal at this level—can never leave that level as long as the ionosphere remains perfectly spherical. The curvature of the ray there is so small that it equals the curvature of the earth. Even rays which are not

quite horizontal at the lower critical level are forced to propagate along it in small pulsations as in a waveguide while they are trapped between two reflecting levels  $\bar{\rho}_r$  and  $\rho_r''$  shown in Fig. 1. (The greater electron density above the lower critical level  $\rho_{\omega}'$  curves the ray more than at  $\rho_{\omega}$ . Consequently the ray must return to  $\rho_{\omega}'$ . Similarly the region below  $\rho_{\omega}'$  deviates the ray back to  $\rho_{\omega}'$ , as below  $\rho_{\omega}$  there are practically no more electrons. Thus the ray is almost rectilinear.) In order to bring the signal from the satellite at this critical level, and then to the earth's surface, there must exist at least two regions in which levels marked by the same electron density are not concentric to the earth's surface. Then, according to the horizontal variation of electron density, both reflecting levels  $\bar{\rho}_r$  and  $\rho_r''$  drift more or less away from each other.<sup>2,4</sup> (The around-the-earth propagation is destroyed when the upper reflecting level  $\bar{\rho}_r$  has withdrawn beyond the maximum of electron density; in this case, the signal escapes. On the other hand, when the lower reflecting level  $\rho_r''$  has withdrawn below the earth's surface, the signal can be received. For  $f \geq 3f_c$ , the ray usually escapes before having attained the earth.)

Let us next follow the best conditions for antipodal receptions.

The probability of reception grows rapidly<sup>3</sup> when the satellite approaches the antipodal point. In this singular case, the transmission can occur with the same probability over all possible great circle paths. The preferred direction of arrival, however, depends first of all upon the ionospheric conditions at the points where the ray first touches the ionosphere and where it leaves it, respectively.<sup>4</sup> The first point depends upon the altitude of the satellite, the second is usually about 2000 km from the observer.

### CONDITIONS AT THE SATELLITE SIDE OF THE RAY PATH

From the point where the ray first enters below the ionosphere, the critical frequency of the layer must increase. That is necessary because the most horizontal part of the ray pencil which has just penetrated through the ionosphere must be prevented from escaping at the next touch with it (occurring from downwards).

### IONOSPHERICAL CONDITIONS AT THE OBSERVER'S SIDE OF THE RAY PATH

At middle latitudes a signal on 20 mc can usually reach the earth's surface only when the levels containing the same electron density are more curved than the earth's surface. Then also the reflecting level is somewhat bent down so that the signal path goes to be more deviated to the earth. Such suitable conditions occur for example when the ray enters into a region where the  $F$  layer just splits into two parts or where both parts  $F_1$  and  $F_2$  still withdraw from each other.

From what has been said, no antipodal reception is probable in the direction from polar regions as in the rare polar ionosphere, the ray would have escaped before having reached middle latitudes. On the other hand, its arrival precisely from the south would presume either a transversal variation in electron density having deviated the ray from the great circle path or a high altitude of the satellite so that the most horizontal rays could enter below the ionosphere just near the rarest part of the polar regions.

Considering the conditions at the observer's side, two directions of arrival seem to be preferred: from the southwest in the forenoon and from the southeast in the afternoon. In both these cases the ray enters into a region where the  $F$  layer grows thicker. The forenoon passes from the southwest, however, will not be frequent as the ray must have passed the region where the critical frequency falls to its morning minimum, which is deep even in the equatorial zone. It is improbable that the ray would not escape there. When passing through the sunset zone the ray can be split by the turbulence or convective streams so that a multipath propagation can never be excluded; but the antipodal reception need not be destroyed by the sunset region.

According to the elimination just made, we have to expect that at Stanford the best conditions for a frequent antipodal reception occur from the southeast during Summer afternoons. They are likely to be suitable—but less frequent—from the southeast during Winter noons.

E. WOYK  
(E. CHOJKOVÁ)  
Academy of Science  
Astronomical Inst. Czech.  
Praha 12, Czechoslovakia

## Apparent Correlation Between Tropopause Height and Long-Distance Transmission Loss at 490 mc\*

A short-term experiment has been carried out in central Ontario to examine 490-mc transmissions over a 640-mile path. UHF transmissions over shorter distances have been studied widely during the past fifteen years, and although the physical process which is responsible for them is not fully understood, it is believed generally that the radio signal from the transmitter is scattered or reflected within the troposphere to ranges well beyond the optical horizon. The region in which this scattering or reflection occurs is assumed to be in the common

\* Received by the IRE, March 30, 1959.

<sup>1</sup> O. K. Garriott and O. G. Villard, Jr., "Antipodal reception of Sputnik III," Proc. IRE, vol. 46, p. 1950; December, 1958.

<sup>2</sup> E. Woyk (Chvojková), "Über den Weltumlauf der Radiostrahlen," Bull. Astro. Czech., vol. 5, pp. 104, 110; November, 1954. And, "The refraction of radio waves by a spherical ionized layer," J. Atmos. Terrest. Phys., in press.

<sup>3</sup> E. Woyk, "Investigation of the ionosphere using signals from the earth satellites," Nature, vol. 182, pp. 1362-1363; November, 1958.

<sup>4</sup> E. Woyk, "Radio waves from the earth satellite and their propagation from the antipodal point," J. Atmos. Terrest. Phys., (just presented).

\* Received by the IRE, February 16, 1959.

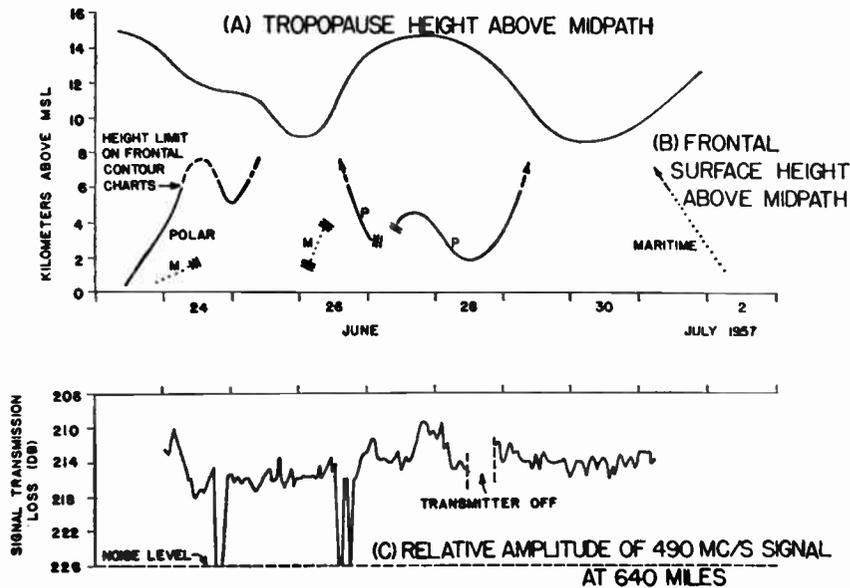


Fig. 1—Measured signal transmission and estimated tropopause height, for the 640-mile path.

volume of the troposphere, which lies above the transmitter and receiver horizons. The question now arises as to what happens to the UHF signal transmission when the receiver is so far from the transmitter that the common volume above the horizon planes lies above the troposphere, within the stratosphere. This is the case when the transmitter and receiver are more than 450 miles apart, the antennas are directed horizontally, and the antenna heights are approximately fifty feet.

It has been established recently that VHF and UHF signals can be transmitted over distances well beyond 450 miles.<sup>1,2</sup> These signals fluctuate in amplitude in a manner not unlike signals transmitted over ranges less than 450 miles. At least three theories have been proposed to explain the physical process of very long range UHF transmission,<sup>3-5</sup> but experimental information to test these theories adequately is lacking.

The experimental radio path in central Ontario was located between Ottawa and Port Arthur. A 10-kw signal was transmitted between 28-foot paraboloidal antennas, and the amplitude of the received signal was recorded continuously for a period of seven days, in June, 1957. Line (C) Fig. 1 illustrates

the fluctuations in hourly median transmission loss over the 640-mile path. It will be seen that three brief intervals occurred when the signal was below the noise level of the receiver. Outside of these intervals, the hourly median signal was at least 10 db above the receiver noise level. Closer

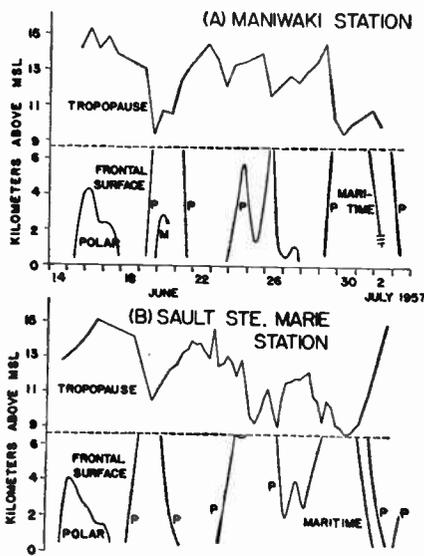


Fig. 2—Tropopause and frontal surface heights at radiosonde stations near the Ottawa-Port Arthur path.

examination indicates that the seven-day period may be divided into four sections according to the type of fluctuation in signal amplitude, although the total range of fluctuation was only about 8 db. There were two intervals when the signal was relatively weak and fluctuations were limited to 2 or 3 db (June 24th P.M. to June 26th A.M., and June 28th P.M. to July 1st A.M.), and two intervals when the signal was relatively

strong and fluctuations were deeper (June 24th A.M., and June 26th P.M. to June 28th A.M.).

The change in height of the tropopause above the radio path during this seven-day period showed an interesting correlation with signal transmission loss. Recently, it has been established that the height-variation of the tropopause follows a regular pattern in the vicinity of a weather front.<sup>7,8</sup> The tropopause height is greater on the warm-air side of the frontal junction than on the cool-air side. Fig. 2 illustrates this relationship as observed above Maniwaki (near the Ottawa end of the radio path) and above Sault Ste. Marie (near the center of the radio path). Here the minimum tropopause height was 9 km at the cool-air side of the tropopause-frontal junction, and the maximum height was 15 km above the warm air mass. The tropopause height above the center of the Ottawa-Port Arthur path was estimated from this information and from the known frontal positions above mid-path, as shown in (A) and (B) of Fig. 1. A comparison of tropopause height in (A) with the four sections of signal transmission loss plotted in (C) (Fig. 1) indicates that the periods of relatively stable transmission and higher loss were concurrent with a low tropopause, and periods of relatively unstable transmission and lower loss occurred when the tropopause was high. It should be added that a similar analysis of 490-mc transmissions over a distance of 380 miles along the same path failed to show this correlation between signal transmission loss and tropopause height above mid-path. The brief intervals of signal fadeout do not appear to be correlated with frontal position or tropopause height.

It is of further interest to compare the average signal-transmission loss as measured between Ottawa and Port Arthur with that reported from the eastern U.S.A. During the one-week interval, the loss over the 640 mile path between Ottawa and Port Arthur was 125 db greater than the free space loss. On a 618 mile path in the eastern U.S.A., the mean loss during the month of January was only 4 db less than this, for a signal frequency of 413 mc.<sup>9</sup> This transmission loss is some 20 db greater than Bullington's average transmission-loss curves indicate it should be for a 640 mile path.<sup>9</sup>

Further details on this experiment are given in a DRTE Project Report.<sup>10</sup> This work was carried out under Defense Research Board project PCC D48-28-01-03, at the Defence Research Telecommunications Establishment.

D. R. HAY  
Department of Physics  
University of Western Ontario  
London, Ontario, Can.

<sup>1</sup> K. A. Norton, P. L. Rice, and L. E. Vogler, "The use of angular distance in estimating transmission losses and fading range for propagation through a turbulent atmosphere over irregular terrain," *Proc. IRE*, vol. 43, pp. 1488-1526; June, 1955.

<sup>2</sup> J. H. Chisholm and J. F. Roche, "Measurements of Signal Levels at U.H.F. and S.H.F. Propagated by the Troposphere over Paths 100 to 618 Miles in Length," *URSI Meeting*, Washington, D. C.; May, 1956.

<sup>3</sup> F. H. Northover, "Long-distance V.H.F. fields. Refractivity profiles containing sharp layers," *Can. J. Phys.*, vol. 33, pt. 2, pp. 316-346; June, 1955.

<sup>4</sup> B. J. Starkey, "Some aircraft measurements of beyond-the-horizon propagation phenomena at 91.3 MC," *Proc. IEE (London)*, pt. B, pp. 761-763; November, 1956.

<sup>5</sup> H. G. Booker and W. E. Gordon, "The role of stratospheric scattering in radio communication," *Proc. IRE*, vol. 45, pp. 1223-1227; September, 1957.

<sup>6</sup> T. J. Carroll and R. M. Ring, "Propagation of short radio waves in a normally stratified troposphere," *Proc. IRE*, vol. 43, pp. 1384-1390; October, 1955.

<sup>7</sup> J. S. Sawyer, "Day-to-day variations in the tropopause—their causes and significance," *Geophys. Memoirs*, Meteor. Office, vol. 11, no. 92; 1954.

<sup>8</sup> B. W. Boville, W. S. Creswick, and J. J. Gillis, "A frontal-jet stream cross-section," *Tellus*, vol. 7, pp. 314-321; August, 1955.

<sup>9</sup> K. Bullington, "Radio transmission beyond the horizon in the 40-to-4000-MC band," *Proc. IRE*, vol. 41, pp. 132-135; January, 1953.

<sup>10</sup> D. R. Hay, J. W. B. Day, and L. A. Maynard, "Report on 495-MC Transmissions over Ranges up to 640 Miles," *Defence Research Telecommunications Establishment*, Ottawa, Canada, Project Report 22-0-13; July, 1958.

### Backward-Wave Oscillations in an Unloaded Waveguide\*

A number of investigators have indicated the possibility of obtaining interaction of a low-energy electron beam with an unloaded circuit by having the electrons traverse a periodic trajectory.<sup>1-3</sup> In this manner the slow-wave structure may be eliminated. Particularly if one is interested in short wavelengths, a slow-wave device has the disadvantages of a fabrication problem, matching difficulties, and the decrease of field strength away from the surface of the structure.

Consider, for example, an electron beam traversing a helical path in a dc longitudinal magnetic field,  $B$ , as illustrated in Fig. 1. Fig. 2 shows a transverse section of the guide and indicates how interaction may occur with the  $TE_{11}$  mode. The magnetic field is adjusted so that the cyclotron frequency is approximately equal to the RF frequency. This means that if an electron is retarded by the field is shown in Fig. 2(a), it will again be retarded by the field upon having rotated  $\pi$  radians as in Fig. 2(b). In this manner a synchronism condition is maintained. In addition to having synchronism, it is also necessary that electrons which are in a phase position to be accelerated do not absorb as much energy as the retarded electrons deliver. A detailed small-signal analysis of this device indicates that the RF fields perturb the electron position in a manner that is appropriate for satisfying this condition.

If the longitudinal motion of the beam is included, then the synchronism condition is slightly altered as in (1).

$$\omega = \frac{\dot{\phi}_0}{1 \pm (z_0/v_p)} \quad (1)$$

where

$\dot{\phi}_0$  = dc angular velocity (cyclotron frequency),

$\omega$  = rf frequency in radians per second,

$z_0$  = longitudinal dc velocity,

$v_p$  = phase velocity for the  $TE_{11}$  mode.

The plus sign in (1) is applicable if the waveguide mode is propagating opposite to the beam motion, and the minus applies when the mode and beam are propagating in the same direction.

An appropriate  $\omega$ - $\beta$  diagram is illustrated in Fig. 3. The solid line in Fig. 3 is the usual dispersion curve for an unloaded waveguide. The dashed lines are obtained by letting  $\phi = (\dot{\phi}_0/z_0)z$ , which expresses the dc relationship between angular and longitudinal motion. Thus, an effective propagation constant of the form  $(\pm p\dot{\phi}_0/z_0 \pm \beta)$  results, where

$$\beta = \omega/v_p$$

$p$  = an integer corresponding to the number of angular variations of the RF field.

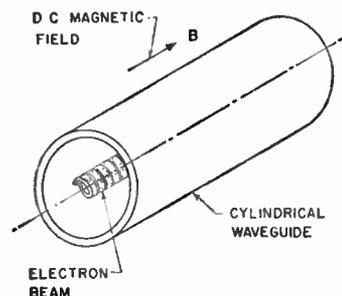


Fig. 1—The beam traverses a helical path in a dc longitudinal magnetic field and interacts with a transverse electric mode in the waveguide.

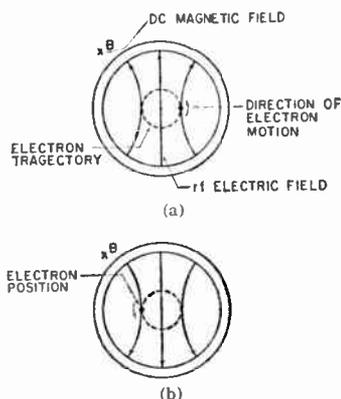


Fig. 2—Synchronism between an electron beam and the dominant waveguide mode may be obtained by having the cyclotron frequency approximately equal to the RF frequency.

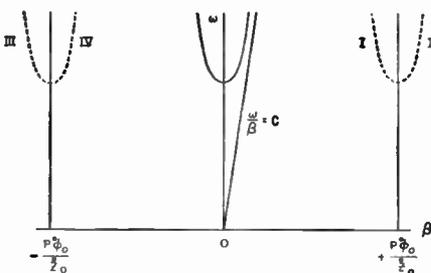


Fig. 3—Dispersion curves for the waveguide mode along a helical trajectory.

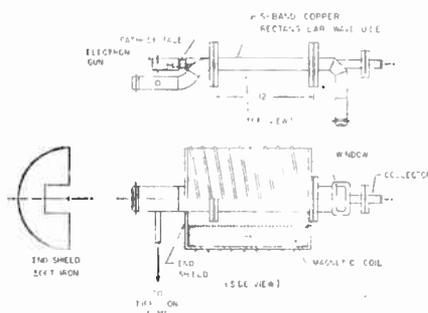


Fig. 4—Experimental setup for obtaining backward-wave oscillations in an unloaded waveguide.

For a  $TE_{11}$  mode,  $p=1$ . Curve I corresponds to a field propagating in the negative  $z$  and positive  $\phi$  directions; curve II corresponds to a field propagating in the positive  $z$  and positive  $\phi$  directions; curves III and IV are fields propagating in the negative  $\phi$  and negative and positive  $z$  directions respectively.

The apparatus illustrated in Fig. 4 was used to obtain experimental evidence for the generation of backward-wave oscillations in an unloaded waveguide. A shielded cathode emits a solid electron beam which then enters a magnetic field several times the Brillouin field. Under these conditions the electrons rotate at half the cyclotron frequency about the center line of the cathode, and at the cyclotron frequency about their axis of rotation. The tube is standard S-band rectangular waveguide which is matched at the collector end and oscillation power measured at the cathode end. Gun perveance is about  $3 \times 10^{-6}$  and the electron beam is accelerated through 1000 volts. The solenoid is 14 inches between end plates. By varying the magnetic field, oscillations were observed from 2.5 to 4.0 kmc at a power level of 0.4 watt and an efficiency of 0.5 per cent. Starting current varied between 80 and 125 ma.

It should be noted that this novel tube has an order of magnitude increase in transverse linear dimensions as compared to a slow-wave device for a comparable beam voltage. In addition, the strongest fields are in the region of the beam and the matching problem is essentially eliminated.

RICHARD H. PANTELL  
Microwave Laboratory  
W. W. Hansen Labs. of Physics  
Stanford University  
Stanford, Calif.

### The Use of Resolvers\*

Resolvers are customarily used for transforming vector coordinates with the use of auxiliary servomechanism. Not generally known is the fact that the resolver itself is quite capable of converting the rectangular components of a vector into polar representation without the benefit of any secondary instrumentation.

If the amplitudes of the carrier waves applied to the stator windings of a resolver are modulated by the rectangular components of an input vector, a stationary field will be set up at an angle equal to the vector phase  $\theta$ , with respect to the stator. There will also be induced voltages in the two rotor windings. Let one of these rotor windings be open-circuited and the other, at an angle  $\phi$  with respect to the stator, be short-circuited. The resultant short-circuit current produces a torque which is proportional to  $\sin^2(\theta - \phi)$  and will set the rotor in motion until it comes to rest at  $\phi = \theta$  or  $\theta - 180^\circ$ , depending on which has the smaller angular displacement from the original position of the rotor. With the rotor at its equilibrium position, induced voltage in the short-circuited rotor winding is zero, while that in the open-circuited one has amplitude equal to the magnitude of the input vector. If rotor angle is at  $\phi = \theta$ , the output voltage is in phase with the

\* Received by the IRE, February 16, 1959.

\* Received by the IRE, February 16, 1959. The research reported in this document was supported by National Science Foundation Grant No. G-6316.

<sup>1</sup> H. Kleinwachter, "Excitation of electromagnetic fields by current waves," *Arch. für Elektr. Über.*, vol. 6, pp. 376-378; September, 1952.

<sup>2</sup> R. Muller, "Space harmonics in electron beams," *Arch. für Elektr. Über.*, vol. 10, pp. 505-511; December, 1956.

<sup>3</sup> E. J. Gorn, "Traveling-Wave Electron Reaction Device," U. S. Patent 2,591,350; April 1, 1952.

stator input carrier waves. If  $\phi = \theta - 180^\circ$ , the output is out of phase with the stator input by  $180^\circ$ .

In many applications, the input vector does not have phase discontinuity, or phase jump. The resolver will then be an extremely useful indicating and control device of phase and magnitude, with the phase of the rotor output always available for checking possible error of  $180^\circ$ . The self-actuating resolver described here does not have transformation accuracy attainable with a servo-driven resolver. But then simplicity alone justifies the use of a short wire in place of even the most rudimentary servo equipment in many instances.

CARSON K. H. TSAO  
Cruft Laboratory  
Harvard University  
Cambridge, Mass.

*Principal European contributions to the discussions of distant (beyond-the-horizon) propagation of ultra short radio waves came from Great Britain, Sweden, France, and Germany.*

G. C. Rider, of Marconi's Research Laboratory, Great Baddow, England, described propagation measurements made during 1956-1958 at  $f=858$  mc over ranges up to 585 km. He found a positive correlation between signal fading rate and wind velocity and negative correlation between signal strength and wind velocity. He also found that when a polar air mass was over the propagation path the signal strength was low. Rider showed plots of the hourly median values of signal strength vs path length (actually given in terms of the angle subtended by the path at the center of the earth).

G. Carlson, of the Research Institute of National Defense, Stockholm, Sweden, described experiments with pulsed signals at  $f=3000$  mc over a distance of 300 km in the middle of Sweden. He found that the median signal was 83 db below the free-space level and that the fading range between the 10 per cent and 90 per cent levels was about 11 db.

F. du Castel and P. Misme, of the Centre National d'Etudes des Telecommunications, Paris, France, discussed the results of many months of measurements in French Equatorial Africa on the propagation of radio waves at  $f=432$  mc over ranges of 170-280 km. They mentioned the difficulty of interpreting their experimental measurements on the basis of turbulent diffusion. They were examining the importance of reflection from layers in the troposphere and from this point of view had obtained theoretical values of the field strengths at great distance which seemed in accord with the principal experimental data.

U. Kuhn, of the Betriebslaboratorium für Rundfunk und Fernsehen, discussed the propagation of VHF and UHF radio waves over irregular terrain. He presented plots showing the dependence of local field strength variations on terrain irregularities for different frequency regions. In essence he investigated the discrepancies between the practical field strength measurements and the theoretical field strength values for the smooth earth.

L. Ames, of the U. S. Air Force Cambridge Research Center, discussed airborne measurements of radio fields of a transmitter operating at  $\lambda=1.36$  m out to ranges exceeding 900 miles and altitudes exceeding 40,000 feet. Representatives of the U. S. Air Force Cambridge Research Center and Rome Air Development Center discussed applications of the troposphere scatter mode of propagation to USAF communications.

*Principal papers by the Europeans on radio propagation experiments involving the ionosphere came from the representatives from Italy and Germany.*

I. Ranzi, of Centro Radioelettrico Sperimentale (G. Marconi), Rome, Italy, described experiments being performed near Rome on oblique back-scatter soundings using a rotating Yagi antenna. He compared skip distances observed by back scatter with skip distances deduced from vertical sound-

ing near the midpoint of the path and found discrepancies.

W. Dieminger, of the Max Planck Institut für Aeronomie, Lindau über Northeim, Germany, reported on combined vertical soundings of the ionosphere, oblique incidence propagation, and back-scatter experiments which were carried out over distances of 2000 km and 8000 km. In particular he discussed the difference between the values of the MUF (maximum usable frequency) derived from the vertical soundings and the values observed in the simultaneous oblique incidence propagation experiments. He also indicated agreement between the oblique incidence propagation experiments and simultaneous back-scatter measurements. Dieminger apparently uses two methods for his back-scatter experiments. In the first method he uses a fixed frequency rotating antenna and obtains the zone of reception as a function of the bearing of the antenna beam. In the second method he has the antenna on a fixed bearing and varies the frequency of his back-scatter radar, thus obtaining the extent of the skip zone as a function of frequency.

B. Beckmann and K. Vogt, of Fernmeldetechnisches Zentralamt, Darmstadt, Germany, reported on back-scatter observations with telegraphy signals. They described a simple procedure depending on back-scatter amplitude observation that is suitable for an evaluation of ionosphere communication conditions. The scheme has been developed by the German post office for application to telegraphic communications. Since with telegraph signals the travel time of the back scatter is small compared with duration of the signal, the back-scatter amplitude affects the envelope of the telegraph signal as observed in the intermediate frequency channel of the receiver. One observes amplitude discontinuities which can best be observed by the "ring method" of display in which the IF voltage of the receiver is applied to one pair of plates of an oscilloscope and the same IF voltage shifted in phase  $90^\circ$  is applied to the orthogonal pair of plates. One then assesses the ionosphere communication conditions from observing on the scope the first discontinuity caused by the back-scatter amplitude. Observations show that the amplitude of the back scatter increases as the size of the skip zone decreases and as the ionosphere attenuation decreases. Since both of these factors determine the propagation conditions, the actual usable communication frequency of the ionosphere can be estimated at the transmitter location. Back-scatter observations taken simultaneously with short pulses and with longer telegraph signals were compared. Finally back-scatter measurements at  $f=25.4$  mc with directive antennas in the direction of North America were compared with a field strength record of the American transmitter WWV at  $f=25$  mc.

*Papers dealing with the forward scattering of radio waves from the ionized trails of meteors were given by representatives from Great Britain and the United States.*

G. A. Isted, of Marconi's Research Laboratories, Great Baddow, England, discussed meteor activity as a factor in ionosphere scatter propagation. Although ionosphere

## International Conference on Radio Wave Propagation\*

Approximately 100 radio scientists from 12 countries attended the Congres International sur la Propagation des Ondes Radio-Electriques held on October 6-11, 1958, in the Palais des Congres at Liege, Belgium. The conference was organized by the Postal and Telecommunication Groups of the Universal and International Exposition at Brussels, 1958. A total of 55 papers were presented at the Congres. About half of these papers were from the United States; the rest came from the European countries. Only three papers came from delegates of the Soviet block countries, two from East Germany, and one from Czechoslovakia. None of the five delegates from the USSR presented any papers or participated in the discussions.

The papers presented at the Congres dealt largely with summaries of the work of several years rather than with discrete details. Major discussions concentrated on the following subjects:

- 1) Distant (beyond-the-horizon) propagation of ultra short radio waves and the application to radio communications.
- 2) Ionospheric investigations—particularly those using back-scatter sounding techniques.
- 3) Forward scattering of radio waves from the ionized trails of meteors.
- 4) Propagation of low-frequency radio waves.

Several of the papers which I considered most significant are summarized in the following pages. Summaries of all the papers are given in the "Resume des Communications" distributed to those who attended the conference. The complete texts of the papers presented at the conference will be published as a special volume by the Academic Press, Inc., New York, N. Y.

\* Received by the IRE, February 16, 1959.

scatter links have been operating now for several years, the observed propagation effects have not all been satisfactorily explained. Ionization by meteors undoubtedly contributes to ionospheric forward scattering. Isted suggests that weather cloud discharges may be capable of producing bursts of ionization similar to those produced by meteors. He has found groups of bursts and thundercloud discharges associated with these burst groups.

J. S. Greenhow and E. L. Neufeld, of the Joddrell Bank Experimental Station, University of Manchester, England, discussed turbulence in the lower  $E$  region of the ionosphere as inferred from meteor echo observations. Under the influence of high altitude winds the ionized trails of meteors show steady drifts due to uniform wind motions and become distorted through the effects of turbulence. Using a single station coherent pulse technique the authors have measured the steady winds and have found a prevailing component and regular 12-hour and 24-hour periodic variations. Using a spaced station coherent pulse technique with two receivers spaced at different distances from the transmitter they can make simultaneous wind velocity measurements at two points along a single meteor trail. From measurements on many trails they infer the wind shears over different heights. The median value of wind shear is found to be 10 m/sec/km with occasional observed values of 100 m/sec/km. The turbulent velocity is about 25 m/sec.

T. J. Keary and H. J. Wirth, of the U. S. Navy Electronics Laboratory, San Diego, Calif., discussed the statistical characteristics of forward-scattered echoes from meteor trails as determined from experimental observations made over a 690 km path from Stanford, Calif., to San Diego, Calif. at  $f=43.5$  mc.

J. Heritage, S. Weisbrod, and W. Fay, of Smyth Research Associates, San Diego, Calif., reported (paper communicated by L. Trolese) on an experimental study of meteor echoes at  $f=200$  mc with the receiving sites at distances of 800, 1300, 1800 km from the transmitter. The transmission path was oriented SE-NW. The shapes of the echoes observed at  $f=200$  mc are found to be extremely varied. The rates of occurrence of echoes showed the usual diurnal variation with highest rates in the morning hours. A comparison was made between rates of occurrence observed at receiving sites located NE of great circle path defined by the sharp transmitter beam with those located SW of it. In general the rates of occurrence were higher at the SW sites during the hours 0700 to 1700 and higher at the NE sites during 1700 to 0700.

*Aspects of the propagation of low frequency radio waves were discussed principally by representatives of the United States and Germany.*

J. R. Wait, of the National Bureau of Standards, Boulder, Colo., reviewed theoretical approaches to the understanding of the propagation of long radio waves over the surface of the earth. The two principal lines of approach are the mode theory and the ray theory. At frequencies  $f > 20$  kc and at long

ranges the mode theory is cumbersome to apply as many modes must be taken into account in computing the mode series to give the field. Hence it is usual to calculate distant fields by the geometrical optics or ray methods. Wait compared his calculated results with experimental data at  $f=16.6$  kc taken by members of the U. S. Navy Electronics Laboratory.

B. Friedman, of the University of California, Berkeley, Calif., discussed from the standpoint of mode theory the field produced by a low frequency vertical electric dipole located between a perfectly conducting earth and an ionosphere assumed to have a sharp lower boundary. He discussed the attenuation of the individual modes.

E. A. Lauter, of the Observatorium für Ionosphären Forschung, Ostseebad, Kühlingsborn, reported on measurements taken over a period of 10 years on the reflection coefficients of the ionosphere at a frequency  $f=245$  kc. He finds that for the undisturbed ionosphere the variation of reflection coefficients can be described by a log-normal distribution. During earth magnetic storms there are strong departures from the normal distribution.

K. Rawer, of the Fernmeldetechnisches Zentralamt, Aussenstelle Ionosphären-Institut, Breisach/Rhein, compared different methods of calculating the field strengths of radio waves reflected from the ionosphere.

T. J. KEARY  
Navy Electronics Lab.  
San Diego 52, Calif.

## A Topological Nonreciprocal Network Element\*

In order to extend equivalent circuit representation to nonreciprocal networks it has been found necessary to augment the set of ideal elements required for reciprocal networks (*viz.*, the independent two-terminal generator, the two-terminal immittance and the ideal transformer) by an essentially two-terminal-pair (two-port) nonreciprocal element. A variety of elements,<sup>1-7</sup> of which the gyrator appears to have gained the widest acceptance, has been proposed for this purpose. These elements are generally charac-

terized by physical (transfer immittance) as well as topological (connective) properties. It is the intention of this communication to indicate that the physical and topological properties may be segregated and the latter accounted for by an element which, in association with physical elements (immittances) may be used to model the more complex nonreciprocal elements so far proposed.

The proposed element is shown in Fig. 1(a); it is an irreducible essentially transmissive three-terminal device 123 which provides unit gain unilateral voltage transmission over path 132 (*i.e.*, from 13 to 32, with 3 common), unit gain unilateral current transmission over path 213 and null transmission over the four remaining paths. Its transmission matrices for paths 132, 213 have unity in the 11, or 22 element positions, with zeros elsewhere. When a signal potential is established at terminal 1 it is transmitted unilaterally, without change, to terminal 2. Dually, when a signal current is set up in the lead to terminal 2 it is transmitted unilaterally, without change (except, under the usual symmetrical sign convention, for reversal of sign) to terminal 3. There must, of course, be a continuous external current path between 23, one point of which will normally form a node-pair with terminal 1 and will form the reference or ground node. The union of these elements with the 123 element comprises a topological space, into which physical, *i.e.*, generative, storage and dissipative, elements may be introduced. It may be shown that the two unit-gain transmission processes are not separable, but that the one necessarily implies the other; they are dual aspects of a single common transfer mechanism, or transfer action (transaction). They are indicated in Fig. 1(a) by directive equality signs, using open arrowheads for voltage transmission, closed heads for current. The signal flow properties are shown graphically in Fig. 1(b), where the available terminal-pairs 12, 23, 31 are connected by directed line segments which are also distinguished by open or closed arrowheads and are labelled with the terminal sequence 132, 213 in the corresponding transmission path.

The two transmission paths may be cascade-coupled in either of two distinct ways. When an immittance  $Y$  is inserted in the lead to terminal 2, between 2 and the ground node 0 [see Fig. 2(a)], it acts as a voltage-to-current converter and defines  $i_2$  at  $i_2 = -Y \cdot v_2 = -Y \cdot v_1$  and, hence,  $i_3$  at  $i_3 = Y \cdot v_1$ . The over-all result is a unilateral coupling between the node-pair 10 and the branch 30 over path 132 through the nonreciprocal 123 element, which is equivalent to a direct unilateral transadmittance action (admittance transaction), as shown by the signal flow graph at Fig. 2(b). Dually, when an immittance  $Z$  is bridged across terminals 1,3 [see Fig. 2(c)] it acts as a current-voltage converter and couples the two unilateral paths in the opposite sequence, the over-all result being equivalent to a direct unilateral transimpedance action (impedance transaction), as illustrated by the flow graph [Fig. 2(d)]. By cascading a 123 element, appropriately oriented, with an ideal transformer of nonunit ratio, unilateral current-current or voltage-voltage transactions of other than unity ratio may be obtained. Thus, all four of the set of transactors introduced by the

\* Received by the IRE, February 18, 1959.

<sup>1</sup> B. D. H. Tellegen, "The gyrator, a new electric network element," *Philips Res. Rept.*, vol. 3, pp. 81-101; April, 1958.

<sup>2</sup> J. Shekel, "The gyrator as a 3-terminal element," *Proc. IRE*, vol. 41, pp. 1014-1016; August, 1953 and "Reciprocity relations in active 3-terminal elements," *ibid.*, vol. 42, pp. 1268-1270; August, 1954.

<sup>3</sup> W. S. Percival, "The graphs of active networks," *Proc. IEE*, vol. 102, pt. C, pp. 270-278; September, 1955.

<sup>4</sup> S. J. Mason, "Docile behavior of feedback amplifiers," *Proc. IRE*, vol. 44, pp. 781-787; June, 1956 and "Topological analysis of linear nonreciprocal networks," *ibid.*, vol. 45, pp. 829-838; June, 1957.

<sup>5</sup> J. L. Merrill, "Theory of the negative impedance converter," *Bell Sys. Tech. J.*, vol. 30, pp. 88-109; January, 1951.

<sup>6</sup> E. F. Bolinder, "Survey of some properties of linear networks," *IRE TRANS. ON CIRCUIT THEORY*, vol. CT-4, pp. 70-78; September, 1957.

<sup>7</sup> G. E. Sharpe, "Axioms on transactors," *IRE TRANS. ON CIRCUIT THEORY*, vol. CT-5, pp. 189-197; September, 1958.

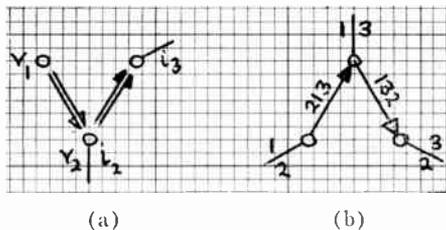


Fig. 1—(a) The proposed nonreciprocal network element and (b) its signal flow graph.

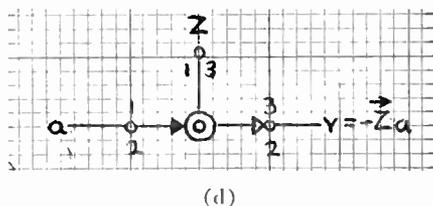
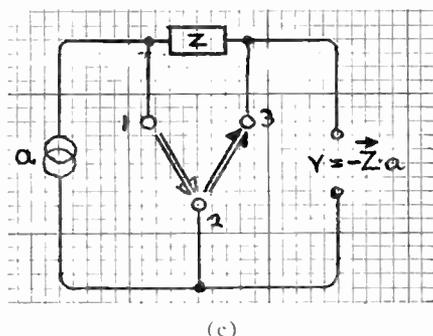
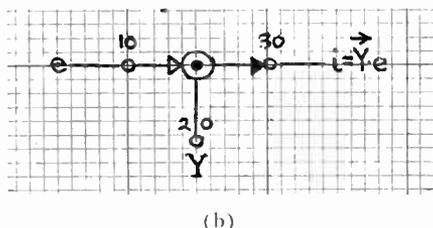
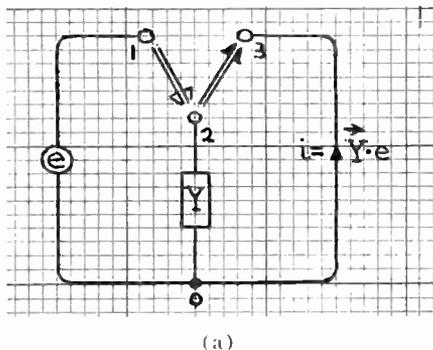


Fig. 2—The dual basic operating circuits of the proposed element: (a) admittance form, (b) flow graph, (c) impedance form and (d) flow graph.

author in a previous communication<sup>8</sup> may be modeled by the one type of nonreciprocal element in association with established reciprocal elements. To obtain complex current or voltage transaction, *i.e.*, unilateral transformation, it is sufficient to connect a unit real immittance across 10 in Fig. 2(a) or in series with the lead to 1 in Fig. 2(c), with  $Y$  and  $Z$  complex, respectively. Immittances

<sup>8</sup> A. W. Keen, "Transactive network elements," *J.I.E.E.*, vol. 3, pp. 213-214; April, 1957. And, "The transactor," *Elec. and Rad. Eng.*, vol. 34, pp. 459-461; December, 1957.

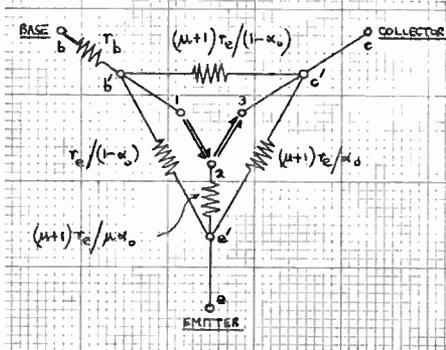
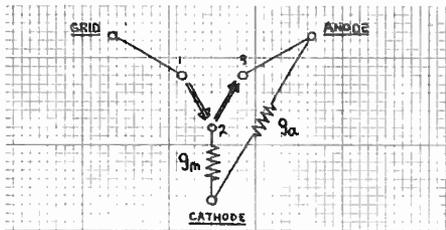


Fig. 3—(a) Models of the linear low-frequency negative-grid thermionic triode and (b) the corresponding transistor.

may be associated with the output side in a similar manner, so that current and voltage amplifiers having any one of the four basic feedback configurations (*viz.* series-series, series-shunt, shunt-series, shunt-shunt) may be formed.

By proceeding in this way one may develop the circuit technique of the 123 element to embrace thermionic tube and transistor circuit technique as particular cases. Models of the small-signal, low-frequency action of the tube and transistor are shown in Fig. 3(a) and 3(b), respectively. In both cases the model consists of a nonreciprocal 123 element imbedded in a reciprocal shell and it will be noted that all of the parameters of the device are assigned to immittances in the latter portion. By successive stripping of the shell immittances, the transistor model reduces in turn to its intrinsic form to the equivalent of a tube, to an admittance transactor and, finally, to the 123 element itself.

An important class of networks arises from closure of the internal transmission path of the 123 element by an external signal-flow path between 3 and 1 with the ground point 0 common. The result is a feedback circuit, the simplest examples of which are the basic thermionic triode circuits.<sup>9</sup> Alternatively, the feedback loop may be formed by a second 123 element coupled to the first in such an orientation that at least three of the four internal transmission paths occur in cascade. Such forms may be used to model any of the proposed bilateral nonreciprocal elements (gyrator, anti-symmetrical amplifier, impedance inverter, etc.). As an example, the series-series feedback form for the gyrator is given in Fig. 4, which

<sup>9</sup> A. W. Keen, "Triode network topology," *IRE TRANS. ON CIRCUIT THEORY* (to be published).

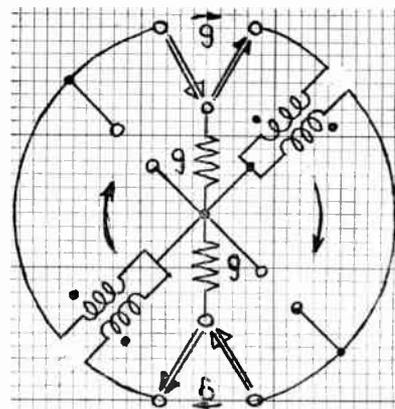


Fig. 4—Model of the admittance form of the gyrator [with  $\epsilon = 1$  and all three terminal pairs (ports) available for external connection].

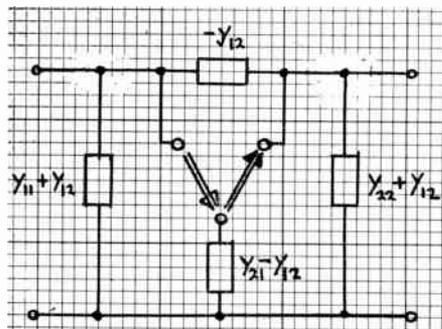
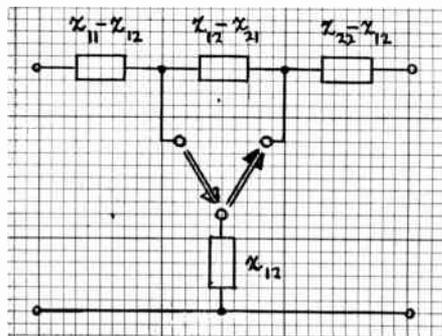


Fig. 5—Models of the general two-terminal pair network. (a) Impedance form and (b) admittance form.

illustrates very clearly its action. Only one of the two ideal transformers used is necessary, the one which provides inversion and thereby avoids the need for a negative real immittance. In the special case of unit  $g$  elements and with all three terminal-pairs brought out for external connection, Fig. 4 becomes a 3-port circulator.<sup>10</sup> The general nonreciprocal two-terminal-pair network may be represented by models containing either one or two 123 elements associated with four distinct immittances in any one of the four feedback configurations, *i.e.*, in forms corresponding to the  $G$ ,  $H$ ,  $Y$  and  $Z$  matrices; the  $Y$  and  $Z$  forms are shown in Fig. 5(a) and 5(b), respectively.

<sup>10</sup> H. J. Carlin, "Synthesis of nonreciprocal networks" *Polytech. Inst. Bklyn. Symp. Series*, vol. 5, pp. 11-44; April, 1955.

There remains the question of a suitable name for the 123 element. The author has used univector and univator in unpublished work but would suggest that UNITOR expresses succinctly the essential characteristics, viz., unilateral signal connectivity, unit gain magnitude and unity of structure.

A. W. KEEN  
Technical College  
Coventry, Eng.

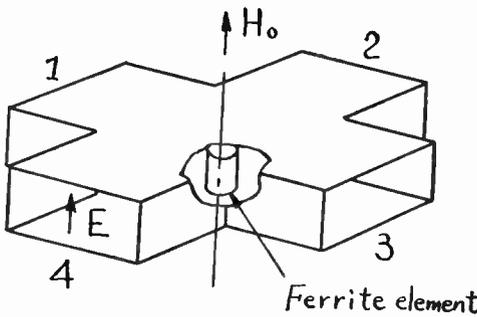


Fig. 1—X circulator.

**X Circulator\***

**INTRODUCTION**

Circulators consisting of junction of rectangular waveguides containing a nonreciprocal element—a properly magnetized ferrite—have been proposed and developed by many authors.<sup>1-3</sup>

In this paper I would like to propose mainly a new four-port waveguide circulator of the structure shown in Fig. 1 which is simple in structure, small in volume, light in weight, and pretty good in characteristics. The operation of this device is based on the same principle as the *Y* circulator<sup>2</sup> and the *T* circulator<sup>3</sup> of the three-port structure, although the exact theory of its operation has not yet been analyzed. Next I would like to discuss the experimental results of the proposed circulator in *X* band.

**EXPERIMENTAL RESULTS**

The ferrite (or garnet) element whose material, shape, and size are proper for the characteristics of a circulator, is inserted with a suitable impedance element at the right place in the usual *H*-type four-port waveguide junction, in which four standard rectangular waveguides in *X* band (WR-90) intersect at right angles at one point. Then a dc external magnetic field is applied to this ferrite element, transverse to the broad side of this waveguide junction by an external electromagnet.

When the wave of 9375-mc frequency enters, for example, through arm 4 from a matched generator, it comes out from arms 1, 2, and 3 terminated by matched detectors. Their transmission losses are shown against the applied dc magnetic field in Fig. 2. When the external dc magnetic field strength *H*<sub>0</sub> corresponding to point *A* or *B* in this figure is applied, the counter-clockwise or clockwise circulator is made up, respectively. In practice barium ferrites having about 300 oersteds are set to this junction; the frequency characteristics of the circulator corresponding to point *B* in Fig. 2 are shown in Fig. 3. In the above-mentioned

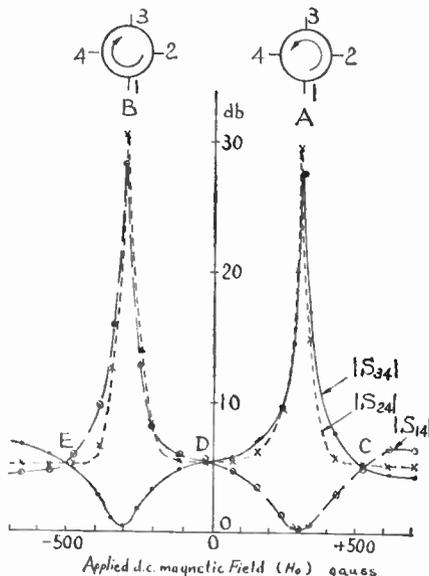


Fig. 2—Transmission loss vs dc magnetic field.

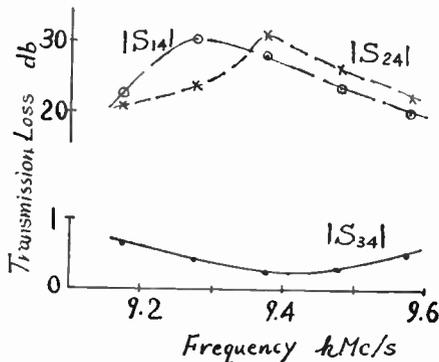


Fig. 3—Transmission loss vs frequency.

example, a rod of manganese magnesium ferrite is placed on the bottom broad wall at the center of the junction as the nonreciprocal element, against the impedance element which is needed to improve the characteristics of the circulator (Fig. 1). When the wave enters from any other arm, the characteristics previously mentioned are kept almost the same.

As the result of the trial to reduce the external dc magnetic field strength proper for the circulator characteristics, the circulator, which needs about 130 oersteds and which has almost the same electrical characteristics, is made by choosing the ferrite

element of the same material which is thinner in diameter and bigger in height, with the same kind of the impedance element.

When the external dc magnetic field strength *H*<sub>0</sub> corresponding to point *C*, *D*, or *E* in Fig. 2 is applied, the four-port matched equal power divider in the nonreciprocal circuit is made up successfully in practice.

In the case of the four-port *H*-type waveguide junction made of four standard rectangular waveguides in *X* band, two of which intersect at 180°, and three of which intersect at 60°, a circulator is made successfully by using almost the same technique. However, its characteristics are inferior to the *X* circulator at present.

SHINICHIRO YOSHIDA  
Matsuda Research Lab.,  
Tokyo Shibaura Electric Co.,  
Kawasaki, Kanagawa-Ken,  
Japan

**A Stability Characterization of the Reciprocal Linear Passive *N* Port\***

It is well known<sup>1-3</sup> that a reciprocal two-port is passive if and only if it is stable under all passive terminations. Does this theorem generalize to the linear reciprocal *n* port? To the author's best knowledge this question has never been answered. In this note we prove that the answer is in the affirmative.

Let *A* be an arbitrary matrix. Then *A'*,  $\bar{A}$ , *A*\* and det *A* stand for the transpose, the complex conjugate, the complex conjugate transpose and the determinant of *A*, respectively. Column vectors are represented by *a*, *b*, *x*, *I*, etc. For a hermitian matrix  $A = A^*$ ,  $A \geq 0$  means that *A* is the matrix of a non-negative form and  $A > 0$  that it is the matrix of a positive form. A diagonal matrix *A* with diagonal elements  $\mu_1, \mu_2, \dots, \mu_n$  is written as  $A = \text{diag} [\mu_1, \mu_2, \dots, \mu_n]$ .

Theorem 1: Let  $Z(i\omega) = Z'(j\omega)$  be the impedance matrix of a reciprocal *n* port *N*. Then *N* is strictly passive, i.e.,  $\text{Re } Z(j\omega) > 0$ , if and only if it is stable.

Proof: Refer to Fig. 1. Suppose  $\text{Re } Z > 0$  and *N* is not stable. By definition there exist *n* complex numbers  $z_1, z_2, \dots, z_n$  with non-

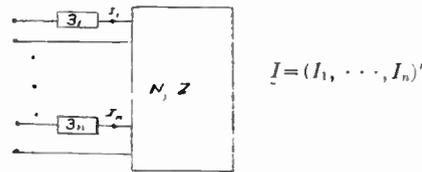


Fig. 1

\* Received by the IRE, February 10, 1959. The reported herein was sponsored by the Air Force Office of Scientific Research under Contract No. AF-18 (603)-105.

<sup>1</sup> F. B. Llewellyn, "Some fundamental properties of transmission systems," Proc. IRE, vol. 44, pp. 1345-1368; October, 1956.

<sup>2</sup> G. Raisbeck, "A definition of passive linear networks in terms of time and energy," J. Appl. Phys., vol. 25, pp. 1510-1514; December, 1954.

<sup>3</sup> E. Folke Bolinder, "Survey of some properties of linear networks," IRE TRANS. ON CIRCUIT THEORY, vol. CT-4, pp. 70-78; September, 1957.

\* Received by the IRE, January 26, 1959; revised manuscript received, February 19, 1959.

<sup>1</sup> For example, see A. G. Fox, S. E. Miller and M. T. Weiss, Bell Sys. Tech. J., vol. 34, pp. 65-103; January, 1955.

<sup>2</sup> This circulator was investigated by M. H. Sirvetz and his collaborators in Research Div., Raytheon Mfg. Co. last spring.

<sup>3</sup> W. E. Swanson and G. J. Wheeler, "Tee circulator," 1958 WESCON CONVENTION RECORD, pt. 1, pp. 151-156.

negative real parts  $r_1, r_2, \dots, r_n$ , so that the linear system

$$(Z + Z_0)I = 0, \quad Z_0 = \text{diag} [z_1, z_2, \dots, z_n], \quad (1)$$

possesses a nontrivial solution  $I$ . From (1),

$$I^*(\text{Re } Z)I = - \sum_{k=1}^n r_k |I_k|^2 \leq 0,$$

a contradiction. Thus stability is necessary.

Sufficiency. For  $n=1$  this is obvious. For arbitrary  $n$  we proceed by induction, that is to say we assume the result to be true for all  $n < k+1$  and show that this implies its truth for  $n = k+1$ .

By hypothesis the  $(k+1)$  port  $N$  is stable from which it is easily seen that the  $k$  port  $N_1$  [Fig. 2(a)] and 1 port  $N_2$  [Fig. 2(b)] that are created by 1) terminating the  $(k+1)$ th port of  $N$  in an arbitrary passive impedance  $z$  and 2) open-circuiting respectively the first  $k$  ports of  $N$  are also stable. Partition  $Z$  as follows:

$$Z = \begin{pmatrix} 1 & & \\ Z_a & Z_b & \\ Z_b' & Z_c & \\ & & k \\ & & & 1 \end{pmatrix} \quad (2)$$

The impedance matrices of  $N_1$  and  $N_2$  are

$$Z_1 = Z_a - \frac{Z_b Z_b'}{z + Z_c} \quad (3)$$

and

$$Z_2 = Z_c \quad (4)$$

respectively. By the induction hypothesis,  $\text{Re } Z_1 > 0, \text{Re } Z_2 > 0$ . In particular, when  $z = \infty, Z_1 = Z_a$  and therefore  $\text{Re } Z_a > 0$ . The object of course is to prove that

$$\text{Re } Z = \begin{pmatrix} \text{Re } Z_a & & \\ \text{Re } Z_b' & \text{Re } Z_c & \\ & & \end{pmatrix} > 0. \quad (5)$$

Because  $\text{Re } Z_a > 0$ , it is only necessary, according to a standard theorem concerning positive forms,<sup>4</sup> to establish that  $\det (\text{Re } Z) > 0$ .

Let  $x$  be any real  $k$  vector. Then

$$\text{Re } x' Z_1 x > 0. \quad (6)$$

Let

$$\begin{aligned} a &= x' Z_a x, \\ c &= Z_b' x, \\ d &= Z_c \end{aligned} \quad (7)$$

and

$$W = x' Z_1 x.$$

From (3),

$$W = a - \frac{c^2}{z + d} \quad (8)$$

and (6) implies that  $\text{Re } W > 0$  for every  $z$  having a non-negative real part. In other words  $W$  maps the closed right half-plane into the open right half-plane. We leave it to the reader to verify that this is the case if and only if

$$(\text{Re } a)(\text{Re } d) - (\text{Re } c)^2 > 0, \quad (9)$$

or substituting from (7), if and only if

<sup>4</sup> Mirsky, "Introduction to Linear Algebra," Clarendon Press, Oxford, Eng., 1955.

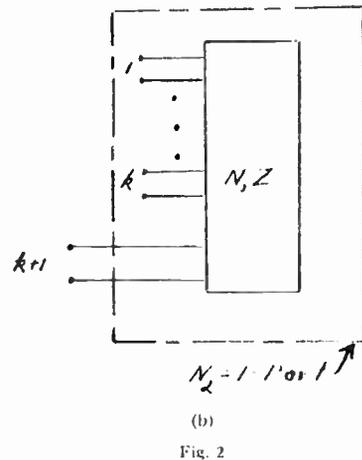
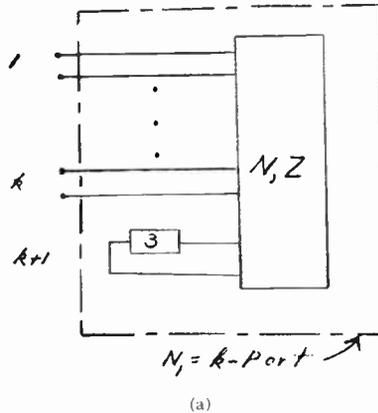


Fig. 2

$$x' [(\text{Re } Z_a)(\text{Re } Z_c) - (\text{Re } Z_b)(\text{Re } Z_b')] x > 0 \quad (10)$$

for every real  $k$  vector  $x$ . Hence the real symmetric  $k \times k$  matrix

$$V = (\text{Re } Z_c)(\text{Re } Z_a) - (\text{Re } Z_b)(\text{Re } Z_b') \quad (11)$$

is the matrix of a positive form whence<sup>4</sup>  $\det V > 0$ . From the identity

$$\begin{pmatrix} \text{Re } Z_a & \text{Re } Z_b \\ \text{Re } Z_b' & \text{Re } Z_c \end{pmatrix} \cdot \begin{pmatrix} 1_k \text{Re } Z_c & 0_{k,1} \\ -\text{Re } Z_b' & 1_1 \end{pmatrix} = \begin{pmatrix} V & \text{Re } Z_b \\ 0_{1,k} & \text{Re } Z_c \end{pmatrix} \quad (12)$$

we get, by taking determinants of both sides,

$$\det (\text{Re } Z) = \frac{\det V}{(\text{Re } Z_c)^{k-1}} > 0. \quad (13)$$

The proof is now completed by induction from  $n=1, \text{Q.E.D.}$

The definition of stability we have adopted is the conventional one: an  $n$  port  $N$  is stable if the port currents are zero under all passive terminations. Evidently lossless networks are not stable. To remedy this defect we introduce the notion of weak stability: an  $n$  port  $N$  is weakly stable if the port currents are zero under all passive terminations whose real parts are simultaneously positive (zero is excluded). It is now possible to give a stability characterization of all reciprocal passive  $n$  ports.

Theorem 2: Let  $Z = Z'$  be the impedance matrix of a reciprocal  $n$  port  $N$ . Then  $N$  is passive, i.e.,  $\text{Re } Z \geq 0$ , if and only if it is weakly stable.

Proof: For any  $\epsilon > 0, Z + \epsilon I_n$  is the impedance matrix of a reciprocal stable  $n$  port. Invoking Theorem 1,

$$\text{Re } b^*(Z + \epsilon I_n)b > 0$$

for any nontrivial  $n$  vector  $b$ . Going to the limit,

$$\text{Re } (b^* Z b) = \lim_{\epsilon \rightarrow 0} \text{Re } b^*(Z + \epsilon I_n)b > 0,$$

i.e.,  $N$  is passive, Q.E.D.

The author wishes to take this opportunity to thank Dr. Carlin of these Laboratories for suggesting the problem.

D. YouA  
Polytech. Inst. Brooklyn  
Brooklyn, N. Y.

### Minimum Transmitter System Weight for Space Communications\*

In a space communications link, the required transmitter power is decreased as the transmitter antenna gain is increased, all other parameters remaining constant. Thus, the link may utilize a light-weight, low-power transmitter with a large, high-gain antenna or a heavy-weight, high-power transmitter with a small, low-gain antenna. We wish to call attention to the fact that for a required transmitter-power-antenna-gain product, there exists an optimum transmitter power and antenna gain from the standpoint of minimizing the weight of the total transmitting system.

The weight of the transmitting antenna,  $W_A$ , may be given by the relation

$$W_A = CD^2 \quad (1)$$

where  $D$  is the diameter of the antenna aperture, and  $C$  is a constant dependent upon the type of antenna used. For a parabolic antenna constructed of aluminum-phenolic honeycomb with a Fibreglass skin one-half inch thick,  $C$  equals 0.16 pound per square foot. (This assumes the thickness is independent of the diameter, which, in weightless space, is reasonable for diameters up to 60 ft.)

The transmitter power is taken to be proportional to the transmitter weight, plus the weight of the power source. The total transmitter weight,  $W_X$ , is expressed as

$$W_X = (K_X + K_B)P_I \quad (2)$$

where  $P_I$  is the transmitter input power,  $K_X$  is the constant relating the transmitter power to its weight, and  $K_B$  is the constant relating the battery power to its weight. Present-day transmitters in the 2000-mc range have an estimated  $K$  of 1.5 pounds per watt. Solar batteries, illuminated 50 per cent of the time, give an estimated value of 0.2 pound per watt for  $K_B$ . This includes any necessary storage batteries.

The communication link requires a certain transmitter-power-antenna-gain product (depending on range, bandwidth, signal-to-noise, etc.) The antenna gain is propor-

\* Received by the IRE, March 12, 1959.

tional to the square of its diameter; therefore, the product,  $M$ , is expressed as

$$M = P_1 D^2 \quad (3)$$

where  $P_1$  is the input power to the transmitter (determined by the transmitter efficiency) and  $D$  is the antenna diameter.

The total weight of the space vehicle's transmitter-antenna combination is from (1) and (2) and expressed as

$$W = W_A + W_X = CD^2 + (K_X + K_B)P_1 \quad (4)$$

Substituting for  $P_1$  from (3) and letting  $K = K_X + K_B$ ,

$$W = CD^2 + \frac{KM}{D^2} \quad (5)$$

By differentiating (5) we find the antenna diameter which leads to the minimum total weight and expressed as

$$D = \left[ \frac{KM}{C} \right]^{1/4} \quad (6)$$

The optimum transmitter power then is determined from (3).

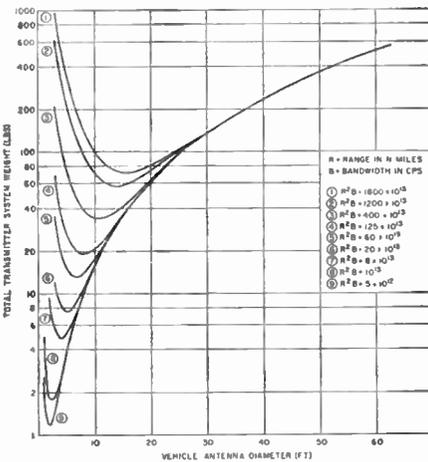


Fig. 1—Total weight vs transmitting antenna diameter with 60-ft receiving antenna at 2000 mc.

Fig. 1 shows a plot of total weights versus antenna diameters. With a frequency of 2000 mc, the assumed parameters were

- $C = 0.147$  pound per square foot
- $K_X = 1.5$  pounds per watt
- $K_B = 0.2$  pound per watt
- $K = 1.7$  pounds per watt

$$M = \frac{(C/N)16k(FT)R^2B\alpha(6080)^2}{E_f G_r \eta}$$

$$\frac{C}{N} = \text{carrier-to-noise ratio} = 30 \text{ db}$$

$$FT = \text{receiver sensitivity} = 30^\circ \text{ K}$$

$$\alpha = \text{system losses} = 10 \text{ db}$$

$$k = 1.4 \times 10^{-23} \text{ joules/}^\circ \text{K}$$

$$R = \text{range in nautical miles}$$

$$B = \text{bandwidth in cps}$$

$$E_f = \text{over-all transmitter efficiency} = 0.2$$

$$G_r = \text{receiving antenna gain} = 79,400$$

$$\eta = \text{efficiency of transmitting antenna aperture} = 0.55.$$

One can see how the optimum diameter varies with the range; therefore, an actual design would be based on the maximum range expected. Other considerations, such as limitations on the transmitter antenna pointing accuracy, could alter the design criteria described above.

The authors wish to thank Walter Kanysky of the Philco Western Development Laboratories' Advanced Development Group for deriving and plotting Fig. 1.

R. S. DAVIES  
C. S. WEAVER  
Advanced Dev. Group  
Western Dev. Labs.  
Philco Corp.  
Palo Alto, Calif.



Fig. 1—University of Illinois moon beam antenna.

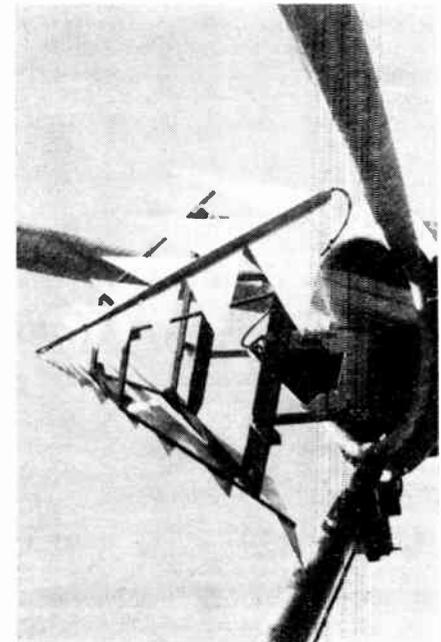


Fig. 2—Log-periodic primary antenna.

### A Log-Periodic Reflector Feed\*

A need for rapidly changing the operating frequency of the University of Illinois "moon bounce" receiving antenna, pictured in Fig. 1, resulted in the installation of the log-periodic primary feed, shown in Fig. 2. The twenty-eight-foot-diameter dish was originally fed with a corner reflector at 150 mc and a sectoral horn at 415 mc. Approximately a day was required in the necessary interchange of feeds when switching from one band to the other. The need for reducing this time to something comparable to that of changing receivers arose when it was decided that measurements at two widely separated frequencies at the same, or nearly the same, time would be useful. With data available at only a single frequency, there is an ambiguity of  $n\pi$  in the determination of the total angle of rotation of the plane of polarization of a wave transmitted to the moon and back through the ionosphere.

The practical realizability of a linearly polarized, unidirectional antenna having essentially constant electrical characteristics over bandwidths of greater than ten to one in frequency was demonstrated in the fall of 1957.<sup>1</sup> Subsequent investigation had provided means of beamwidth control and improved impedance characteristics so that little development was required in designing the new feed.

The primary antenna is a nonplanar log-periodic antenna.<sup>2</sup> It is made up of two plane sheet-metal elements described in terms of the geometry of Fig. 3. The values of the various parameters used for this installation are  $R_1 = 46$  in,  $\tau = 0.7$ , and  $\alpha = 45^\circ$ . The elements are mounted on a bakelite pyramidal frame at an included angle  $\psi = 60^\circ$ . The antenna is fed between the vertices of the two elements by means of a coaxial cable brought to the apex along one of the elements with its outer conductor bonded to the element. Note that no balun is required. It is the

property of log-periodic antennas of this type that at all frequencies greater than some low limit, determined by the maximum tooth length, only a portion of the elements is excited. As the frequency is varied, the current distribution along the elements remains nearly constant in terms of the wavelength.

The patterns of antennas of the type shown in Fig. 3 are characterized by a  $\cos^2 \theta$  shape in one hemisphere, with nearly equal principle plane beamwidths. The front-to-back ratio is approximately 15 db when the elements are mounted at an included angle  $\psi$  of about  $2\alpha$ . The beamwidth under this condition is shown in Fig. 4 for several values of  $\alpha$ . A dotted line through the points shows the variation to be nearly linear. It is possible to pick an  $\alpha$  from this figure which corresponds to a particular illumination taper for a given  $f/d$ . By reducing the included angle  $\psi$ , the beamwidth in the  $H$  plane may be increased without changing the  $E$ -plane pattern so that equiamplitude contours in a plane perpendicular to the beam axis are elliptical rather than circular. The antenna

\* Received by the IRE, March 2, 1959.  
<sup>1</sup> D. E. Isbell, "Non-Planar Logarithmically Periodic Antenna Structures," Seventh Annual Symposium on the USAF Antenna Research and Development Program, Abstract, Monticello, Ill.; October, 1957. Also published by University of Illinois, Urbana, Antenna Lab. Tech. Rep. No. 30; February, 1958.  
<sup>2</sup> Patent Pending, Ser. No. 768297.

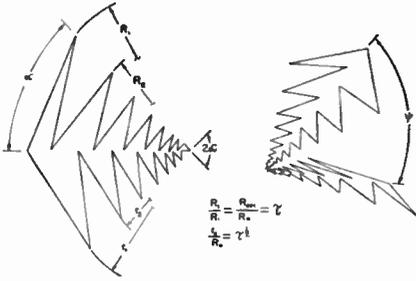


Fig. 3—Primary antenna parameters.

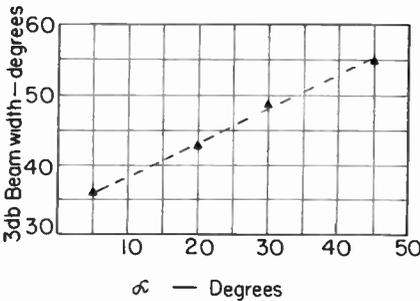


Fig. 4—Primary antenna beam width vs  $\alpha$ ,  $\psi = 28$ .

may thus be used to illuminate apertures other than circular with a reasonable degree of control in the amplitude at the edges.

The parameter  $\tau$  corresponds to the periodicity of the structure. It has not been found to be critical and a value in the range 0.6 to 0.8 works well with  $\alpha$  in the range  $60^\circ$  to  $30^\circ$ . If  $\alpha$  is reduced while  $\tau$  remains constant, the teeth tend to flatten out and lose their identity in which case the antenna ceases to function properly. This is avoided if the ratio  $\tau$  is made to approach unity as  $\alpha$  becomes small. If the aspect ratio of an individual tooth is defined as the ratio of its height to its width at the base, and  $\tau$  is adjusted to provide an aspect ratio near unity or greater, the antenna will function properly. This condition is met if

$$\frac{2\tau^{1/2} \sin \frac{\alpha}{2}}{1 - \tau} \geq 1.$$

A one-fourteenth scale model of the reflector and feed was used to measure secondary patterns over a band corresponding to the full scale range of 105 to 430 mc. The primary antenna was designed to provide a 10-db amplitude taper at the edges of the reflector as a compromise between gain and side lobe level. The half-power beamwidth of the secondary patterns is shown as a function of frequency in Fig. 5 for both *E* and *H* plane. The beamwidth decreases as frequency is increased in the expected manner up to about 325 mc beyond which it rises and then falls again. The problem here is primarily one of scaling. It has been found difficult to construct log-periodic antennas of the type here employed which provide the characteristic smooth, axially symmetric patterns much above 4000 mc. At about 4500 mc, 321 mc full scale, the primary patterns of the model feed were beginning to deteriorate as a result of this difficulty.

Taking as a reasonable approximation of the gain the ratio of 25,000 to the product of

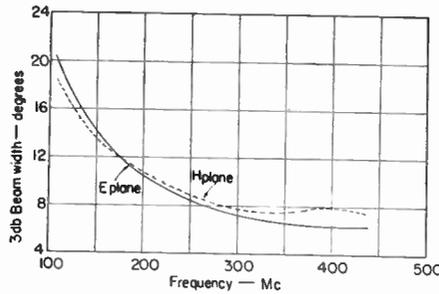


Fig. 5—Secondary pattern beam width.

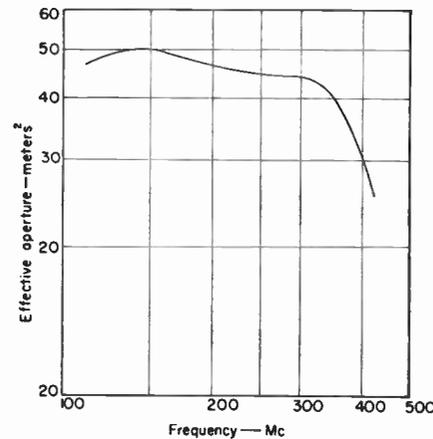


Fig. 6—Estimated effective aperture.

the half-power beamwidth in degrees and multiplying this result by  $\lambda^2/4\pi$ , a value for the effective aperture is obtained. The estimated effective aperture is plotted as a function of frequency in Fig. 6. Note that it is fairly constant in the band below 325 mc where the primary patterns of the model have not deteriorated.

The installation of the log-periodic feed in the University of Illinois antenna was completed in July, 1958, and has been in use since that time. In addition to "moon bounce" reception, it has been used for satellite tracking at 108 mc, and there are plans for its use in the range near 1000 mc.

DWIGHT E. ISBELL  
Dept. of Elec. Eng.  
University of Illinois  
Urbana, Ill.

### A Torsional Magnetostrictive Delay Line\*

Pulses may be propagated along a wire delay line either in a longitudinal or a torsional mode. The torsional mode is preferred because of its lower velocity of propagation and its negligible dispersion factor.

Hitherto published methods of propagating torsional pulses use either barium

titanate ceramic transducers<sup>1</sup> or longitudinal to torsional mode transformers.<sup>2</sup> Neither of these methods is readily adaptable for either a continuously variable or a multitap delay line. To overcome this disadvantage, a third technique may be used. In this technique, which is based upon the magnetostrictive phenomenon known as the Wiedemann effect, the conventional longitudinal coil transducers of a magnetostrictive delay line are replaced by toroidal coil transducers.

To generate the torsional pulse, a wire of magnetostrictive material, such as nickel, is initially biased within the transmitter toroid by both a longitudinal and a circular magnetic field in order to place the wire under torsional stress. The application of a pulse of electrical current to the transmitter toroid will change the amplitude of the circular magnetic field and produce a torsional pulse. This pulse will propagate along the nickel wire at a time delay of 8  $\mu$ sec per in. As the pulse passes through the receiver toroid, it will cause an incremental change in the amplitude of a biasing circular magnetic field which can be detected as a voltage pulse across the receiver toroid.

A basic delay line of this type is shown in Fig. 1. To provide the maximum magnetic flux linkage, the toroids are wound on bobbins of phenolic or equivalent material. The nickel wire threads through the toroids in close proximity to the individual turns of wire so as to obtain maximum magnetic coupling between the toroid and the nickel wire. The biasing circular magnetic field may be obtained by passing a dc current through the nickel wire, as shown in Fig. 1.

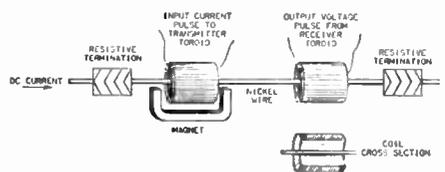


Fig. 1—Torsional magnetostrictive delay line with toroidal transducer coils.

Alternately, by introducing a dc current bias separately in each toroid, one can achieve independent amplitude control of the biasing circular magnetic field at each transducer. The biasing longitudinal magnetic field at the transmitter transducer is obtained from an external magnet.

The operational characteristics of the torsional magnetostrictive delay line are similar to those of the conventional longitudinal delay line in that maximum amplitude and minimum width of the output voltage pulse occur simultaneously when the width of the input current pulse is equal to the time of propagation through identical transducers. However, since the torsional time delay per unit length is greater than

<sup>1</sup> R. N. Thurston, and L. M. Tornillo, "Coiled wire torsional wave delay line," 1958 IRE NATIONAL CONVENTION RECORD, Pt. 2, p. 109.

<sup>2</sup> G. G. Scarrott, and R. Naylor, "Wire-type acoustic delay lines for digital storage," Proc. of the IRE, vol. 103, Pt. B, Supplement No. 3, Paper No. 2027M, p. 497; March, 1956.

\* Received by the IRE, February 27, 1959.

the longitudinal time delay by a factor of 1.6, the length of the torsional transducer must be reduced in the same ratio for a given width of input current pulse. For an input current pulse of 2  $\mu$ sec, for example, the length of the torsional transducer is 0.25 in. while the length of the longitudinal transducer is 0.4 in. However, it should be noted that in practice, the actual physical lengths must be reduced because of magnetic flux fringing at the ends of the transducers.

The author wishes to thank R. DeWitt of the ITT Laboratories for suggesting that the Wiedemann effect could be used to generate a torsional pulse in a magnetostrictive delay line.

A. ROTHBART  
ITT Labs.  
Nutley, N. J.

### Enhanced Correlator Output Through Optimum Delay-Line Processing\*

In dealing with stochastic time functions, it is often desirable to compute the cross-correlation function

$$\psi_{12}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f_1(t)f_2(t - \tau)dt, \quad (1)$$

where  $T$  is the integration time and  $\tau$  is the delay, to obtain a measure of the statistical correlation between  $f_1(t)$  and  $f_2(t)$ . We see from (1) that the operations of time delay, multiplication, and averaging are involved. The time delay and multiplication are physically realizable, but it is obviously not possible to average over infinite time. The errors due to finite observation intervals have been investigated by Davenport<sup>1</sup> and others, who have found for a "perfect integrator averaging filter" and very large  $T$  that the signal-to-noise ratio varies directly as  $\sqrt{T}$ .

In many physical situations, the stochastic time functions whose cross-correlation function we wish to compute may have been derived through processes employing the Doppler shift in which unequal spectrum shifts between  $f_1(t)$  and  $f_2(t)$  are introduced. In this case, excluding the effects of correlated noise, the cross-correlation function will vanish in the limit as  $T \rightarrow \infty$ . This is the aspect of correlation with which this paper is concerned, that is, the restoration of the stochastic time functions to their original spectral relationship.

Consider the stochastic variables  $f_1(t)$  and  $f_2(t)$  and their Fourier transforms  $F_1(\omega)$  and  $F_2(\omega)$ . Let  $f_2(t)$  incur a spectrum shift,  $\alpha$ , such that its transform is now  $f_2(\alpha\omega)$ , and the inverse transform is

$$F_2^{-1}(\alpha\omega) = f_2\left(\frac{t}{\alpha}\right). \quad (2)$$

A spectrum shift in the frequency domain becomes a time compression or expansion in the time domain.

If the effects of noise are not considered, there will be no correlation should

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f_1(t)f_2(t - \tau)dt \neq 0 \quad (3)$$

for then

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f_1(t)f_2\left(\frac{t - \tau}{\alpha}\right)dt = 0. \quad (4)$$

The total spectrum change is

$$\Delta\omega = (\alpha - 1)\omega. \quad (5)$$

The inverse transform of  $\omega$  is the unit doublet  $u''(t)$  which in this case is  $dt$ . Hence the inverse transform of  $\Delta\omega$  is

$$F^{-1}(\Delta\omega) = (\alpha - 1)dt \quad (6)$$

and the total delay time change,

$$\Delta\tau = \int_0^t (\alpha - 1)dt. \quad (7)$$

Thus if the delay in (1),  $\tau$ , is varied in accordance with the negative of (7), the total effective change in  $\tau$  will be zero.<sup>2</sup> The converse effect in the frequency domain will be the restoration of the original spectral relationship between  $F_1(\omega)$  and  $F_2(\omega)$ .

### CONCLUSION

The rate at which the delay of a correlator is varied is of prime importance. If the condition imposed by (7) is not met exactly, the integrand of the correlation function will tend to become oscillatory about zero, and a degradation in signal-to-noise ratio will result. The exactness with which (7) is implemented depends on the specific spectrum and also on the integration time,  $T$ . The longer the integration time, the closer  $\Delta\tau$  must approximate (7). If a prior knowledge of (7) is not available, then all possibilities must be tried.

H. SCHIMMEL  
Columbia University  
Hudson Labs.  
Dobbs Ferry, N. Y.

<sup>2</sup>  $\tau = \tau_0 + \Delta\tau$  where  $\tau_0$  is the delay at  $T = 0$ .

### Radiometer Circuits\*

Although the description given by the author<sup>1</sup> about the "alternate system" shown in his Fig. 1 is not exactly clear, I am convinced that he has made a rather serious error.

If Mr. Graham is correct, he has found a relatively simple way to eliminate noise in a finite time. For, if we follow his reasoning, by splitting a signal up into more and more

parts, adding a new "uncorrelated" noise to each part, until the signal in each channel becomes infinitesimal and the noise continues to increase in proportion, we end up with an infinite signal-to-noise ratio.

The error which I believe Mr. Graham has made is not a new one. When information theory was in its infancy, some of our leading authorities fell into the same trap. The IRE would perform a service by pointing out the fallacy in Mr. Graham's letter so as to prevent engineers who are not experts on the subject of noise from falling into a similar trap.

Here, I believe, is the fundamental error. In thinking of cross-correlating two uncorrelated noise voltages, it is erroneous to assume that the resultant output noise in a finite integration time reduces to zero (independent of the delay  $\tau$ ). Furthermore, whenever the input signal is split into two channels, it is incorrect to assume that the average signal power in each channel is increased. Instead, the true situation is that the signal-to-noise ratio in each channel is reduced by virtue of the split input signal and cannot again be increased over single channel operation once nonlinear rectification has taken place.

A more detailed discussion of the problem may be found.<sup>2</sup>

SAMUEL F. GEORGE  
Radar Div.  
U. S. Naval Res. Lab.  
Washington 25, D. C.

### Author's Comment<sup>3</sup>

I believe Mr. George has incorrectly followed my reasoning. The figures presented in my letter<sup>1</sup> concern only the comparison of my proposed alternate system to the system originally given in a previous article.<sup>4</sup> The basic difference between these two systems is that the incoming signal power is wasted one half the total time in the square wave Dicke system, while in the alternate system, this power is utilized by a second receiver. There is no question involved of splitting the received signal  $S(t)$  into more and more parts, but only the concept of utilizing the entire signal the entire time.

MARTIN GRAHAM  
The Rice Institute  
Houston, Tex.

<sup>2</sup> S. F. George, "Time Domain Correlation Detectors vs Conventional Frequency Domain Detectors," Naval Res. Lab., Washington, D. C., Rep. No. 4332; May 3, 1954.

<sup>3</sup> Received by the IRE, March 4, 1959.

<sup>4</sup> S. J. Goldstein, "A comparison of two radiometer circuits," Proc. IRE, vol. 43, pp. 1663-1666; November, 1955.

### The Cause of Aurora Borealis\*

The February, 1959 special issue of IRE PROCEEDINGS on the nature of the Ionosphere, calls to mind an item that may be of

\* Received by the IRE, March 12, 1959.

\* Received by the IRE, March 2, 1959. Hudson Labs. Contribution No. 40. This research was sponsored by the Office of Naval Research, Washington, D. C.

<sup>1</sup> W. B. Davenport, Jr., "Correlator Errors Due to Finite Observation Intervals," M.I.T. Res. Lab. of Electronics, Cambridge, Mass., Tech. Rep. No. 191; March 8, 1951.

\* Received by the IRE, February 12, 1959.

<sup>1</sup> M. Graham, "Radiometer circuits," Proc. IRE, vol. 46, p. 1966; December, 1958.

interest. One of the earliest uses of a vacuum tube was in connection with a study of auroral displays. During a survey of cathode ray tube applications made by the writer some thirty years ago, the following description of an early cathode ray tube was noted. It appeared in the second volume of the *Annual of Scientific Discovery* published in 1851, having been picked up from an earlier report in 1850.

ON THE CAUSE OF AURORA BOREALIS

With regard to the origin of the aurora borealis, it seems natural to attribute it to the electric field contained in the atmosphere, which at great heights, where the air is rarified, must become luminous, as under the receiver of an air-pump and in the barometer vacuum; this hypothesis would acquire a great probability if we succeeded in proving by direct experiments that magnetism exerts an influence upon electric light." This extract from a memoir by Morlet on the aurora borealis induced M. de la Rive to communicate the following experiment, showing the influence of magnetism on the light produced by ordinary electrical discharges. "I introduce into a glass globe, by one of the two tubulures with which it is furnished, a cylindrical iron bar, of such length that one of its extremities reaches nearly to the center of the globe, while the other extends a short distance out of the tubulures. The bar is hermetically sealed in the tubulure and covered throughout its length, except at its two ends with an insulating and thick layer of wax. A copper ring surrounds the bar above the isolating surface of its internal part the nearest to the side of the globe; from this ring proceeds a conducting rod, which, carefully isolated, traverses the same tubular as the iron bar, but without communicating with it, and terminates externally in a knob. When by means of a stopcock adjusted to the second tubulure of the globe, the air in it is rarified, the knob is made to communicate with one of the conductors of an electrical machine, and the external extremity of the iron bar with the other so that the two electricities unite in the interior of the globe, forming between the internal extremity of the iron bar and the copper ring which is at its base a more or less regular fascicle of light. But if the external extremity of the iron bar is placed in contact with one of the poles of a strong electromagnet taking good care to preserve the isolation the electric light takes a very different aspect. Instead of issuing, as before, from the different points of the surface of the terminal parts of the iron bar it is emitted only from the points which form the contour of this part so as to constitute a continuous luminous ring. This is not all; this ring and the luminous jets which emanate from it have a continuous movement of rotation around the magnetized bar, now in one direction, now in another, according to the electric discharges and the direction of the magnetization. Lastly, more brilliant jets appear to issue from this luminous circumference, without being confounded with those which terminate on the ring and from the fascicle. As soon as magnetization ceases, the luminous phenomena becomes again what it was previously, and what is generally in the experiment known as the electrical egg.—*Silliman's Journal*; May, 1850

RALPH R. BATCHER  
240-02 Forty-Second Ave.  
Douglaston, N. Y.

$$\psi \cong \psi_0 e^{-(\gamma^2 - k^2)^{1/2} z} \cong \psi_0 e^{-\gamma h} \quad (1)$$

In the above,  $k = 2\pi/\lambda$  where  $\lambda$  is the wavelength in air and  $\gamma$  is the complex propagation constant given by

$$\gamma = [i\mu\omega(\sigma + i\epsilon\omega)]^{1/2}$$

in terms of the angular frequency  $\omega$ , conductivity  $\sigma$ , permeability  $\mu$ , and permittivity  $\epsilon$ .

Eq. (1) is also valid in a local sense for a gently undulating surface if the local radius of curvature is large compared to both  $|\gamma^{-1}|$  and  $h$ . This statement, if not obvious, follows immediately from the work of Leontovich.<sup>2</sup>

In most cases of practical interest, the displacement currents are negligible compared with the conduction currents. Thus  $\sigma \gg \epsilon\omega$  and consequently

$$\gamma \cong (i\sigma\mu\omega)^{1/2} = (1 + i)/\delta$$

where  $\delta = (2/\sigma\mu\omega)^{1/2}$  is the classical "skin depth."

We now wish to calculate the field at any depth in a homogeneous conductor which has a gently undulating or wavy interface. For illustration, the interface between the conductor and the air above it is taken to be sinusoidal in form. This might be a fair representation for the ocean. This idealized situation is illustrated in Fig. 1. The waves are two dimensional in form so that, in terms of a rectangular coordinate system, the interface is described by

$$z = D \cos(2\pi x/L) = D \cos \beta x \quad (2)$$

where  $D$  is the amplitude,  $L$  is the period, and  $\beta = 2\pi/L$ .

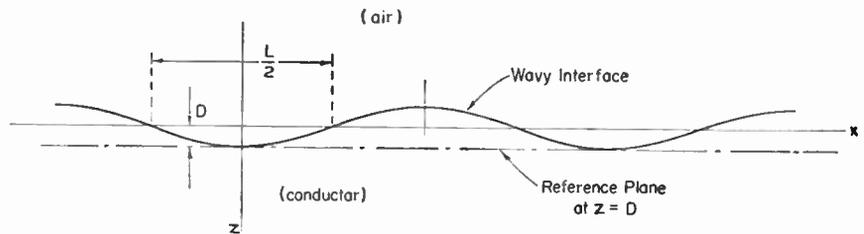


Fig. 1.

It is now assumed that the field in the air above the interface is of the form

$$\psi_0 e^{-ik_1 x} e^{-ik_2 z}$$

being a uniform plane wave traveling in a direction parallel to the mean interface, i.e., the plane  $z=0$ . This assumption is justified if the free-space wavelength  $\lambda$  is large compared to the period  $L$  of the waves and if the height of the waves is small compared to their period, i.e.,  $D \ll L$ . Under these assumptions the field  $\psi_r(x)$  along a reference plane  $z=D$  in the conductor can be written:

$$\psi_r(x, y) \cong \psi_0 e^{-ik_1 x} e^{-ik_2 D} e^{-\gamma D(1 - \cos \beta x)}$$

The latter exponential factor accounts for the attenuation of a wave traveling vertically, or near vertically, downward from the interface to the reference plane  $z=D$ . This concept is valid since  $D$  is always much

less than the radius of curvature of the interface.

The solution is facilitated by rewriting  $\psi_r(x, y)$  in the series form

$$\psi_r(x, y) = \psi_0 e^{-ik_1 x} e^{-ik_2 D} e^{-\gamma D} \sum_{n=-\infty}^{+\infty} I_n(\gamma D) e^{i\beta n x} \quad (3)$$

using an addition theorem<sup>3</sup> for modified Bessel functions  $I_n(\gamma D)$  of order  $n$  and argument  $\gamma D$ . Now the field  $\psi(x, y, z)$  is a solution of

$$\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \gamma^2 \right) \psi(x, y, z) = 0 \quad (4)$$

for  $z \geq D$  and, therefore, it is of the form

$$e^{i(\beta n - k_1)x} e^{-ik_2 y} e^{-\alpha_n z}$$

where

$$\alpha_n = [(\beta n - k_1)^2 + k_2^2 + \gamma^2]^{1/2}$$

Since  $\psi(x, y, z)$  is to reduce to  $\psi_r(x, y)$  for  $z=D$ , it easily follows that

$$\psi(x, y, z) = \psi_0 e^{-ik_1 x} e^{-ik_2 y} e^{-\gamma D} \times \sum_{n=-\infty}^{+\infty} I_n(\gamma D) e^{i\beta n x} e^{-\alpha_n(z-D)} \quad (5)$$

which is of the required form. It is immediately seen that, if  $D$  approaches zero,

$$\psi(x, y, z) \cong \psi_0 e^{-ik_1 x} e^{-ik_2 y} e^{-\gamma z} \quad (6)$$

as expected. In general, when  $D$  is finite, the field is no longer characterized by a single exponential term.

The general formula is simplified somewhat if it be noted that in cases of practical

The Calculation of the Field in a Homogeneous Conductor with a Wavy Interface\*

It is well known that electromagnetic waves are attenuated exponentially below the plane interface of a homogeneous conductor. For example, if radio waves are propagating along the surface of a flat well-conducting ground, the field component  $\psi$  at depth  $h$  is related to the same field component  $\psi_0$  at the interface by the simple relation<sup>1</sup>

interest  $\beta \gg k_1$  and  $k_2$ . This inequality is satisfied if  $\lambda \gg L$ . Then

$$\psi(x, y, z) = \psi_0 e^{-ik_1 x} e^{-ik_2 y} e^{-\gamma D} \times \sum_{n=0}^{\infty} \epsilon_n I_n(\gamma D) \cos(\beta n x) e^{-(\beta n)^2 + \gamma^2}^{1/2} (z - D) \quad (7)$$

where  $\epsilon_0 = 1$  and  $\epsilon_n = 2$  for  $n = 1, 2, 3$ . For computation<sup>4</sup> we can make use of the series formula

$$I_n(\gamma D) \cong \frac{(\gamma D)^n}{n! 2^n} + \frac{(\gamma D)^{n+2}}{(n+1)! 2^{n+2}} + \dots \quad (8)$$

which converges rapidly if  $|\gamma D| < 1$ . In fact, if  $|\gamma D| \ll 1$  or if  $D \ll \delta$ , the field in the conductor can be written

<sup>3</sup> N. W. McLachlan, "Bessel Functions for Engineers," Oxford University Press, New York, N. Y.; 1934.

<sup>4</sup> Also note that

$$I_n(\gamma D) \cong I_n(i^{1/2} Z) = \frac{1}{i^n} [\text{ber}_n Z + i \text{bei}_n Z]$$

where  $Z = |\gamma D| = \sqrt{2} D/\delta$  and where  $\text{ber}_n$  and  $\text{bei}_n$  are tabulated by McLachlan, *op. cit.*

\* Received by the IRE, December 16, 1958.  
J. R. Wait, "Theory of electromagnetic surface waves over geological conductors," *Geofisica Pura e Applicata*, vol. 28, pp. 47-56; 1954.

<sup>2</sup> M. A. Leontovich, "Approximate boundary conditions," in "Investigations on Radiowave Propagations," Printing House of Academy of Sciences, Moscow, U.S.S.R., pt. 11, pp. 5-12; 1948.

$$\psi(x, y, z) = \psi_0 e^{-ik_x x} e^{-ik_y y} \times [e^{-\gamma z} + (\gamma D) e^{-\gamma D} e^{-\sqrt{\gamma^2 + \beta^2}(z-D)} \cos \beta x] \quad (9)$$

+ higher order terms in  $|\gamma D|$ .

This clearly shows that the perturbation due to the ripples is proportional to the amplitude  $D$  of the waves. It is also noted that the first perturbation term varies sinusoidally with  $x$  following the form of the wavy interface.

This problem was suggested to me by A. D. Watt.

JAMES R. WAIT  
National Bureau of Standards  
Boulder, Colo.

### Phase Relationships at the Output of a Short Slot Hybrid Junction\*

Prompted by conflicting statements in the literature,<sup>1-3</sup> a study has been made of the phase relationships at the output ports of short slot couplers. From a consideration of the phase velocity of the modes established in the coupling region of the short slot hybrid, it can be shown that, with an input to port 2 (Fig. 1);

- 1) for the *sidewall* coupler, the output of the main guide (port 4) will *lead* the output of the auxiliary guide (port 3) by 90 degrees,
- 2) for the *topwall* coupler, the output of the main guide will *lag* the output of the auxiliary guide by 90 degrees.

This conclusion was expressed by Mattingly, *et al.*,<sup>3</sup> but contradicts Riblet<sup>1</sup> and Strom.<sup>2</sup>

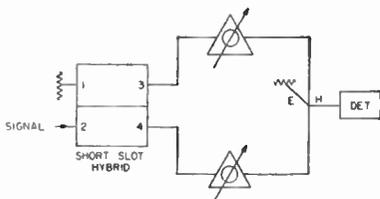


Fig. 1—Microwave bridge circuit for measuring phase relationships.

The findings have been verified by two experimental measurements. One method involved the use of the microwave bridge circuit shown in Fig. 1. With approximately equal lengths in the bridge arms (within 0.1 wavelength), excitation of port 2 of the short slot coupler will produce some output from the  $H$ -arm of the magic- $T$ . By changing the length in either arm of the bridge and observing the change in the output of the

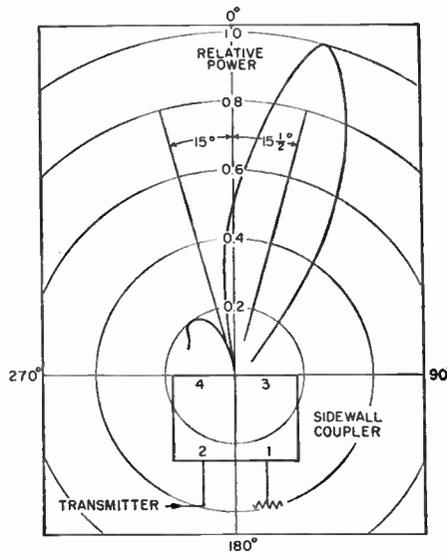


Fig. 2—Radiation pattern for sidewall coupler showing phase lead in main guide.

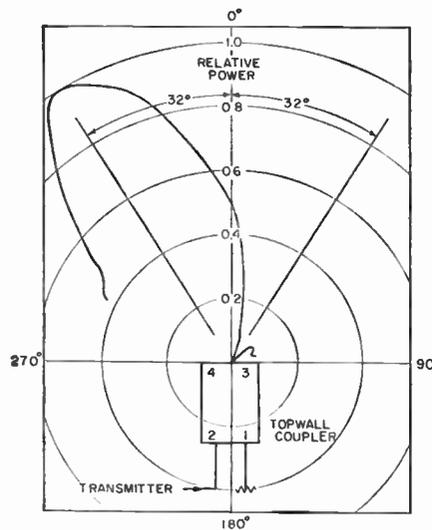


Fig. 3—Radiation pattern for topwall coupler showing phase lag in main guide.

magic- $T$ , the phase relationship of the outputs of the short slot coupler can be immediately deduced. This experiment shows that in-phase conditions at the magic- $T$  are produced when the length of the arm connected to port 3 is decreased for the sidewall coupler and increased for the case of the topwall coupler, indicating a lagging phase from the auxiliary arm (port 3) of the sidewall coupler and a leading phase from the same part of the topwall coupler.

Another experimental method, used to resolve the question of output phase relationships, required taking antenna patterns of an open-ended short slot coupler excited from one opposing port. In this case, the hybrid ports radiate as a two element array with a 90 degree phase difference between elements. This phase difference will shift the point of maximum intensity away from the array axis in the direction of the element which has the lagging excitation. The angle off axis will, of course, be determined by the array geometry and the frequency of operation. In addition to a point of maximum in-

tensity, a null is to be expected in the radiation pattern corresponding to the angle at which the signals from the two elements are 180 degrees out of phase. Trigonometric considerations require that, for a 90 degree phase difference, the angle of the null with respect to the axis should be the negative of the angle of the maximum also referred to the axis. The antenna patterns shown in Figs. 2 and 3 illustrate the shift which is to be expected from the phase relationships obtained analytically. In addition, the angular relationship of the maximum and minimum is clearly evident.

The absolute phase relationships have been found to be in contradiction to those of Riblet<sup>1</sup> and Strom<sup>2</sup> but, a rigorous analysis of various microwave circuits using both concepts has shown that the difference in solutions is primarily one of polarity. It should be noted that the relative difference in output phase is the important factor to be recognized in designing microwave circuits which utilize both types of couplers.

ROBERT D. TOMPKINS\*  
Radar Div.  
Naval Research Lab.  
Washington 25, D. C.

### Radio Reflections from Satellite-Produced Ionization\*

In the past few months a great deal of interest has been aroused concerning the effects of earth orbiting satellites on the upper atmosphere or ionosphere. It has been speculated that enough ionization could be produced by the motion of a satellite through the upper atmosphere to produce a plasma of ionized particles capable of returning electromagnetic waves to earth. If such a plasma exists, its critical frequency appears to be below the VHF region since various types of radar have thus far indicated cross sectional areas which are close to the skin dimensions of the tracked orbiting vehicles.

Several attempts have been made to prove (on a scientific basis) that such low frequency effects exist. Observers have reported both positive and negative results by making observations of signals received from continuously transmitting radio stations. The principal source of energy in several cases has been the 20 mc WWV transmitter.

In October, 1958, the authors under the sponsorship of General Electric's Light Military Electronics Department initiated a project for making such observations. However, the search for reflected signals was not limited to 20 mc; observations were also made on the 5, 10, and 15 mc WWV transmitters, which seemed to be well located for such use. From preliminary data obtained it appears that disturbances do exist which are correlated with satellite

\* Received by the IRE, April 15, 1959.

\* Received by the IRE, December 29, 1958.  
<sup>1</sup> H. J. Riblet, "The short-slot hybrid junction," *Proc. IRE*, vol. 40, pp. 180-184; February, 1952.  
<sup>2</sup> L. D. Strom, "Noise cancellation in microwave mixers," *Tele-Tech and Electronic Indus.*, vol. 15, pp. 186-188; March, 1956.  
<sup>3</sup> R. L. Mattingly, B. McCabe, and M. J. Traube, "The split reflector technique for broadband impedance matching of center-fed antennas without pattern deterioration," *IRE WESCON 1957 CONVENTION RECORD*, vol. 1, pt. 1, pp. 231-235.

passages. Through the use of the various frequencies, more than one type of correlated disturbance has been observed. On the 20 mc frequency a fast flutter phenomenon has been observed, as has been previously reported. However, at 15 mc, and especially at 10 mc, two very different types of disturbances have been found.

The first phenomenon observed at 10 mc differs from the flutter-type reflections obtained at 20 mc in that the signals received exhibit characteristics which change frequency. This frequency change is similar to that which would be obtained from a transmitter aboard the satellite itself and for this reason has been called a "Doppler" reflection. This reflection has been observed to produce an audible beat with the WWV carrier received by sky wave.

The Doppler-shifted reflected signal, when observed on a spectrum analyzer, is seen to consist of a smear of frequencies 50-200 cycles wide, moving as would be expected, from a few hundred cycles gradually through zero beat, and up again. As might be anticipated, observations indicate that the Doppler is heard more often at night. The density of the plasma partially surrounding the orbiting vehicles is probably enough for reflections at this frequency but it is obscured during the day because of ionospheric influence.

The second observed phenomenon is in the form of a delayed disturbance. This effect is a rough audio note containing components between 60 and 400 cycles. It is delayed approximately 8-10 minutes after known satellite passages (and after the Doppler disturbance). The frequency distribution remains relatively constant with time, and for this reason the effect has been called a "rumble." The duration of "rumble" varies from a few seconds to nearly one minute. There are several clues regarding the explanation of the rumble, but nothing conclusive as yet.

The effects described have been heard on 10 and 15 mc. However, generally speaking the disturbances observed on 15 mc have been much less frequent and considerably weaker. It has been found that the rumble is more consistent in its appearance than the "Doppler" and on the lower passes of the satellites it generally appears both day and night. As yet the rumble has never been observed previous to a satellite passage.

To date no extensive effort has been made to explain the Doppler and the rumble disturbances although a large number of observations have been made on satellites 1958 delta 1 and 2 (the two components of Sputnik III). The data which have been obtained have been compared with the scheduled passes of the satellites as well as with the Doppler signal taken directly from the 1958 delta 2 transmitter and are presently being analyzed more extensively.

An interruption in the study was recently made to apply the not yet proven technique on the lost satellite, 1959 beta (Discoverer I). This was done because the radio of the satellite was for all practical purposes unable to report the satellite's position, and the orbit plane was such that it was not visible over most of the area of the earth. Day and night listening periods were set up in Ithaca and Schenectady,

N. Y. Extremely good results were obtained using the 10 mc reflected signals. For a period of over 2 weeks consistent returns were obtained with a periodicity that correlated very closely with the period predicted from the launch data obtained from the Air Force. The rumble effect was very consistent. The Doppler returns, although not obtained on every pass of the satellite, definitely existed and at times lasted for up to 5 minutes.

The purpose of continuing the study now in progress is to determine the cause of the phenomenon. However, from the work accomplished thus far it appears that observed signal disturbances are definitely correlated with actual satellite passes.

C. R. ROBERTS  
P. H. KIRCHNER  
D. W. BRAY  
General Electric Co.  
Advanced Electronics Center,  
Cornell University  
Ithaca, N. Y.

10<sup>9</sup> high with respect to the frequency derived from the UT 2 second (provisional value) as determined by the US Naval Observatory. The atomic frequency standards remain constant and are known to be constant to 1 part in 10<sup>9</sup> or better. The broadcast frequency can be further corrected with respect to the USA Frequency Standard, as indicated in the table below. This correction is *not* with respect to the current value of frequency based on UT 2. A minus sign indicates that the broadcast frequency was low.

The WWV and WWVH time signals are synchronized; however, they may gradually depart from UT 2 (mean solar time corrected for polar variation and annual fluctuation in the rotation of the earth). Corrections are determined and published by the US Naval Observatory.

WWV and WWVH time signals are maintained in close agreement with UT 2 by making step adjustments in time of precisely plus or minus twenty milliseconds on Wednesdays at 1900 UT when necessary; no time adjustment was made during this month at WWV and WWVH.

NATIONAL BUREAU OF STANDARDS  
Boulder, Colo.

### WWV Standard Frequency Transmission\*

Since October 9, 1957, the National Bureau of Standards radio stations WWV and WWVH have been maintained as constant as possible with respect to atomic frequency standards maintained and operated by the Boulder Laboratories, National Bureau of Standards. On October 9, 1957, the USA Frequency Standard was 1.4 parts in

### The Need for Revision of Transistor Terminology and Notation\*

Congratulations to Armstrong for his excellent article on the need for revising transistor terminology and notation.<sup>1</sup>

Like Armstrong, I also have been teaching a course in transistor electronics and am in hearty agreement with most of his remarks. However, I am not quite in agreement with his statement that for quotation, the *h* parameters should be converted into whatever set will be most immediately useful to the user. How does a manufacturer know that the designer will be concerned only with the transmission parameters? Armstrong himself points out earlier in his article that the *Y* and *Z* matrices also are useful.

This brings up a closely related point: the failure of most manufacturers to publish adequate data on the transistors they make. Frequently one will receive a data sheet which shows a circuit designed around a given transistor, its temperature, voltage, and current limits and perhaps the *h* parameters at a particular operating point. This is all well and fine if the user is interested only in the particular circuit shown. Suppose he has another application in mind, however? Perhaps he is interested in linearity about some different operating point, or perhaps in low-frequency noise fluctuation or some other quantity.

Such added information assuredly will require more work on the part of manufacturers, but should net them correspondingly increased applications, which I am certain they will welcome. To me, adequate infor-

WWV FREQUENCY†

1959 March	Vs NBS‡ Atomic Standards 30-Day Moving Average Seconds pulses at 15 Mc	Vs Atomichron at WWV Measuring Time 1 Hour 2.5 Mc	Vs Atomichron at NRL Measuring Time 56 Minutes 2.5 Mc
1	-37	-40	
2	-37	-40	36
3	-37	-40	36
4	-36	-39	36
5	-35	-39	36
6	-35	-40	36
7	-34	-39	
8	-33	-39	
9	-33	-39	36
10	-32	-39	36
11	-32	-39	35
12	-32	-39	35
13	-32	-38	34
14	-32	-39	
15	-32	-39	
16	-32	-38	34
17	-32	-38	34
18	-32	-38	35
19	-32	-38	33
20	-33	-38	
21	-33	-38	
22	-32	-37	
23	-32	-38	35
24	-32	-38	35
25	-32	-38	35
26	-32	-38	35
27	-32	-38	34
28	-31	-38	
29	-31	-37	
30	-31	-37	35
31	-31	-38	36

\* Received by the IRE, April 24, 1959.

† WWVH frequency is synchronized with that of WWV.

‡ Method of averaging is such that daily adjustments appear in the data on the day they are made.

\* Received by the IRE, January 5, 1959.

<sup>1</sup> H. L. Armstrong, "On the need for revision in transistor terminology and notation," *Proc. IRE*, vol. 46, pp. 1949-1950; December, 1958.

mation is just as important as uniform terminology, if not more so.

To return to the problem of terminology and notation, I have observed also a lack of uniformity in sign conventions. These should be standardized along with the notation.

In conclusion, I agree with Armstrong that it is undesirable to try to standardize things too quickly in a new field. However, transistor electronics has advanced to the point that standardization is imperative—now.

BERNHARD E. KEISER  
RCA Laboratories  
Princeton, N. J.  
Formerly with Missouri Res. Labs., Inc.  
St. Louis, Mo.

Author's Comment<sup>2</sup>

Naturally, the author is happy to hear that Keiser agrees with many of his views on transistor notation and terminology. Perhaps if enough of us feel the same way, something may be done to end the present confusion.

Actually, I think we are in essential agreement. My quarrel with the *h* parameters is not with them as pieces of information—but rather with the notation. To present them as matrix elements, when in fact they will probably never be used in matrix analysis, seems useless. In particular, it requires two subscripts. If, then, one or two additional subscripts should be required to distinguish stages, operating conditions, etc., the result is quite a mess.

I could not agree more with Keiser on the frequent dearth of information on transistor types. The same situation seems to arise sometimes with new tube types. I have sometimes wished to use, in pulse circuitry, tubes such as sweep tubes and damper diodes, intended originally for use in television receivers. Only too often the only information readily available is that dealing with that use alone.

Certainly it should be safe to do some standardizing on transistor electronics now. My comments on this matter were intended to suggest that in some really new field, say cryotrons or masers, we should go very cautiously in trying to standardize terminology and notation.

May I add here one final word of explanation about my comment on the *B* and *H* fields? I am aware that recently a number of excellent articles and books have made the correct distinction here. Actually, my comments were prompted by exasperation at reading a proposed standard, which called the field *H* when it should have been *B*. After all, if the standards do not treat these matters correctly, who is to show the way?

H. L. ARMSTRONG  
Dept. of Physics  
Queen's University  
Kingston, Ont., Can.  
Formerly with Pacific Semiconductors,  
Culver City, Calif.

Semantics and Kirchhoff's Current Law\*

Kirchhoff's current law is usually stated thus: the sum of the currents entering a node equals the sum of the currents leaving; or the algebraic sum of the currents entering a node is zero.

Thus, if any part of a network is enclosed by a surface the algebraic sum of the currents entering the surface is zero. In the case of lumped elements, the set of branches crossing such a surface forms a cut set and if the network possesses a dual, the dual of the cut set is a tie set. A tie set is associated with a loop and this loop may be considered as a graphical analog of the tie set. The graphical analog of a cut set is the surface referred to above and the purpose of this note is to suggest a name for it.

Since a circuit diagram is usually drawn on a flat surface, the surface in question degenerates to a closed path that encloses a part of the network. It is, in fact, a circuit defined by the branches it crosses. An appropriate name for this circuit is ambit.<sup>1</sup> This is a perfectly good English word that means a circuit, circumference, or boundary. As applied to networks in general, an ambit is a closed surface that separates a network into two parts. It is sufficient for most purposes to think of it as a closed path formed by cutting branches. The branches crossed by an ambit are said to belong to the ambit or to be contained in it. For example, a set of branches contained in an ambit forms a cut set.

A few examples will illustrate the utility of the ambit concept. Evidently, Kirchhoff's current law may be stated—the sum of the currents entering an ambit equals the sum of the currents leaving it or, the algebraic sum of the currents entering an ambit is zero. Concerning duality, we may say that two graphs are duals if to every ambit in one there corresponds a loop in the other and vice versa.

The ambit notion is especially useful when writing equilibrium equations using as variables node-pair voltages other than node-to-datum voltages. Fig. 1 illustrates this point. Here a tree has been selected (solid lines) and the tree-branch voltages are used as variables. Ambits are selected in such a way that each contains only one tree branch. To write the equilibrium equations, a current-law equation is written for each ambit, expressing the branch currents in terms of the branch conductances and the appropriate voltages. Thus, for ambit one we have

$$E_1G_1 + (E_1 - E_2)G_5 + (E_1 - E_2 + E_3 + E_4)G_7 = J_1 \quad (1)$$

in which each term represents a current flowing away from the ambit. Similarly, for ambit two we obtain

$$E_2G_2 + (E_2 - E_1)G_5 + (E_2 - E_1 - E_3 - E_4)G_7 + (E_2 - E_3)G_8 = 0 \quad (2)$$

\* Received by the IRE, February 9, 1959.  
<sup>1</sup> This particular word was suggested by H. E. Meadows, Jr., Bell Telephone Labs., Murray Hill, N. J.

In this equation positive currents are considered to be flowing into the ambit. This is done in order to make the coefficient of *E*<sub>2</sub> positive. Now, collecting similar terms in these equations gives

$$E_1(G_1 + G_5 + G_7) - E_2(G_5 + G_7) + E_3G_7 + E_4G_7 = J_1 \quad (3)$$

$$-E_1(G_5 + G_7) + E_2(G_2 + G_5 + G_7 + G_8) - E_3(G_7 + G_8) - E_4G_7 = 0 \quad (4)$$

and the two remaining equations are readily found to be

$$E_1G_7 - E_2(G_7 + G_8) + (G_3 + G_6 + G_7 + G_8) + E_1(G_6 + G_7) = 0 \quad (5)$$

$$E_1G_7 - E_2G_7 + E_3(G_6 + G_7) + E_1(G_1 + G_6 + G_7) = -J_2 \quad (6)$$

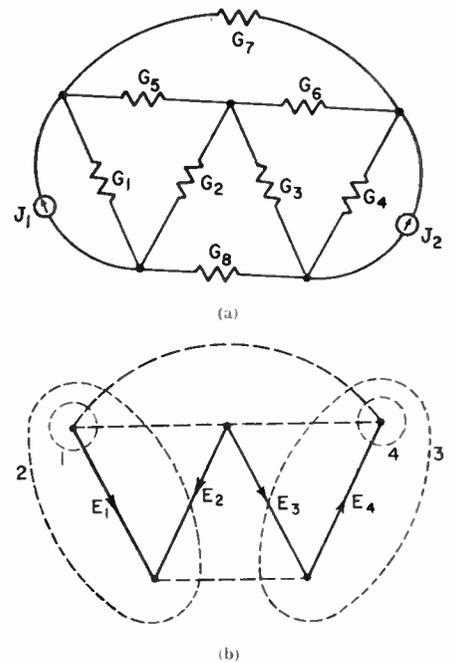


Fig. 1—A network and its graph showing ambits associated with tree branches. Reference arrows show positive directions for tree-branch currents. The *E*'s are the corresponding tree-branch voltages. *J*<sub>1</sub> and *J*<sub>2</sub> are current sources.

It is easy to devise a set of rules for writing the final equations by inspection. Noting that each equation is of the form

$$\sum_{k=1}^n G_{sk}E_k = J_s \quad (7)$$

it is apparent that *G*<sub>sk</sub> is the sum of all conductances common to ambits *s* and *k*. *G*<sub>ss</sub> is always positive. *G*<sub>sk</sub>, for *s* ≠ *k*, is positive if the reference arrows for *E*<sub>s</sub> and *E*<sub>k</sub> have the same direction along a tree-branch path from ambit *s* to ambit *k*; otherwise, it is negative. *J*<sub>s</sub> is positive if its direction is opposite to the current reference direction for *E*<sub>s</sub>; otherwise it is negative.

The rules are readily modified to include voltage as well as current sources, no source transformations being required. Also, if the

<sup>2</sup> Received by the IRE, January 30, 1959.

tree is selected so that its branches have a common node, the ambit equations are identical to the familiar node equations.

B. J. DASHER  
School of Elec. Eng.  
Georgia Inst. Tech.  
Atlanta, Ga.

**A Compact Antenna Switch for Scintillation Measurements\***

In meter wave radio interferometers modulation of discrete radio source signals may be obtained by alternately reversing the phase of the signal from one of the antennas. This is usually accomplished by introducing one half-wavelength of cable into one of the antenna leads. A more compact switch may be constructed if a phase shift of  $\pi/2$  radians is introduced into each of the antenna lines in turn. This method is not generally recommended since the parallel impedance of the two antennas may not be the same for each position of the switch resulting in changes in the internal noise power of the receiver at the switching frequency.<sup>1</sup> However, when recording scintillating signals from strong radio sources, the effect of such changes is small and virtually constant.

The switch, which was designed to be part of a portable scintillation recorder, is shown in Fig. 1. The two antenna cables are

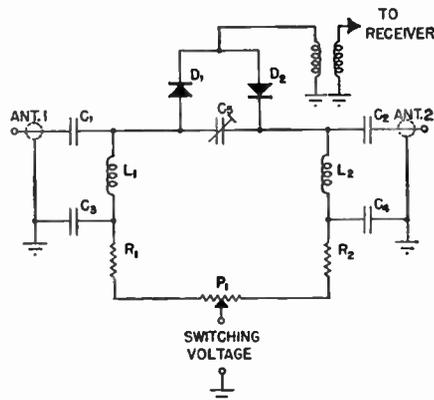


Fig. 1.

joined by an artificial quarter-wave section of line consisting of  $L_1$ ,  $L_2$ , and  $C_3$ , the receiver input being alternately connected to either end of this line by the switching diodes  $D_1$  and  $D_2$ . The capacitors  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$  are blocking capacitors whose reactance should be small at the signal frequency and large at the switching frequency. The resistors  $R_1$  and  $R_2$  limit the maximum forward current through the diodes and  $P_1$  is a balancing potentiometer which is adjusted for zero modulation of the background signal.

\* Received by the IRE, February 9, 1959.  
<sup>1</sup> M. Ryle, "A new radio interferometer and its application to the observation of weak radio stars," *Proc. Roy. Soc., pt. A*, vol. 211, pp. 351-375; March, 1952.

$D_1$  and  $D_2$  are 1N91 junction diodes. At 50 mc the 1N91 has a low resistance of the order of several ohms when biased in the forward direction. With reverse bias the diode exhibits a capacitive susceptance of about 5 millimhos. In the design, the forward resistance may be neglected and the reverse susceptance assumed to be part of  $C_3$ . The diodes should be relatively well matched although small differences may be compensated by adjustment of  $P_1$  and if necessary by trimming capacitors.

Measurements were made on a switch designed to operate at 50 mc. In the frequency range 40 mc to 60 mc, the phase shift obtained was found to be very closely that which would be produced by a corresponding length of coaxial cable, while the insertion loss was less than 1.5 db. The noise produced by the switch is low since the diodes are mismatched in both the forward and reverse directions.

The author wishes to thank G. A. Harrower, Director of the Queen's University Radio Observatory for providing the facilities for testing the above switch.

W. D. RYAN  
Dept. Elec. Eng.  
Royal Military College  
Kingston, Ont., Can.

**An Historical Note on Graphical Representation\***

It is always interesting and somewhat surprising to find similar mathematical tools in various fields of engineering and science. One of the most striking examples of this is obtained from the diagrams in Sommerville's book,<sup>1</sup> which was written in 1914 and recently republished. These diagrams are identical to the bicircular chart, the Smith Chart, and Deschamps' construction for the iconocenter of a loss circle. The bicircular chart is the representation of concentric non-euclidean circles all of which cut orthogonally a pencil of lines from a common vertex. The loci of constant VSWR are the concentric circles, and the loci of constant phase angles are the lines of the pencil centered at  $Z_0$ . The Smith Chart is similar (geometrically) to the bicircular chart except that the vertex of the pencil is moved to a point at infinity (the rim of the unit circle). Therefore, the constant resistance circles (or horocycles) are all centered at their point of intersection. The lines of the pencil from the center of the non-euclidean concentric circles represent the constant reactance lines of the Smith Chart.

Deschamps' construction for the iconocenter determines the non-euclidean center of the loss circle.

This short note is written to illuminate

\* Received by the IRE, February 25, 1959.  
<sup>1</sup> D. M. Y. Sommerville, "The Elements of Non-Euclidean Geometry," Dover Publications, Inc., New York, N. Y., pp. 177-179.

some of the interesting relationships between the graphical representations employed in microwave networks and those of the non-euclidean geometry.

It is felt pertinent to do so in view of the recent interest in the subject.

D. J. R. STOCK  
L. J. KAPLAN  
Dept. of Elec. Engrg.  
New York University  
New York, N. Y.

**Binary Multiplication in Digital Computers\***

One of the most basic arithmetic operations required from a Digital Computer is that of binary multiplication; yet, the multiplier in most computers does not actually go through the process of multiplying, but rather through a process of successive additions, the final sum of which is equal to the respective product. Actually, by using successive additions in order to multiply two numbers, the computer imitates the human "multiplier" in a somewhat primitive though faster way.

The human mind would multiply two binary numbers in the following fashion (Example 1): the multiplication of the multiplicand by each digit of the multiplier produces an "intermediate product." Such products are placed one underneath the other so that each "intermediate product" associated with a digit of the multiplier is below, and shifted one digit to the left, in respect to intermediate product associated with the next higher significant digit of the multiplier. Furthermore, if the digit in the multiplier is "1," its corresponding "intermediate product" is the exact duplicate of the multiplier, while when it is "0," its corresponding "intermediate product" is merely a row of zeroes. All the "intermediate products" are then added, and the result is the final product.

*Example 1*

multiplicand .....	110										
multiplier .....	011										
intermediate products .....	<table border="0" style="margin-left: 100px;"> <tr> <td>110</td> <td></td> </tr> <tr> <td>110</td> <td></td> </tr> <tr> <td>000</td> <td></td> </tr> <tr> <td colspan="2" style="border-top: 1px solid black;"></td> </tr> <tr> <td>10010</td> <td></td> </tr> </table>	110		110		000				10010	
110											
110											
000											
10010											
final product .....	10010										

The Digital Computer operates in a similar fashion, except that it shifts to the right the "intermediate product" corresponding to the lower significant figure of the multiplier, rather than shifting to the left the one associated with the higher significant figure. The basic difference is that the final addition is not done after all "intermediate products" are obtained, but an "intermediate addition" is done after each "intermediate product" is registered, the last one of such "intermediate sums" being the final product (Example 2).

\* Received by the IRE, December 29, 1958.



### Near-Zone Power Transmission Formulas\*

In his paper<sup>1</sup> the author gives some results of work on energy transmission from a transmitter to a receiver, previously published by me in a more general statement.

The essential results I obtained can be summarized as follows:

Let  $\vec{E}_1, \vec{H}_1$  be the complex electromagnetic field existing when unit power normalized wave is radiated by transmitter antenna (A). Let  $\vec{E}_2, \vec{H}_2$  be the complex field existing when unit power is radiated by receiver antenna, (B). The receiver antenna is linked to the detector by a transmission line matched at each end. The field in this receiver transmission line when unit power is radiated by (A) is the field of normalized wave propagating from antenna to the detector multiplied by a complex number  $\bar{T}_{12} = T_{12}e^{i\theta}$ , which is the transmission coefficient from (A) to (B).  $\frac{1}{2}T_{12}^2$  is the received power when a normalized wave is transmitted,  $\theta$  is the phase of the received wave. In the same way, the transmission coefficient from (B) to (A) can be defined.

I have shown that if  $S_2$  is a surface immediately close to the receiver,

$$\bar{T}_{12} = \frac{1}{4} \int_{S_2} (\vec{E}_1 \times \vec{H}_2 + \vec{H}_1 \times \vec{E}_2) \cdot d\vec{s}. \quad (1)$$

If  $S_1$  is a surface close to the transmitter

$$\bar{T}_{21} = \frac{1}{4} \int_{S_1} (\vec{E}_2 \times \vec{H}_1 + \vec{H}_2 \times \vec{E}_1) \cdot d\vec{s}. \quad (2)$$

\* Received by the IRE, February 10, 1959.  
<sup>1</sup> M.-K. Hu, "Near-zone power transmission formulas," 1958 IRE NATIONAL CONVENTION RECORD, vol. 6, Pt. 8, pp. 128-135.

If  $\vec{\epsilon}$  and  $\vec{\omega}$  are symmetric

$$\nabla \cdot (\vec{E}_1 \times \vec{H}_2 + \vec{H}_1 \times \vec{E}_2) = 0. \quad (3)$$

Therefore, in this case,

$$\bar{T}_{12} = \bar{T}_{21} = \frac{1}{4} \int_S (\vec{E}_1 \times \vec{H}_2 + \vec{H}_1 \times \vec{E}_2) \cdot d\vec{s}, \quad (4)$$

$S$  being any surface surrounding completely  $b$  without surrounding (A).

The preceding relationships are of great interest because the medium is not necessarily homogeneous as it is in the Kirchhoff and Kottler relationships. Eqs. (1) and (2) assume no restriction on the medium between (A) and (B). Eq. (4) allows heterogeneity of any kind if  $\vec{\epsilon}$  and  $\vec{\omega}$  are symmetrical. Huyghen's principle is a particular case of these very general and rigorous relationships.

Two important cases are encountered for which  $\vec{\epsilon}$  and  $\vec{\omega}$  are not symmetrical; therefore,  $\bar{T}_{12} \neq \bar{T}_{21}$ .

In electron tubes it is found that

$$\bar{T}_{12} - \bar{T}_{21} = \int_V (\vec{E}_1 i_2 - \vec{E}_2 i_1) \cdot d\vec{v},$$

$V$  being the volume between (A) and (B),  $i_1$  and  $i_2$  being the currents produced by normalized waves transmitted by (A) and (B).

In the ferrite case,

$$\bar{T}_{12} - \bar{T}_{21} = \frac{K\omega\mu_0}{2} \int_V \vec{H}_0 \cdot (\vec{H}_1 \times \vec{H}_2) \cdot d\vec{v},$$

$\vec{H}_0$  being the unit vector in the direction of the applied magnetic field. This last relationship has also been found by Heller of the Lincoln Laboratory, Massachusetts Institute of Technology.

I have found relationships giving the

variation of  $\bar{T}$  with respect to frequency. If  $\vec{\epsilon}$  and  $\vec{\omega}$  are symmetrical, this relationship has a simple form:

$$\frac{\delta \bar{T}}{\delta \omega} = \frac{1}{2} \int_V (\vec{E}_1 \cdot \vec{\epsilon} \cdot \vec{E}_2 - \vec{H}_1 \cdot \vec{\omega} \cdot \vec{H}_2) \cdot d\vec{v}.$$

I used (4) and made the first demonstration in the early months of 1956, in the case of a surface wave problem. Reports have been given to the Air Research and Development Center of Cambridge, Mass. (AF Contracts Nos. 61-514-933 and 61-514-1149), which sponsored this work, in July, 1956; November, 1956; May, 1957; and November, 1957. I gave a speech on this topic, in September, 1956, at the "Colloque international sur la propagation," in Paris, France. A paper summarizing the speech has been given to *Onde Electrique* in June, 1957, and published in December, 1957. Other speeches have been given in October, 1957, at the "Congrès international des circuits et antennes hyperfréquences," and in October, 1958, at the Meeting on Propagation, in Liege, Belgium. On this same topic, a paper entitled "Compte rendu à l'Académie des Sciences" was published on August 12, 1957.

These theorems can be considered very general ones in physics. They can be extended to other fields than electromagnetic theory; I applied them to wave mechanics and fluid mechanics. In electromagnetics I have used them to develop a theory of surface wave and propagation problems, particularly concerning tropospheric scatter propagation. A synthesis of all the work will be published in the *Annales de Radio-électricité*.

JEAN ROBIEUX  
Dept. de Physique Appliquée  
C.S.F., France

## Contributors

Leonard E. Alsop was born on April 10, 1930, in Oroville, Calif. He received the B.A. degree from Columbia College, New York, N. Y., in June, 1951, and the M.A. degree in physics from Columbia University, New York, N. Y., in 1956. From 1951 to 1954, he served on active duty with the U. S. Navy as a communication's officer. At present, he is a research assistant in the Columbia Radiation Laboratory at



L. E. ALSOP

Columbia University.  
Mr. Alsop is a member of the American Physical Society and of Sigma Xi.



Lloyd V. Berkner (A'26-M'34-SM'43-F'47), for a photograph and biography, please see page 338 of the February, 1959 issue of PROCEEDINGS.



Robert W. Berry was born in Atlanta, Ga., on October 27, 1928. In 1950, he received the B.S. degree in chemistry with high honors from The Clemson Agricultural College, Clemson, S. C. He received the Ph.D. degree in chemistry from Michigan State University, East Lansing, in 1956.



R. W. BERRY

From 1951 to 1953, he served as a lieutenant in the Chemical Corps, with

duties as a physical chemist at the Chemical and Radiological Laboratories. He joined the staff of the Bell Telephone Laboratories, Murray Hill, N. J., in 1956 as a member of the technical staff in the component development department.

Dr. Berry is a member of the American Chemical Society, Sigma Xi, Sigma Pi Sigma, and Phi Kappa Phi.



E. Folke Bolinder (A'50-M'55) was born at Uppsala, Sweden, on August 11, 1922. He received the Civilingenjör degree in electrical engineering in 1945 and the degree of Teknologie Licentiat in 1954, both at the Royal Institute of Technology, Stockholm.

After a year of military service as a special engineer in the Swedish Air Force, where he worked on UHF antennas and cir-

culits, he worked from 1946 to 1951 on different microwave and pulse technique projects for the Swedish defense, combining



E. F. BOLINDER

the work with graduate studies at the Royal Institute of Technology. In 1951 he became a Fellow of the Sweden-America Foundation under whose auspices he worked (during 1952 and 1953) in transient synthesis as research assistant at the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge. During the winter of 1954-1955, he carried out research at the Instituto Nacional de la Investigación Científica, Mexico City. From the summer of 1955 to the end of 1957, he was a research staff member of the Research Laboratory of Electronics, M.I.T., working with geometrical methods in the microwave field. Since the beginning of 1958, he has had the position of physicist at the U. S. Air Force Cambridge Research Center, Bedford, Mass.

Mr. Bolinder is a member of RESA.



D. G. Brennan (A'48-M'51-SM'57) was born in Waterbury, Conn., on April 9, 1926. From 1942 to 1944, he was engaged in



D. G. BRENNAN

broadcast and communication system engineering and from 1944 to 1947, he was with the U. S. Army Signal Corps, principally as Chief of Communications, Southern Islands Command, Southwest Pacific Theater. From 1947 to 1949, he was chief engineer of Radio Station WWCO, Waterbury, Conn. In 1949-1951, he was an instructor at the Ward School of Electronics of the University of Hartford, Conn. He has been at the Massachusetts Institute of Technology, Cambridge, Mass., since 1951, and the M.I.T. Lincoln Laboratory since 1953, where he has conducted ionosphere research, electronic circuit and system design, and mathematical research. He is currently engaged in basic research in mathematics, and some occasional teaching in the M.I.T. Department of Mathematics, while continuing his association with the Lincoln Laboratory. He was a Gerard Swope Fellow during the 1955-1956 academic year.

He is a registered professional engineer in Connecticut, a member of Sigma Xi and the Institute of Mathematical Statistics, and an Institutional Member of the American Mathematical Society.



James F. Gibbons was born on September 19, 1931, in Leavenworth, Kan. In 1953, he received the B.S. degree in electrical

engineering from Northwestern University, Evanston, Ill., where he was a co-winner of the Eshback Award for the outstanding engineering graduate.



J. F. GIBBONS

He received the M.S. degree in electrical engineering in 1954, and the Ph.D. degree in 1956 at Stanford University, Stanford, Calif. He was a National Science Foundation Fellow from 1953-1955 and an RCA Fellow in Electronics in 1956. After completion of his

graduate studies, he was awarded a Fulbright Fellowship to Cambridge University, Cambridge, Eng., for the year 1956-57, where he studied nuclear magnetic resonance. Since 1957, he has been an assistant professor of electrical engineering at Stanford University; his major interest being in solid-state devices. During the year 1957-58, he also worked half-time with the Shockley Transistor Corporation, where he is presently a consultant.

Dr. Gibbons is a member of Phi Eta Sigma, Pi Mu Epsilon, Eta Kappa Nu, Tau Beta Pi, Sigma Xi, and an honorary member of the Western Society of Professional Engineers.



Joseph A. Giordmaine was born in Toronto, Ontario, Can. on April 10, 1933. He received the B.A. degree in physics and chemistry at the University of Toronto in 1955, and the A.M. degree in physics at Columbia University, New York, N. Y., in 1957.



J. A. GIORDMAINE

Since 1955, he has been research assistant and Esso Fellow at the Columbia Radiation Laboratory, working in the fields of solid-state and maser physics, and radioastronomy.

Mr. Giordmaine is a member of the American Physical Society, Sigma Xi, and the American Astronomical Society.



Robert D. Gold (A'54-M'58) was born on June 21, 1931, in Brooklyn, N. Y. He received the B.S. and M.S. degrees in electrical engineering from the City College of New York, N. Y., in 1953 and Cornell University, Ithaca, N. Y., in 1957, respectively.



R. D. GOLD

He spent three months with the RCA Victor Division before entering the U. S. Army in 1953. Assigned to the Electronic Warfare De-

partment under the Scientific and Professional Personnel Program, he worked on feasibility studies of radar countermeasure systems. He joined the RCA Laboratories, Princeton, N. J., in 1957, where he has worked on television circuits and kinescopes, and in the field of semiconductor devices. He is currently a lecturer in electrical engineering at the City College of New York.

Mr. Gold is a member of Tau Beta Pi and Eta Kappa Nu.



Douglas J. Hamilton (A'53) was born in Canton, Ohio, on December 6, 1930. He attended the College of Wooster, Wooster, Ohio. He received the B.S. degree from Case Institute of Technology, Cleveland, Ohio, in 1953 and the M.S. degree from the University of California, Los Angeles, in 1956.



D. J. HAMILTON

From 1953 to 1957, he was employed by Hughes Aircraft Co., Culver City, Calif., where he was engaged in the development of transistor circuitry for airborne digital computers. In 1957, he joined the General Electric Company Computer Laboratory, Palo Alto, Calif., where he was engaged in the development of transistor circuitry for magnetic core and magnetic drum memories. Since 1958, he has been a research assistant at Stanford Electronics Labs., Stanford University, Stanford, Calif., where he is studying for the doctorate.

Mr. Hamilton is a member of Eta Kappa Nu, Tau Beta Pi, and an associate member of Sigma Xi.



Cornell H. Mayer (M'47) was born in Ossian, Ia., on December 10, 1921. He received the B.S. degree from the State University of Iowa, Iowa City, in 1943, and the M.S. degree in electrical engineering from the University of Maryland, College Park, in 1951. From 1943 to 1948, he did design and development work on microwave receivers and antennas at the U. S. Naval Research Laboratory in Washington, D. C., and, since 1948, has been engaged in research in radio astronomy.



C. H. MAYER

Mr. Mayer is a member of the American Astronomical Society, URSI, and the Scientific Research Society of America.



H. T. McAleer (S'51-A'54-M'59) was born in Boston, Mass., on August 17, 1930. He received the B.S. and M.S. degrees, in

electrical engineering, in 1953, from the Massachusetts Institute of Technology, Cambridge, Mass.



H. T. McALEER

In 1953, he joined the General Radio Company in Cambridge, Mass., as a development engineer in the Signal Generator Group, concerned with the problems of standard frequency generation in the UHF range. From 1954 to 1956, he served as an officer in the U. S. Army

Signal Corps, at the Signal Corps Engineering Laboratories in Fort Monmouth, N. J. His work there consisted of research and development on crystal filters, and frequency measurement and synthesis equipment. In 1956, he rejoined the General Radio Company as a development engineer in the Frequency Control Group, concerned with pulse and digital equipment and the general problems of frequency control.

Mr. McAleer is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.



Louis Pensak (A'42-SM'47) was born on June 16, 1911, in New York, N. Y. He received the B.S. degree from Long Island University, Brooklyn, N.Y., in 1932 and spent the next year as a graduate assistant at the University of Pittsburgh, Pittsburgh, Pa. In 1936, he received the M.S. degree in physics from New York University, N. Y.



L. PENSACK

He joined the cathode-ray tube development section of RCA in 1937 and in 1940, was transferred to the Research Department of RCA Victor Division. Since 1941, he has been with RCA Laboratories, Princeton, N. J., where he has worked on kinescopes and storage tubes and solid state devices.

Mr. Pensak is a member of Sigma Xi.



John R. Pierce (S'35-A'38-SM'46-F'48), for a photograph and biography, please see page 455 of the March, 1959 issue of PROCEEDINGS.



Charles Ferencz Pulvari (SM'54) was born in Karlsbad, Austria-Hungary in July, 1907. He received the diploma in electrical and mechanical engineering at the University of Technical Sciences in Budapest, Hungary.

After graduation, he was associated with the Research Laboratories of the Telephone Manufacturing Company in Budapest for three years. He then joined the Hungarian Radio and Communications Corporation as a chief engineer and later became technical

managing director. In 1937, he became a consultant on television for Scophony, Limited, London. After the Second World



C. F. PULVARI

War, he organized his own research laboratory in the switching and communications field and began research in solid-state physics. He came to the United States in 1949 and joined George Washington University, Washington, D. C., where he pursued research on electro-static information storage devices. In 1951, he became a member of the staff of the Catholic University of America, Washington, D. C., and is presently a professor of electrical engineering and head of the Solid-State Research Laboratory. He has contributed to the development of ferroelectric materials and their various applications and holds numerous U. S. patents.

Prof. Pulvari is a member of Sigma Xi.



William Shockley (SM'51-F'55), for a photograph and biography, please see page 1314 of June, 1958 issue of PROCEEDINGS.



Doris J. Sloan was born in Summit, N. J., on January 9, 1936. In 1957, she received the B.S. degree in chemistry from Mount



D. J. SLOAN

American Chemical Society.



Martin C. Steele was born on December 25, 1919, in New York, N. Y. He received the B.S. degree in chemical engineering in 1940 from Cooper



M. C. STEELE

Union Institute of Technology, New York, N. Y. He received the M.S. and Ph.D. degrees in physics in 1949 and 1952, respectively, from the University of Maryland, College Park. During the years 1942 through 1946, he was in the U. S. Army. After his discharge, he worked at the Naval Research Laboratory, Washington, D. C., as a research physicist in the field of solid state physics. His major interests there were in superconductivity and magnetic properties of metals at low temperatures. He was the head of the Cryomagnetic

Research Section at the Naval Research Laboratory. Since 1955, he has been at RCA Laboratories, Princeton, N. J., doing research in semiconductor physics.

Dr. Steele is a member of Sigma Xi and the American Physical Society.



Charles H. Townes (SM'58) was born in Greenville, S. C., on July 28, 1915. He received the B.A. degree in modern languages and the B.S. degree in physics from Furman University, Greenville, in 1935, the M.A. degree in physics from Duke University, Durham, N. C., in 1937, and the Ph.D. degree in physics from the California Institute of Technology, Pasadena, in 1939.



C. H. TOWNES

From 1939 to 1947, he was a member of the technical staff of the Bell Telephone Laboratories. At Columbia University, New York, N. Y., he was an associate professor of physics from 1948 to 1950, Director of the Columbia Radiation Laboratory from 1950 to 1952, and Chairman of the Physics Department from 1952 to 1955. Since 1950, he has been a professor of physics.

His research has been in the fields of microwave spectroscopy, hyperfine effects, nuclear moments and structure, molecular structure, atomic time standards, electronic devices, masers, and radioastronomy.

Dr. Townes is a member of the Scientific Advisory Board of the Air Force, the Council of the American Physical Society, the National Academy of Sciences, and the American Academy of Arts and Sciences. He was the recipient of the Research Corporation Annual Award for 1958 and co-recipient of the Morris Liebmann Memorial Prize of the IRE for 1959.



Jerald A. Weiss was born in Cleveland, Ohio, on June 9, 1922. After some undergraduate training and industrial work in chemistry, he received the B.A. and M.A. degrees in 1949, and the Ph.D. degree in physics in 1953 at Ohio State University, Columbus.



J. A. WEISS

He joined the faculty of the University of Wyoming, Laramie, in 1949, as instructor in mathematics, and in 1951 returned to Ohio State as a graduate student. In 1953, he joined the Bell Telephone Laboratories, Murray Hill, N. J., in work relating to microwave magnetic materials and devices.

Dr. Weiss is a member of the American Physical Society, the Mathematical Association of America, Phi Beta Kappa, and Sigma Xi.

# Scanning the Transactions

**Information theory** unquestionably is regarded as one of the great conceptual advances of the radio engineering field. At the same time, quite a few authorities have expressed doubt from time to time as to whether much of the basic work being done in this field is leading to any worthwhile results. In a recent TRANSACTIONS editorial, a well-informed "outsider" from the field of mathematics had these provocative observations to make on this fundamental question:

"In spite of all the suggestive work by Wiener, Shannon, and their successors, the main thing that strikes an outsider is that there are so few theoretical results. In fact almost every time a writer proves an assertion connecting the capacity of a channel with the entropy of a source, his paper  $P_n$  is succeeded by a paper  $P_{n+1}$  which, instead of generalizing or extending the results of  $P_n$ , is devoted to pointing out and correcting some defect or insufficiency in it. The paper  $P_{n+1}$ , in its turn, receives the same harsh treatment, and so on. Moreover, in this presumably convergent process of purging and purifying, the theorems become more and more attenuated and inapplicable as their hypotheses become more restrictive.

"Even more extraordinary is the fact that this process of organizing what seems to be the very basis of the subject seems to have no effect whatever on its applications! Can it be that the existence of a mathematical basis is irrelevant, and that the basic principle is the very idea that there is a context in which the word 'information' is accepted by general agreement and used in an intuitive way, and that no more is needed?" (J. L. Doob, IRE TRANS. ON INFORMATION THEORY, March, 1959.)

**Public telephones on airplanes** may soon become an established service for air travelers. Two Bell System operating companies have been conducting commercial trials since September, 1957, with very satisfactory results. These trials are using 15-watt 450-mc FM airborne transmitters, with interconnection to the regular telephone network being provided by two land base stations, one in Chicago and the other near Detroit. Thus, progress in the telephone system has, so to speak, reached new heights. (R. V. Crawford, "Transmission tests on a trial system of telephone service for aircraft," IRE TRANS. ON VEHICULAR COMMUNICATIONS, April, 1959.)

**Body-snatching.** The increased demand for scientists and engineers over the past few years has made the technical recruiter a frequent and familiar sight on college campuses and at technical meetings throughout the country. He has also become a frequent and familiar target for criticism in the press. *Time* magazine once summed up an engineering convention with the caustic caption, "Recruiting? Forbidden. Piracy? Okay." More recently, the *New York Times*, perhaps moved by the sudden influx of employment ads, took notice of the opening of another meeting with the words, "The annual body-snatching of electronics engineers . . . began yesterday."—a salvo that was quoted a week later on the editorial page of *The Nation*. The prevailing impression seems to be that recruiters live only by the law of the jungle and that engineers are either helpless prey or, at least, can always be had for a price. Midst all these slings and arrows it is, therefore, a refreshing and welcome change to find recruiting spoken of, for once, as a respectable pursuit carried on by intelligent, responsible people. It is also reassuring to find that a recent survey of 20,000 engineers and scientists reveals that of 15 factors contributing to job satisfaction, they placed the following at the top of the list, in order of importance:

- 1) Interest potential of work.
- 2) Integrity of management.
- 3) Opportunity to discover and do creative work.
- 4) Opportunity to move up in the organization.
- 5) Caliber of supervision.
- 6) Living conditions.
- 7) Pay.

Evidently, engineers and scientists are not as materialistic as some would have us believe, and job satisfaction and professionalism are still strong motivating factors. (J. Kurshan, "Recruiting the technically creative—a dual responsibility," IRE TRANS. ON ENGINEERING MANAGEMENT, March, 1959.)

**Should we say "data" or "information"?** Probably most of us can safely use the two words interchangeably. There is, however, a distinction in meaning which becomes important to preserve when speaking of data-handling or information-processing systems. "Data" are the raw material from which "information" is derived. The differentiation is important only when there exists a desire to know *specific* facts, in contrast to a situation where any kind of facts are interesting as long as it is news. Thus, we can say that communicated facts are 1) "data," if they are not called for and do not contain any "news" value for the receiver at that time, 2) "news," if it is not called for but arouses an interest, or 3) "information," if it has been called for and increases the knowledge of the interrogator at that time. (E. A. Keller, "Data or information," IRE TRANS. ON INDUSTRIAL ELECTRONICS, April, 1959.)

**Communication between machines,** and between men and machines, has grown in volume almost to the point where it exceeds the amount of communication traffic between humans, so rapid have been the strides in digital data translation and transmission techniques. The most recent development in this fast-moving area is a transistorized special purpose computer, called MAUDE, that can decode hand-sent Morse code messages. MAUDE has thus far successfully decoded between 90 per cent and 95 per cent of the messages sent her by 184 operators, and its developers feel she can be a practical addition for a site with heavy traffic. Moreover, MAUDE has shed light on the interesting question of the extent to which a decoder (either a man or a machine) needs to know the language being encoded, in addition to knowing the code itself. It was found that MAUDE, even though she knows Morse code, would need to be taught more sophisticated language rules, including a word vocabulary, before she could be as good a decoder as a man. (B. Gold, "Machine recognition of hand-sent Morse code," IRE TRANS. ON INFORMATION THEORY, March, 1959.)

**A new twist in memory devices** makes use of a magnetic effect first observed a century ago involving twisted magnetic rods. Although the new element, appropriately called the "Twister," is still in the experimental stage, it has shown great potentialities for the future. In fact, it is expected that the Twister will someday probably replace many magnetic core and drum types of memories in computers and electronic switching systems where a rapid-access, high-capacity memory is necessary. The Twister makes use of the fact that a magnetic rod or wire under a mechanical torsion will produce a voltage between its ends when it is magnetized by a field acting along its longitudinal axis. It is interesting to note that in 100 years this is probably the first important use that has been found for this phenomenon. (D. A. Ellerbruck, "A new memory device—the Twister," IRE TRANS. ON COMPONENT PARTS, March, 1959.)

Graduate degrees in engineering were first awarded in the U. S. in 1873. Since then, graduate student enrollments have greatly increased, particularly in the last two decades, so that today well over 100 institutions annually award more than 1000 advanced degrees in electrical engineering alone. During this period the conditions under which graduate work is given have undergone considerable change. The Master's degree, for instance, which once was exclusively a research degree, has had to be modified to include advanced design as a proper subject for a thesis, in order to accommodate the greatly enlarged number of students pursuing this degree. In fact, some institutions today do not even require a thesis for a Master's degree. The advent of government and industry-sponsored research in universities has had a pronounced effect on graduate study programs, as have the off-campus graduate programs which an increasing number of companies now offer their technical employees. These and many other aspects of present-day graduate study were the subject of a unique and

fruitful conference of electrical engineering educators, held last summer by Syracuse University and the IRE Professional Group on Education. The meeting gave educators an unusual opportunity to clarify their understanding of these highly important matters, and has given the PGE an opportunity to publish a record of their discussions for the further enlightenment and benefit of our profession. (Sagamore Conference on Electrical Engineering Education, IRE TRANS. ON EDUCATION, April, 1959.)

Electroluminescence, despite being a subject of only fairly recent interest, has been written about extensively in the technical literature. Because of both this recent interest and the voluminous literature, a recent bibliography on the subject is worthy of special note. The number of references contained in it is a surprising 720, attesting to the thoroughness with which this valuable document was compiled. (H. F. Ivey, "Bibliography on electroluminescence and related topics," IRE TRANS. ON ELECTRON DEVICES, April, 1959.)

## Report of the Secretary—1958

TO THE BOARD OF DIRECTORS  
THE INSTITUTE OF RADIO ENGINEERS, INC.

Gentlemen:

Again comes the Secretary's Report for the year 1958. General growth continues, membership being up from 64,773 at the end of 1957 to 71,361. (See Fig. 1 and Table I.) The percentage of voting members to the total has risen. In fact, five years ago 28% of the total were voting members whereas now they number 59%. The number of Students has gone up 11%. Over the past five years Students have risen in number from 5556 to 14,961.

Six new Sections bring the total number to 100.

Your Secretary notes with satisfaction the 18% growth of the IRE publications as to editorial pages, the improved format of the PROCEEDINGS, the continued widespread activities of the Technical Department, the increases in size and activities of the Professional Groups, the larger Annual Convention, and the bigger figures displayed on the financial statements.

Respectfully submitted,



Haraden Pratt  
Secretary

January 31, 1959

### Fiscal

A condensed summary of income and expenses for 1958 is shown in Table II, and a balance sheet in Table III.

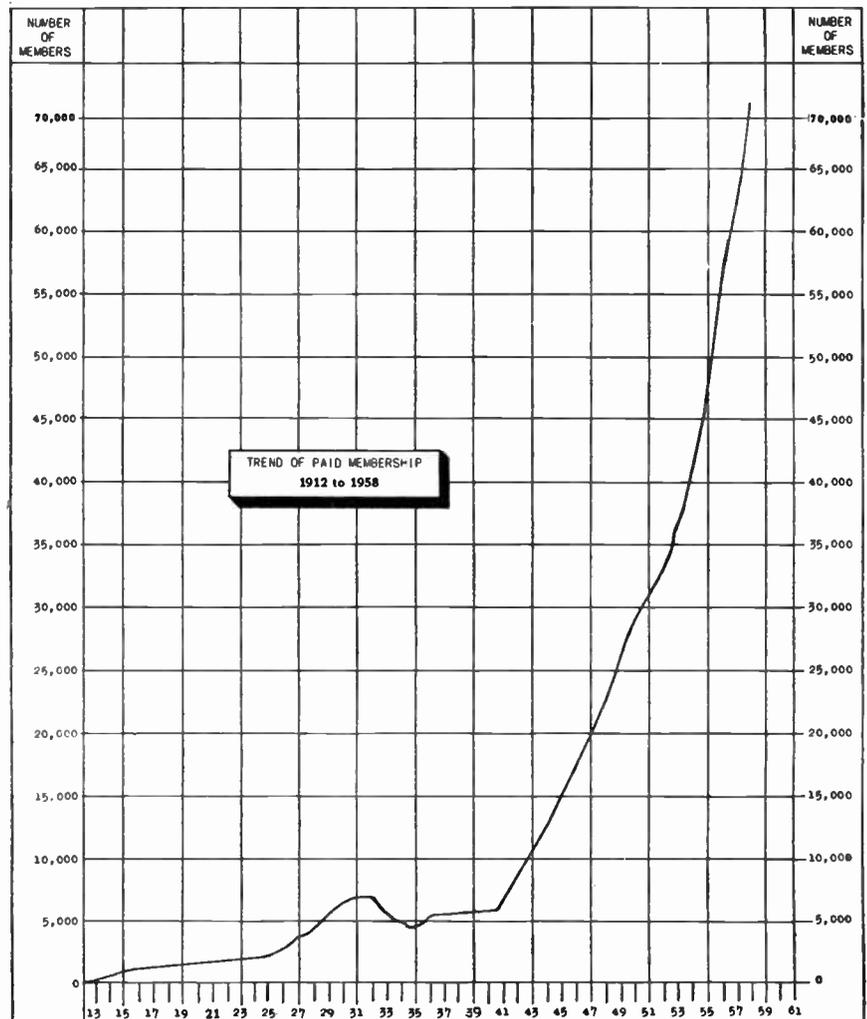


Fig. 1

TABLE I  
COMPARISON OF TOTAL MEMBERSHIP DISTRIBUTION BY GRADES, 1956-1958

Grade	As of Dec. 31, 1958		As of Dec. 31, 1957		As of Dec. 31, 1956	
	Number	% of Total	Number	% of Total	Number	% of Total
Fellow	770	1.1	700	1.1	635	1.1
Senior Member	8,536	12.0	7,685	11.9	6,486	11.6
Member	32,373	45.4	26,115	40.3	19,110	34.6
Associate	14,721 †	20.6	16,827 †	25.9	18,879*	34.0
Student	14,961	20.9	13,446	20.8	10,384	18.7
TOTALS	71,361		64,773		55,494	

\* Includes 388 Voting Associates.  
† Includes 352 Voting Associates.  
‡ Includes 330 Voting Associates.

TABLE II  
SUMMARY OF INCOME AND EXPENSE, 1958

<i>Income</i>		
Advertising		\$1,243,535
Member Dues and Convention		1,395,384
Subscriptions		179,631
Sales Items: Binders, Emblems, etc.		167,195
Investment Income		39,181
Miscellaneous Income		1,783
TOTAL INCOME		\$3,026,712
<i>Expense</i>		
PROCEEDINGS Editorial Pages		\$ 486,036
Advertising Pages		626,080
Directory		249,359
Section Rebates		71,820
Student Program		105,292
Professional Group Expense		182,407
Sales Items		122,499
General Operations		501,520
Convention Cost		412,779
TOTAL EXPENSE		\$2,757,792
Reserve for Future Operations—Gross		\$ 268,920
Depreciation		22,344
Reserve for Future Operations—Net		\$ 246,576

TABLE III  
BALANCE SHEET—DECEMBER 31, 1958

<i>Assets</i>		
Cash and Accounts Receivable		\$ 645,997
Inventory		34,244
TOTAL CURRENT ASSETS		\$ 680,241
Investments at Cost		1,490,478
Building and Land at Cost		912,317
Furniture and Fixtures at Cost		241,989
Other Assets		77,786
TOTAL		2,722,570
TOTAL ASSETS		\$3,402,811
<i>Liabilities and Surplus</i>		
Accounts Payable		\$ 72,519
TOTAL CURRENT LIABILITIES		\$ 72,519
Deferred Income		879,073
Professional Group Funds on Deposit		217,822
TOTAL LIABILITIES		1,169,414
Reserve for Depreciation		81,112
Surplus Donated		595,287
Surplus		1,556,998
TOTAL SURPLUS		2,152,285
TOTAL LIABILITIES AND SURPLUS		\$3,402,811

**Editorial Department**

*PROCEEDINGS OF THE IRE*

Several important changes were adopted by the Editorial Board during 1958 in order to increase the value and utility of the PROCEEDINGS to the membership. A new monthly feature, called "Scanning the Transac-

tions," was inaugurated with the February issue in order to single out and interpret for the general reader some of the significant or unusual work that is reported in IRE publications he may not normally see, i.e., the TRANSACTIONS, the NATIONAL and WESCON CONVENTION RECORDS, and the STUDENT QUARTERLY. In April the table of

TABLE IV  
VOLUME OF PROCEEDINGS PAGES

	1958	1957	1956	1955
Editorial	2199	1868	1996	2060
Advertising	2169	2700	2800	2372
TOTAL	4368	4568	4796	3956

TABLE V  
VOLUME OF TRANSACTIONS PAGES

	1958	1957	1956	1955
Groups Publishing	26	24	23	21
No. of Issues	81	75	69	56
No. of Pages	5388	5372	5044	3504

contents was expanded to two pages and transferred from the editorial section to the front of the magazine so that it may be found more readily. At the same time the IRE News and Notes Section, because it is not of permanent reference value, was moved to the front portion of the advertising section.

In addition to improvements in format and content, the number of editorial pages was increased 18% over the previous year, reaching an all-time high of 2199 pages, as shown in Table IV and Fig. 2. At the same time the over-all total for all IRE publications reached a new high of 119 issues totaling 15,140 pages.

The increase in PROCEEDINGS editorial output was due primarily to the appearance of a special issue on Radio Astronomy in January, another on Transistors in June, timed to coincide with the tenth anniversary of the transistor, and a special series of reports in July on the URSI Twelfth General Assembly. In addition, the "Correspondence" section continued to expand, with 180 letters to the Editor being published during the year as compared with 146 the previous year.

The volume of material reviewed for PROCEEDINGS remained about the same as the previous year. A total of 294 papers comprising 2195 pages were considered. Of these, 31% were accepted, 35% were referred to the TRANSACTIONS for publication consideration, and 34% were rejected. The number of papers published, 183, was substantially larger than the 1957 total of 130. Seven IRE Standards also appeared during the year.

*TRANSACTIONS*

The year 1958 saw the TRANSACTIONS output of the Professional Groups continue to increase, as shown in Fig. 2 and Table V. The total number of papers and letters published, 799, represented half of the total number published in all IRE publications during the year (1610). The number of Groups employing letterpress printing increased to 19, and the proportion of pages printed by letterpress increased from 66% to 80%.

*IRE CONVENTION RECORDS*

The practice of publishing CONVENTION RECORDS for the National and WESCON

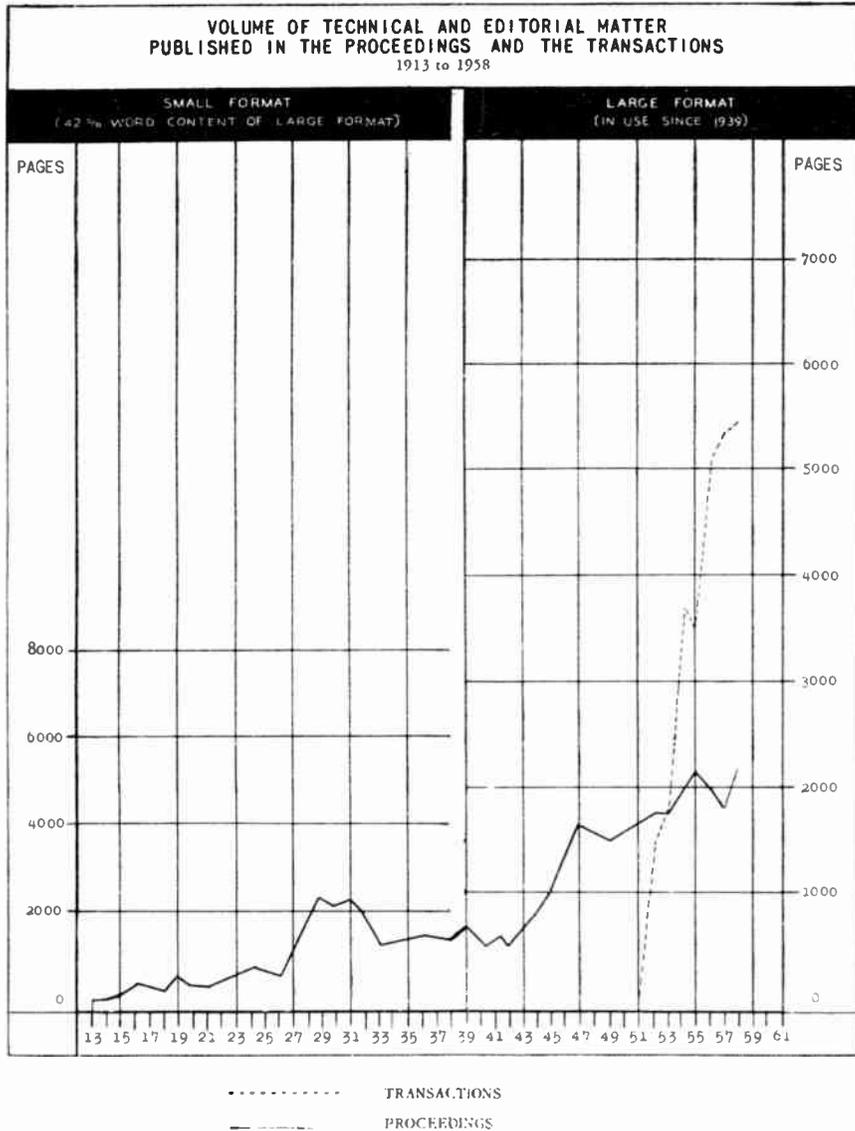


Fig. 2

Conventions was continued. The 1958 IRE NATIONAL CONVENTION RECORD, containing 252 papers and 15 abstracts totaling 2144 pages, was issued in ten parts. The IRE WESCON CONVENTION RECORD contained 180 papers and 14 abstracts totaling 1576 pages, and was issued in nine parts.

#### IRE STUDENT QUARTERLY

Four issues, totaling 208 pages, were sent free to IRE student members during the year. In addition, approximately 20,000 free copies of the September issue were distributed to all non-IRE junior and senior electrical engineering students.

#### IRE DIRECTORY

The 1959 IRE DIRECTORY was published in November, containing 1152 pages including covers, of which 516 were membership listings and information and 636 were advertisements and listings of manufacturers and products.

#### CONFERENCE PUBLICATIONS

The *Proceedings of the 1957 Eastern Joint*

*Computer Conference*, sponsored jointly by the IRE, AIEE, and Association for Computing Machinery, was published by the IRE Editorial Department in June, 1958. The issue contained 264 pages.

### Technical Activities

#### Technical Committees

During 1958 these groups held 262 meetings of which 234 were held at IRE Headquarters and 28 throughout the nation.

Seven Standards, having been approved by the Standards Committee and the IRE Board of Directors, were published in the PROCEEDINGS in 1958, and reprints are now available to the public.

In addition an Index to IRE Standards on Definitions of Terms, 1942-1957 was published in the February PROCEEDINGS.

IRE is directly represented on 33 Committees of the American Standards Association and sponsors three: the ASA Sectional Committee on Radio and Electronic Equip-

ment, C16; the ASA Sectional Committee on Sound Recording, Z57; and the ASA Sectional Committee on Nuclear Instrumentation, N3. One IRE Standard received approval of the American Standards Association as an American Standard in 1958, and is now available overseas through the International Standards Organization.

IRE Technical Committees actively participated in international standardization in 1958 by reviewing and preparing comments on documents for the United States National Committee of the International Electrotechnical Commission.

#### Appointed IRE Delegates on Other Bodies

The IRE appointed delegates to a number of other bodies for the one-year period—May 1, 1958 to April 30, 1959 (as listed on page 44A of the October, 1958 issue of PROCEEDINGS).

The Annual Spring Meeting of the International Scientific Radio Union (URSI) was held on April 24, 25, and 26, 1958, in Washington, D. C. The Fall meeting was held on October 21 and 22, 1958, in University Park, Pennsylvania. The XIIth General Assembly of URSI was held on August 22 through September 5, 1957, in Boulder, Colorado. The XIIIth General Assembly of URSI will be held on September 5 through 15, 1960, at the University College of the University of London, London, England.

During 1958, the Executive Committee of the U. S. Preparatory Committee of the International Radio Consultative Committee (CCIR) held six meetings. At these meetings, the representatives of the fourteen Study Groups summarized and reported on their activities. Numerous responses to the questions under review by the Study Groups have been received in IRE during 1958. Lists of all material received from these organizations were distributed quarterly to the Chairmen of IRE Technical Committees and Professional Groups, as well as to The Joint Technical Advisory Committee.

In October, 1958, the Executive Committee of the CCIR established a working committee, called "Government-Industry Arrangements Committee" to prepare ways and means and to administer the CCIR Ninth Plenary Assembly. This will be held at the Hotel Biltmore, Los Angeles, California, April 1 through 29, 1959. Two meetings have been held to date and the committee structure is being formed.

In June, 1958, the Department of State organized a committee to prepare U. S. Proposals for the International Telecommunication Union Plenipotentiary Conference (PC/PC) to be held in Geneva in October, 1959. At present two Ad Hoc Committees have been established, one will deal with the administration of ITU funds and the other concerns proposals to amend the Radio, Telegraph, and Telephone Regulations between International Conferences. Several meetings have been held to date.

#### The Joint Technical Advisory Committee

The Joint Technical Advisory Committee held a total of eight meetings for the period July 1, 1957 through June 30, 1958. A Tenth Anniversary dinner was held in May, 1958.

Volume XV, the cumulative Annual Report of the JTAC proceedings was published in 1958. This includes in Section I: Official Correspondence between the Federal Communications Commission and The Joint Technical Advisory Committee (IRE-EIA). Also included were other items of correspondence pertinent to the activities of the JTAC. Section II of the Report contained approved Minutes of Meetings of The Joint Technical Advisory Committee for the period July 1, 1957 through June 30, 1958.

The JTAC Subcommittee, Cooperative Interference Committees (57.1) was established on January 24, 1957 to enlist the cooperation of the IRE Sections and Professional Groups in the publicizing of the "Cooperative Interference Committees project" of the JTAC. This Subcommittee completed its work assignments and was discharged on July 25, 1957.

The JTAC Subcommittees on the Study of Forward Scatter Propagation, Ionospheric Scatter and Tropospheric Transmission (55.1) are nearing the completion of their work assignments and final reports are expected shortly.

The JTAC Subcommittee on Study of Interference from Arc Welders (54.2) was disbanded in January, 1957. The Federal Communications Commission recently requested a progress report of the current status of the RF Arc Welding Industry. A review reveals that the Joint Industry Commission (NEMA) is filing a petition with the FCC, recommending that the proposed rule making on high frequency arc welders be made permanent, due to an appreciable reduction in the number of interference cases reported to the Commission. As soon as JTAC receives a copy of the petition a summary of JTAC's activities will be submitted to the Commission.

The JTAC Subcommittee on Study of Single Sideband Transmission (56.1) completed a résumé (Single Sideband Analysis) and was discharged on January 16, 1958.

#### *The International Electrotechnical Commission (IEC)*

The International Electrotechnical Commission Subcommittee 12-1 on Measurements met in Paris, France from March 24 to 31, 1958. IRE secured the services of a delegate to represent the United States, and the IRE Technical Committees prepared this delegate on all items to be discussed at this meeting.

The annual meeting of the International Electrotechnical Commission was held in Copenhagen, Denmark, July 1 to 4, and in Vasteras and Ludvika, Sweden, July 1 to 5, 1958. IRE did not actively participate in these meetings, since there were no meetings of IEC Technical Committee 12 on Radio Communication, or Technical Subcommittee 12-1 on Measurements scheduled at this time.

A list of all documents and material received in the Office of the IRE Technical Secretary from the IEC was distributed to the Chairmen of all Professional Groups, Technical Committees, and Subcommittees.

#### *Professional Group System*

*General:* There are currently 28 Professional

Groups operating actively within the IRE.

Approximately 60% of all IRE members have taken advantage of the Professional Group System which now has a total membership of over 72,000. Included are 5408 Student members of the IRE who have joined the Groups at the special Student member rate of \$1.00 annually. Under the newly instituted Affiliate Plan, 279 scientists and medical doctors, whose major interests lie in fields other than electronics, have affiliated with a number of the Professional Groups.

All of the Groups have levied publications fees and their members are receiving the pertinent Group TRANSACTIONS regularly. In addition, a large number of company, university, and public libraries have subscribed to the TRANSACTIONS of all the Groups. There is also a demand for individual Group subscriptions and individual copies of the TRANSACTIONS from outside sources.

Financial and editorial assistance were among the many services rendered by Headquarters to the Groups during 1958. The Office of the Technical Secretary provided administrative services for Group operations, the planning of meetings, advance publicity, and the recording and mailing for all activities, including 882 mailings to Group members during the year.

*Symposia:* The procurement of papers and actual management of national symposia are entirely in the hands of the Professional Groups. Each of the Groups, with the exception of G-28, sponsored one or more technical meetings during this year, in addition to Technical Sessions at the IRE National Convention, the WESCON, The National Electronics Conference and other joint meetings, for a total of 47 meetings of national import in 1958.

*Publications:* During the year, 25 Groups published 81 TRANSACTIONS containing 5388 pages. Since publication began in 1951, 392 issues (26,092 pages) have appeared. Full details on Group TRANSACTIONS are included in the Report of the Editorial Department.

*Professional Group Chapters:* 242 Professional Group Chapters have been organized by Group members in 54 IRE Sections. Chapter growth is continuing at a healthy rate. The Chapters are meeting regularly and sponsoring meetings in the fields of interest of their associated Groups in the various Sections.

### Section Activities

We were glad to welcome six new Sections into the IRE during the past year. They are as follows: Anchorage, Colombia, Erie (formerly Subsection), Quebec (formerly Subsection), South Carolina (formerly Charleston Subsection), and Western Michigan.

The total number of Sections is now 100.

The Charleston, Erie and Quebec Subsections became full Sections in the year 1958.

The Subsections of Sections now total 28, the following being formed in 1958: Fairfield County (Connecticut), Merrimack Valley (Boston), and Santa Ana (Los Angeles).

A growing major activity of many Sections and the larger Subsections in recent years is the publication of a local monthly Bulletin to fulfill the need for announcing to the Section members the increasing activities of the Section, including 1) Section meetings, 2) Professional Group Chapter meetings, and 3) information on the local and national level of interest to the Section member.

Forty-nine of the Sections and Subsections are now issuing these monthly publications.

### Student Branches

The number of Student Branches formed during 1958 was 17. The total number of Student Branches is now 168, 115 of which operate as joint IRE-AIEE Branches, and 9 as Student Associate Branches.

Following is a list of the Student Branches formed during the year: Academy of Aeronautics, Brigham Young University, Broome Technical Community College, California State Polytechnic College (Kellogg-Voorhis Campus), Capitol Radio Engineering Institute, Catholic University of America, University of Houston (College of Engineering), Houston University (College of Technology), Los Angeles State College of Applied Arts and Sciences, Lowell Technological Institute, Loyola University of Los Angeles, Oregon Technical Institute, Provincial Institute of Technology and Art, Tennessee Agricultural and Industrial State University, University of Western Ontario, and Youngstown University.

It is with deep regret that this office records the death of the following members of the IRE during the year 1958.

#### *Fellows*

Barkely, William J. (M'29, SM'43, F'56, L'56)  
Gustafson, Gilbert E. (A'27, M'38, F'40)  
Hallborg, Henry Emanuel (F'45, L'57)

#### *Senior Members*

Aranson, Ludwig (M'33, SM'43)  
Burnside, Don G. (A'12, SM'46, L'56)  
Campbell, Charles A. (SM'52)  
Crom, George Curtis (A'22, M'24, SM'43, L'57)  
Daly, George M. (A'44, SM'52)  
Donnelly, Clarence P. (SM'57)  
Exon, Frank C. (A'38, SM'47)  
Fitzsimmons, David P. (SM'51)  
Fox, Benjamin (SM'54)  
Giguere, Edmour F. (SM'52)  
Gretener, Edgar (SM'57)  
Hasenberg, Werner (SM'53)  
Jacobsen, Lance R. (SM'57)  
Johnston, Harry C. (SM'53)  
Link, Louis J. (A'30, SM'57)  
Lippincott, Donald K. (SM'43)  
Metcalf, Herbert E. (M'35, SM'43)  
Scott, Milton C., Jr. (A'31, SM'50)  
Talmage, Franklin E. (A'41, SM'49)

Toulon, Pierre Marie Gabriel (SM'47)  
Wick, Edwin L. (SM'46)  
Wiegand, Erich P. (SM'56)

#### Members

Armstrong, Albert B. (A'50, M'55)  
Bosley, Clement T. (S'54, M'57)  
Bruno, Theodore P. (M'54)  
Buffington, Guy B. (M'58)  
Creutz, Patrick M. (M'57)  
Davis, Don L. (M'52)  
Dembinski, Leonard J. (A'49, M'55)  
Dionne, James H. (S'50, A'52, M'56)  
Durkee, Chauncey H. (M'49)  
Dymond, Clifton W. (A'36, M'55)  
Earley, John M. (A'48, M'54)  
Easley, Lowry E. (M'47)  
Elliott, L. W. (A'39, M'55)  
Emmons, Albert W. (S'48, A'49, M'55)  
Fountain, Henry F. (A'48, M'55)  
Francis, M. Clifford (A'44, M'55)  
Gould, Roderick (S'55, M'58)  
Handler, Joseph E. (M'47)  
Jacobs, Herbert, Jr. (S'43, A'47, M'55)  
Johnson, Harold S. (M'50)  
Joyce, Maurice V. (S'51, A'52, M'56)  
Kaplan, Jack (S'44, A'50, M'55)  
Kelly, Joseph M., Jr. (A'43, M'55)  
Kenyon, F. Ralph (A'38, M'55)  
Kohler, Ernest, Jr. (A'31, M'55)  
Koziuk, Frank S. (M'56)  
Lyet, Louis F. (M'57)  
Marsh, Mendole D. (M'53)  
McCoy, John G. (S'49, A'51, M'56)

Morris, Robert (A'46, M'55)  
Muller, John S. (A'51, M'56)  
Museus, Hjalmer U. (A'47, M'55)  
Palmer, Vinson L. (S'39, A'41, M'44)  
Paluthe, Wilfred F. S. (S'56, M'57)  
Pegram, George B. (A'12, VA'39, M'55, L'47)  
Price, Osborne I. (A'43, M'55)  
Read, Cyrus T. (A'43, M'55)  
Robidoux, Finley T. (A'46, M'55)  
Scheck, James A. (M'58)  
Seabury, Edwin M. (M'50)  
Shubert, H. Neal (A'46, M'55)  
Stautner, Paul L. (M'48)  
Sykes, Stephen D. (M'55)  
Trinkle, Wilmer S. (M'49)  
Van Nostrand, Albert W. (S'55, M'57)  
Watts, Wilbur E. (M'46)  
Wellendorf, Joseph J. (A'36, M'45)  
Wiegerinck, Hendrikus T. J. (M'54)  
Williams, Robert E. (A'44, M'48)

#### Voting Associates

Jones, Benjamin F., Jr. (A'39, VA'39)  
Lyon, Henry H. (A'14, VA'39)  
Waite, Samuel A. (A'22, VA'39)

#### Associates

Alonso, Carlos J. (A'51)  
Barsky, Norton (S'53, A'54)  
Becker, Otto F. (A'57)  
Benson, William J. (A'56)  
Brabec, Rudolph J. (A'57)

Bradish, Walter M. (A'51)  
Buck, Harold W. (A'56)  
Chenoweth, William G. (A'51)  
Crouser, Earl E. (A'53)  
Dennison, George L. (A'54)  
Driggs, Lee O. (A'44)  
Fuchs, George W. (S'50, A'53)  
Gnewotta, Milton R. (S'56, A'57)  
Grinberg, Eduardo L. (A'51)  
Johnson, William H. (A'55)  
Lambert, Gerald M. (S'50, A'53)  
Marcinko, Frank C. (A'46)  
McMillan, Dayton, N. (A'53)  
O'Reilly, Thomas F. (A'56)  
Remorenko, Michael S. (A'52)  
Roka, Edward G. (A'54)  
Ross, Nicholas E. (A'57)  
Schwartz, Chris N. (A'55)  
Scrimshire, John B. (S'43, A'45)  
Stoffel, Lester L. (A'35)  
Taylor, Marvin A. (A'56)  
Van Allen, John W. (A'52)  
Walker, Richard D. (A'53)

#### Students

De Vries, Ramon A. (S'57)  
Larue, Jean (S'57)  
Mitchell, Don E. (S'57)  
Mooney, Philip I. (S'56)  
Neumann, Richard G. (S'56)  
Noble, Robert G. (S'57)  
Perrone, Gerard T. (S'57)  
Shay, Lynn (S'57)  
Sullivan, Thomas D. (S'57)

## Books

### Electrical Engineering Materials, by A. J. Dekker

Published (1959) by Prentice-Hall, Inc., Englewood Cliffs, N. J. 197 pages + 4 index pages + ix pages. Illus. 6 x 9. \$8.00.

In this short book, the author undertakes to provide the subject matter for a one-quarter course in those aspects of materials and their behavior of interest to electrical engineering students. His approach is to discuss the necessary general aspects of atomic and molecular structure, and from these to show how the properties and behavior may be explained. This reviewer feels that Dr. Dekker has done a masterful job in the selection and presentation of the material in his 200 page book.

In his Preface, Dr. Dekker notes that his expected audience probably has no background in quantum mechanics, except perhaps in a passing manner. For this reason he has used classical or semiclassical models, often supplemented by arguments based on more advanced principles, to provide a coherent picture. In his words, "I feel that the lack of rigor implied by these models is outweighed by their usefulness in providing the student with a reasonable amount of insight into the physical mechanisms which underlie the properties of materials."

The book contains seven chapters. Chapter 1, "Atoms and Aggregates of Atoms," provides a general introduction to atomic and molecular nomenclature and properties. Chapter 2, "Dielectric Properties of Insulators in Static Fields," contains an account of the various types of polarization, and their relation to the dielectric properties of materials. Chapter 3, "Behavior of Dielectrics in Alternating Fields," examines the frequency response characteristics of dielectrics. Chapter 4, "Magnetic Properties of Materials," is in two parts, the first of which contains background material; the second section addresses itself to the atomic interpretation of the magnetic properties of materials. Chapter 5, "The Conductivity of Metals," relates the conductivity and its temperature dependence through the free electron gas model, and the electron motion in the periodic lattice structure of matter. Chapter 6, "The Mechanism of Conduction in Semiconductors," provides a discussion of the mechanism underlying the action of semiconductor devices. Chapter 7, "Junction Rectifiers and Transistors," contains a condensation of the important properties of semiconductor devices.

Some readers may perhaps protest the occasional lack of rigorous justification of

some developments, or may take issue with Dr. Dekker in occasionally using relations without proof, where the proof may be quite involved. However, if one keeps in mind his stated limited purpose and the student for whom this book has been written, I think that he has succeeded remarkably well.

SAMUEL SEELY  
Case Institute of Technology  
Cleveland, Ohio

### Liquid Scintillation Counting, edited by Carlos G. Bell, Jr. and F. Newton Hayes

Published (1958) by the Pergamon Press, 122 E. 55 St., N. Y. 22, N. Y. 292 pages + xi pages. Illus. 6 x 9. \$10.00.

This volume is the transcript of a symposium which was held at Northwestern University in August, 1957. The liquid scintillation technique has recently found wide application for quantitative measurement of beta emitting radioisotopes. Hence this conference was held to bring together instrument designers from research laboratories, commercial manufacturers of liquid scintillation apparatus, and a broad spectrum of those who make use of the techniques. Many others who belong to these three categories will welcome this written record. The review

articles and many of the specific papers will be of interest to anyone who may be concerned with the design or use of scintillation equipment. This volume, together with the Scintillation Counter Symposia papers published in the IRE TRANSACTIONS ON NUCLEAR SCIENCE (Vol. NS-5, No. 3, NS-3, No. 4) constitute an excellent basic library on the subject.

There are 35 papers which cover the basic processes involved in liquid scintillators, design of present day laboratory and commercial counting systems, the chemistry required to get the elements of interest from various samples into the scintillating liquid, applications of the technique in different fields (archeology, biology, chemistry, industry, medicine, physics), and some examples of unusual systems such as the anti-neutrino experiment of Cowan and Reines which used 3600 gallons of liquid scintillator and 110 five-inch photomultiplier tubes.

There is some duplication of material and variation in the quality among the chapters, but the material is well arranged and well presented.

W. A. HIGINBOTHAM  
Brookhaven National Lab.  
Upton, N. Y.

### The Theory and Design of Magnetic Amplifiers, by E. H. Frost-Smith

Published (1959) by John Wiley and Sons, Inc., 440 Fourth Ave., N. Y. 16, N. Y. 443 pages + 6 index pages + 4 bibliography pages + 34 appendix pages + xix pages. Illus. 5 1/2 x 3 1/2. \$12.50.

This is the first book in a new Automation and Control Engineering Series coming from Great Britain. As the author points out in the preface, the aim of the book is to give an account of the theory of magnetic amplifiers and to link the theory with the design in such a way that it will be of value not only to the professional engineer but also to the university student. Dr. Frost-Smith is most likely to attain this goal.

The book opens with an introduction to basic magnetic circuits. The next two chapters deal with the transductor (this term was adopted in Great Britain after the Second World War to denote a plain saturable reactor). Steady-state and dynamic performance are examined. Here, as well as in subsequent chapters, occasional use is made of a square-wave supply voltage. Such a voltage shape greatly simplifies the graphical solution of transient problems, and is a very useful tool for didactic purposes.

Chapter 4 is entitled the "Self-Excited Transductor" (designated in the U. S. by saturable reactor with external feedback). The core material is assumed to have a flux-current loop of zero loop width, consisting of the familiar three straight lines, with saturation branches being horizontal. The loop width of the flux-current loop is, however, an important parameter affecting the saturable reactor with external feedback. While the assumption of zero loop width facilitates the analytical treatment, the assumption also introduces discrepancies with the observations made on actual amplifiers, since core materials with zero loop width are nonexistent. The next chapter is titled "Transductors with Auto-Excitation"

(called self-saturated amplifier in the U. S.). The assumption regarding the flux-current loop is the same as in Chapter 4, and the treatment follows previously established methods. The effects of ac and dc inductive loads, and of dc capacitive loads, are treated in Chapter 6.

Chapters 7 and 8 deal with the dynamics of magnetic amplifiers. Time lags in the amplifier, as well as in inductive and capacitive loads, are discussed. The treatment is in the form of operational calculus, augmented by presentations in the gain-phase angle-frequency domain. Oscillograms and other measured data support the text convincingly. The following chapter gives a good account of amplifiers with half-cycle response. Chapter 10, "Balanced Magnetic Amplifier Circuits," deals effectively with various forms of push-pull circuitry. Chapter 11 is called "Low Level Amplification and Multi-Stage Amplifiers." A good account is rendered of various forms of drift, and the even-harmonic detector, both single-ended and push-pull. Very little, however, is said about multi-stage amplifiers although the chapter title carries this topic in its heading. A brief discussion of magnetic modulators is presented in Chapter 12.

"Construction and Design of Magnetic Amplifiers" is the title of Chapter 13, a chapter which is likely to become the most popular of the book. It contains a vast amount of practical information and also several design procedures which were so sorely missed in the literature. Concluding the book is a good selection of numerous applications of magnetic amplifiers.

In summary, this very interesting book covers a vast area of the magnetic amplifier art with the emphasis divided about evenly between theory and practice. The book is very well written, and many illustrations support the text. The terminology is that of Great Britain, however, which may lead to the occasional confusion of the U. S. reader. The very high degree of competence of the author makes itself felt on nearly every page.

This book will make many friends for Dr. Frost-Smith among practicing engineers and students alike. Congratulations are offered to the author for accomplishing such fine work.

H. F. STORM  
General Electric Co.  
Schenectady, N. Y.

### Progress in Metal Physics, Vol. 7, 1st Ed., edited by Bruce Chalmers

Published (1958) by Pergamon Press, Inc., 122 E. 55 St., N. Y. 22, N. Y. 395 pages + 12 index pages + viii pages. Illus. 6 X 9 1/2. \$16.00.

The "Progress in Metal Physics" series presents authoritative reviews of specialized subjects in the fields of physical metallurgy and metal physics. These reviews strive to summarize and appraise the existing state of knowledge in diverse areas. It is the hope of the editors that this series of timely reviews will help research workers keep abreast of developments in their own and closely related fields.

The present volume has the same high quality and pleasant balance as its predecessors. Included in this volume are the follow-

ing five articles: 1) "Equilibrium, Diffusion, and Imperfections in Semiconductors," by J. N. Hobstetter (64 pages), is written on an introductory level, and contains a good deal of background material. It is designed to familiarize specialists in metal physics with the theoretical and experimental advantages of studying such phenomena as equilibrium, diffusion of atoms, and the generation and motion of dislocations and other imperfections in semiconductors. 2) "The Physical Metallurgy of Titanium Alloys," by R. I. Jaffee (100 pages), is a highly detailed account of titanium alloys, with emphasis on their crystallographic, metallurgical, thermal, and mechanical properties. A wealth of experimental data is presented, as well as a critical analysis of this data. The general impression is that the alloying of titanium is still largely an empirical science. 3) "Thermodynamics and Kinetics of Martensitic Transformations," by Larry Kaufman and Morris Cohen (81 pages), is a comprehensive and highly readable survey of the physical chemistry of martensitic (displacive or shear-like) reactions. There has been much noteworthy progress in the understanding of such reactions during the past decade, and this is capably summarized here. 4) "The Stored Energy of Cold Work," by A. L. Titchener and M. B. Bever (91 pages), is a well-organized and interestingly written account of the process of deformation by cold working, the nature of the cold-worked state, and the mechanisms of the restoration processes that occur in cold-worked metals. In addition to a well-documented survey of pertinent experimental investigations, there is a thoughtful section devoted to overall conclusions, a feature lacking in a few of the other articles in this volume. 5) "The Properties of Metals at Low Temperatures," by H. M. Rosenberg (55 pages), is a compact account of the electrical and thermal conductivity, the specific heat, and diverse mechanical properties of metals and alloys at low temperatures, including the superconducting phase. A clear and well-balanced picture of the subject emerges from this paper.

This volume forms a welcome addition to the review literature. Each of the articles is well documented, and the value of the volume as a whole is enhanced by author and subject indices. It is expected that each of the articles will prove useful to a wide circle of readers.

FRANK HERMAN  
RCA Laboratories  
Princeton, N. J.

### A Facsimile Edition of George Green's Essay on Electricity and Magnetism, by Stig Ekelöf

Published (1958) by Wezata-Melins Aktiebolag, Göteborg, Sweden. 72 pages + vii pages. 8 X 10 1/2. Half Cloth Binding. The book will be sent postage free to any country on remittance of U. S. \$5.00 or the equivalent in Swedish Kronor to Institute for Theoretical Electricity, Chalmers Institute of Technology, Göteborg, Sweden.

This most remarkable book is an original contribution to mathematics and to the theories of electricity and magnetism made 131 years ago by George Green, a 34 year old self-taught miller's son from Notting-

ham. The exposition is a model of clarity and directness. Because of this and the fact that all modern books on advanced calculus and electrostatics still use the terms introduced by George Green, the essay is very readable. It should certainly be most stimulating and inspirational to students struggling with Green's theorem, Green's functions, etc., in courses on advanced calculus and partial differential equations. Every college library should have a copy.

On Page 1 Green creates a name "poten-

tial function" to facilitate thought and communication of ideas. The symbol for it,  $V$ , follows. On Page 3 he obtains the properties of potential functions in the interior of a charged surface and at infinity and establishes the continuity of potential functions across the surface. On Page 10 he enunciates and proves a theorem which we now know as Green's theorem. Generalizations of this theorem follow. Then come the expressions for the potential functions in the interior (or the exterior) of a closed surface in terms of

their values or the values of their normal derivatives on the surface. It is here that Green introduces functions now known as Green's functions. The remainder of the essay is divided about equally between the applications of the general results to electrostatics and magnetostatics.

Congratulations to Professor Ekelöf on recalling from oblivion this mathematical masterpiece.

S. A. SCHELKUNOFF  
Bell Telephone Labs.  
Murray Hill, N. J.

## Abstracts of IRE Transactions

The following issues of TRANSACTIONS have recently been published, and are now available from the Institute of Radio Engineers, Inc., 1 East 79th Street, New York 21, N. Y. at the following prices. The contents of each issue and, where available, abstracts of technical papers are given below.

Sponsoring Group	Publication	Group Members	IRE Members	Non-Members*
Component Parts	CP-6, No. 1	\$2.65	\$3.95	\$7.95
Education	E-2, No. 2	1.10	1.65	3.30
Electron Devices	ED-6, No. 2	1.70	2.55	5.10
Engineering Management	EM-6, No. 1	0.70	1.05	2.10
Industrial Electronics	IE-9	1.00	1.50	3.00
Information Theory	IT-5, No. 1	1.00	1.50	3.00
Instrumentation	I-8, No. 1	1.95	2.90	5.85
Vehicular Communications	VC-12	1.60	2.40	4.80

\* Libraries and colleges may purchase copies at IRE Member rates.

### Component Parts

VOL. CP-6, NO. 1, MARCH, 1959

Information for Authors (p. 1)

Who's Who in PGCP—Louis Kahn, Member Administrative Committee July, 1958 to July, 1961 (p. 2)

An Introduction to Inorganic Dielectrics—K. H. McPhee (p. 3)

The characteristic behavior of inorganic dielectrics under thermal, electrical, and mechanical stress is explained in terms of the atomic and molecular structure of matter. Nonconductors and conductors are treated as different areas of a continuous materials spectrum.

The nature and properties of constituent raw materials in representative ceramic bodies are discussed. The chemistry of their combination is analyzed and Gibbs phase rule is applied.

The phenomenon of plasticity in clay-water systems and its application to forming and shaping are examined. The manufacture of bodies containing varying amounts of plastic ingredients is described, and the influence of forming techniques on resultant properties of these bodies is shown.

A few of the more common dielectric ma-

terials are listed, and a number of good design practices are given.

Components for Submarine Telephone Cable Repeaters—M. C. Wooley (p. 34)

The first transatlantic, Alaskan, and Hawaiian Submarine Telephone Cables are provided with repeaters spaced about 37.5 miles apart. Each repeater consists of a 3-stage electron tube amplifier and contains approximately 60 components. Because of their inaccessibility for repair or replacement, it is desired that these systems operate for the order of 20 years without failure of any component. This requires a degree of reliability for both active and passive components and orders of magnitude higher than are achieved in more conventional systems.

Electron tubes and high-voltage capacitors subject to "wear out" have been the subject of intensive development and life testing for more than 15 years. These studies indicate that at the low cathode temperature and low cathode current density used in the electron tubes, satisfactory thermionic performance can be expected for more than 20 years. Likewise, tests on the high-voltage capacitors indicate that, in spite of the use of a high dielectric stress, the probability of a wear-out failure in 20 years is low.

These estimates, however, do not include the probability of catastrophic or sudden failure of the components. Since practicable test programs cannot measure the extremely low catastrophic failure rate expected of these components, it must be attained through the use of 1) reliable types, 2) close control of raw materials, 3) simple designs which are easy to make and inspect, 4) careful manufacture, and 5) thorough inspection.

This procedure is restrictive and involves close attention to all details in both design and manufacture. Consequently, unusually detailed specifications were required for raw materials, processes, and components. Complete records of all operations and inspections were kept to insure against omissions or errors. Specially selected operators were trained and required to demonstrate ability to produce satisfactorily before placed on production work. Manufacture was carried out under exceptionally clean conditions. The product was inspected at each stage of manufacture so that hidden defects would not be missed. The resulting slow and painstaking manufacture disclosed many defects which would have been overlooked in normal production.

Although no numerical value can be attached to the degree of reliability attained, experience in production and service to date indicate that it is unusually high. Over 270 million component hours of service without a failure have been accumulated to date, which indicates, with 90 per cent certainty, a failure rate of less than 1 in 10,000 per year.

A New Memory Device—The Twister—D. A. Ellerbruch (p. 42)

Although the "Twister" memory device is still in the experimental stage at the Bell Laboratories, great expectations for its use in the future may be considered. Very favorable performance results have been obtained in the past, and if the results are as favorable in the future, it may eventually replace many of the magnetic core and drum types of memory systems. Physical size of the Twister memory device may be small due to the diameter of the magnetic wires used and the high-storage density of an array.

Contributors (p. 45)

### Education

VOL. E-2, NO. 2, APRIL, 1959

Special Issue—Sagamore Conference on Electrical Engineering Education, May 30-June 1, 1958

**Editorial**—Wilbur R. Lepage (p. 32)

**Advanced Education—Past, Present, and Future**—F. Hamburger, Jr. (p. 33)

A brief historical review of education beyond the baccalaureate with particular emphasis on graduate study in electrical engineering is followed by an examination of present trends. Consideration is given to such topics as admission requirements, programs of study, graduation requirements, and residence requirements as they affect students. Staff factors such as qualifications, load, compensation, research, and advancement are considered. The problems that must be examined to assure a healthy future are listed to stimulate continuing discussion.

**Industry's Contributions and Needs in Graduate Education**—David F. Kline (p. 40)

The General Electric Company provides, within its Advanced Technical Programs for Engineers and Scientists, a logical sequence of courses for the development of these men. Within this framework of courses is the Honors Program for Graduate Study, which enables the motivated engineer or scientist to complete requirements for the Master's degree. The opportunity is also present to continue further his educational development through the Advanced Engineering Program, the Creative Engineering Program, or a Tuition Refund Program.

**The Challenges in the Development of Graduate Programs**—Ernst Weber (p. 41)

In order to protect the integrity of graduate degrees, the basic objectives of graduate study must be examined and interpreted in the widest sense, taking into consideration changing social conditions. This paper traces the development of graduate engineering education from World War I, with its original attachment to the established pattern of graduate study in university science departments. It considers the place of research in graduate study and the problems arising from the need for research sponsorship today. Also, attention is focused upon the classification problems arising from the large number of part-time graduate students enrolled in engineering schools today.

**Current Developments in Patterns of Graduate Study**—Newman A. Hall (p. 45)

During recent years considerable attention has been given to an analysis of the changing patterns of graduate study in engineering. Reviews by the American Society for Engineering Education, the University of Chicago under a Carnegie Grant, and the Ford Foundation have been directed to the problem. A primary observation has indicated a trend towards greater diversity in patterns and a closer correlation with current requirements of research and development in industry and government. These patterns, including evening and various forms of off-campus training, are described and their effect on the nature of the graduate program is examined. The effect of direct and indirect industry support is analyzed, and the changing role of the faculty in cases where close coordination with industry exists is examined. Current trends are evidently enriching and strengthening the value of the graduate program, particularly through immediate contact by both student and faculty with research and development applications. There exists, however, a consequent overemphasis on immediate needs, and an ambiguity of long-range academic objectives. The critical financial needs of students, faculty, and engineering educational institutions are emphasized by the demands of new patterns and current high standards of academic endeavor.

**Off-Campus Resident Graduate Programs at Company Centers**

**The Syracuse-IBM Program**—R. A. Galbraith (p. 52) and J. R. Lakin (p. 53)

**The R.P.I.-United Aircraft Program**—

Warren C. Stoker (p. 55) and G. C. Barnes (p. 58)

**The NYU-Bell Labs Program**—S. S. Shamis (p. 60) and S. B. Ingram (p. 62)

**Question and Answer Period Following Panel Discussion** (p. 64)

**Contributors** (p. 66)

## Electron Devices

VOL. ED-6, NO. 2, APRIL, 1959

**Frequency Characteristics of a Semiconductor Rectifier at Voltages Greater than  $kT/q$** —L. Depian, W. E. Newell, and A. G. Milnes (p. 125)

In this paper, the frequency characteristics of a half-wave  $p-n$  junction rectifier with a resistance load are analyzed for applied voltages greater than  $kT/q$ . Beginning with the derivation of an implicit integral equation which describes the large-signal operation of this circuit, the effect of frequency on the rectification efficiency is then obtained from an iterative numerical solution of this equation using an IBM 650 computer. Contrary to the small-signal theory developed by Kalashnikov and Penin, it is found that at higher applied voltages the half-power frequency is very dependent on voltage as well as load resistance, carrier lifetime and reverse saturation current. The relationships between the half-power frequency and these parameters are shown graphically, and the results are compared to the small-signal theory for small applied voltages.

**A Lumped Model Analysis of Noise in Semiconductor Devices**—R. N. Beatie (p. 133)

A clear picture of the noise properties of semiconductor devices may be gained through a study of lumped models. When a semiconductor device is represented by a lumped model it is found that all the noise may be accounted for by associating a noise-current generator with each of the conductances appearing in the lumped model. The expressions describing these noise-current generators are easily found, and the determination of the device noise properties involves only lumped model analysis.

Such analysis applied to a diode model and a transistor model yields results which are in agreement with the results of other authors, and with experimental measurements cited by them. The mechanisms considered in this paper do not, however, account for the so-called  $1/f$  noise.

**A Physical Theory of Junction Transistors in the Collector-Voltage-Saturation Region**—C. Huang (p. 141)

A physical theory has been formulated for the operation of junction transistors in the "collector-voltage-saturation" region or "on" region. Transistor characteristics in this region are important for switching applications. Class A or Class B amplifiers, as well as other large signal applications. The formulation is based on the physical consideration that in the "collector-voltage-saturation" region the collector-base junction is forwardly biased, and that the injection level is high.

Two-dimensional distributions of carrier densities, current densities, and electric field are obtained for separate portions of the base region. Using these distributions, theoretical expressions are derived for the characteristics of  $p-n-p$  and  $n-p-n$  transistors including saturation voltage, base input voltage, and dc current amplification factor. Good agreement between theoretical and experimental results indicates that the approximations used in the theory are valid. Numerical calculations have been carried out for the saturation voltage, base input voltage, and dc current amplification factor for different geometries and material properties. The calculation shows the use of the theory for quantitative designs of transistor characteristics.

**Surface and Geometry Effects on Large Signal Base Input Voltage and Input Resistance of Junction Transistors**—C. Huang, C. M. Chang, and M. Weissenstern (p. 154)

A design theory for large signal base input voltage and input resistance of junction transistors is presented with emphasis on the effect of surface recombination and transistor geometry.

Two-dimensional distributions of minority carrier density and electric field are obtained for the separate portions of the base region. These distributions are then expressed in terms of emitter current and are used to derive expressions for the large signal base input voltage and resistance. The dependence of carrier distribution on surface recombination velocity and transistor geometry is illustrated by curves.

The theory is corroborated by a series of experiments carried out with  $p-n-p$  power transistors. The parameters varied are surface recombination velocity (by baking), emitter diameter, base ring diameter, dice thickness and base region resistivity. The measured base input voltage and resistance are plotted and compared with calculated values based on the theory presented. Good agreement is found between calculated and measured values. The results deviate in several interesting respects from the values predicted by the small signal or low injection level theory. Based on this theory, design considerations for the large signal input voltage and input resistance are discussed.

**On the Theory of DC Amplification Factor of Junction Transistors**—S. Wang and T. T. Wu (p. 162)

According to Webster's phenomenological theory, the base current consists of three terms arising from 1) surface recombination, 2) bulk recombination, and 3) non-unity emitter efficiency. In this paper, the diffusion equation inside the base region of an alloyed transistor is solved, and an analytic expression for the dc amplification factor is presented in terms of arbitrary bulk lifetime and surface-recombination velocity and the geometry of the transistor. The surface-recombination term depends on the distribution of the injected carriers along the emitter-base boundary, and hence, it increases as the emitter current increases on account of the base-resistance bias effect which tends to concentrate the minority carriers near the periphery of the emitter. The falloff in beta at high-emitter current is attributed mainly to an increase in the surface-recombination term, contrary to Webster's theory. A comparison with Webster's theory is made on units having the same bulk property but different surface treatments. Experimental results give strong support to the present theory.

**An Alloy-Diffuse Silicon High Current Transistor With Fast-Switching Possibilities**—D. Navon and P. Debeurs (p. 169)

A method is described for the fabrication of a high-power silicon transistor capable of operating at case temperatures exceeding  $100^\circ\text{C}$ . Large-area transistors can be produced having uniform base widths of a few ten-thousandths of an inch. The base is established by the diffusion of impurities from a double-doped alloyed contact on the surface of the silicon. This device gives a common emitter current gain of greater than 30 at 10 amperes collector current and can dissipate 100 watts if the case temperature is held below  $55^\circ\text{C}$ . In addition it has constant common emitter current gain out to 200 kc. Hence this technique produces a power device combining both good current gain and frequency characteristics.

**High-Level Transistor Operation and Transport Capacitance**—K. E. Mortenson (p. 174)

A two-dimensional analysis of high-level transistor operation is presented which includes the effects of an extended base region,

internal emitter biasing,  $\gamma$ -falloff, unequal collector and emitter dimensions, and surface recombination.

The transistor model considered is directly appropriate to the strip-type geometry, but also yields results which are approximately valid for the ring- and dot-type structures under certain conditions. Transforming the geometry permits a solution to be obtained for the charge-density distribution in the base as well as the current density distribution at the emitter and collector junctions. From these relations, both the collector and emitter transport (diffusion) capacitances are also determined.

Two complete numerical evaluations of the theoretical results are given, first for a symmetrical unit with equal emitter and collector dimensions, and second for an unsymmetrical unit with the collector dimension 24 per cent greater than that of the emitter. It is indicated that an appreciable fraction of the total base charge can exist external to the emitter and collector, particularly for very high-level operation, causing large increases (1.5 to 3 times the one-dimensional values) in both the emitter and collector transport capacitances, particularly for units having grossly extended base regions and low-surface recombination velocities. Further shown is the effect of increasing the collector dimension over that of the emitter; the capacitances are appreciably lowered and the transport efficiency (and thus the current gains) is increased.

Finally, some collector transport capacitance measurements are presented covering the entire operating range which tend to substantiate the theoretical results.

#### Impedance and Dispersion Characteristics of the Flattened Helix—C. C. Johnson (p. 189)

Curves are presented which give the dispersion and impedance characteristics of a flattened helix in the presence of a dielectric material and a metal shield. The impedance curves are very similar to the corresponding curves for a round helix whose circumference is equal to the perimeter of the flattened helix. The flattened helix has dispersion characteristics similar to a round helix whose diameter equals the flattened helix thickness.

#### An Investigation and Application of the Contraround Helix—J. E. Nevins, Jr. (p. 195)

An investigation of the contraround helix and the determination of its characteristics has been completed. By means of a hollow-beam traveling-wave tube and perturbation methods, the fundamental and minus-one space harmonics impedances have been measured and a number of space harmonics of the symmetric and antisymmetric modes identified.

The fundamental impedance was measured to be equal or greater than the sheath-helix impedance depending, however, upon the thickness and width of the helix tape. The minus-one space harmonic impedance was measured to be larger than the fundamental impedance for  $ka > 0.65$ .

The contraround helix has been employed as a high-power X-band amplifier (8.5–9.5 kmc). The characteristics of the tube described are 25-db gain, power out  $> 1$  kw, efficiency 7–10 per cent.

#### Bibliography on Electroluminescence and Related Topics—H. F. Ivey (p. 203)

#### Propagation of Space-Charge Waves in Diodes and Drift Spaces—H. G. Kosmahl (p. 225)

It has generally been found that propagation of space-charge waves may be described 1) in space-charge-limited gaps by Gauss' hypergeometric functions; 2) in gaps with no space charge but with an accelerating homogeneous field by Bessel functions with real argument; and 3) in spaces free of dc fields by simple trigonometric functions. The equations derived become simpler through the selection of the dc velocity  $\mu$ , rather than the time  $t$ , as the independent variable.

#### Disturbances in a Multi-Velocity Plasma—J. R. Pierce and J. A. Morrison (p. 231)

The single-velocity treatment of ac disturbances in electron flow is often accurate enough even when the flow actually has an infinitely broad but peaked velocity distribution, such as a Maxwellian or a bell-shaped distribution. In seeking a better approximation, one is tempted to expand in terms of the moments of the velocity distribution and to disregard terms beyond that involving the second moment. This leads to a dispersion equation describing waves which have no physical existence. A linearized analysis by transform methods predicts a field which oscillates in an exponentially damped manner. This is not the oscillation of a normal mode, however, for various velocity classes of the charge distribution have ac densities which grow with time. This indicates that linear expression will hold over a finite interval only.

A simple example which compares a multi-stream analysis with an analysis based on the charge density in phase space (the distribution function) indicates the same sort of failure of both of the linearized theories at large times. Thus, it appears that the failure is characteristic of the problem rather than of the method of solution. The method of solution is a matter of choice.

#### A Self-Excited Drift-Tube Klystron Frequency Multiplier for Use in Generating Millimeter Waves—W. H. Cortet, Jr. (p. 236)

Theoretical and experimental considerations are presented in this paper for a self-excited frequency multiplier in which the fundamental oscillator is an integral cavity two-gap klystron and a harmonic resonator is placed at an appropriate position in the drift tube between the two gaps. This is a new device and it is called a drift-tube klystron frequency multiplier abbreviated as DTKFM.

A general design procedure is given, based on both theoretical equations and experimental characteristics. Operation of the self-excited DTKFM with fundamental oscillations at 4 and 1 cm and harmonics at 1 cm and 5 mm respectively, is presented. The device is capable of generating larger CW output power at millimeter wavelengths than many existing devices.

#### Experimental Notes and Techniques

##### A High-Frequency Alpha Cutoff Jig—L. M. Terman (p. 242)

##### Magnetron Cathode Arm Radiation and Luminescence—R. A. White and A. Bamford (p. 242)

##### Program of 1958 Electron Devices Meeting, October 30–31, Washington, D. C.—(p. 243)

##### Contributors (p. 254)

## Engineering Management

### VOL. EM-6, NO. 1, MARCH, 1959

#### An Oscilloscopic View of the Electrical Industry by Investment Capital—C. M. Bower (p. 1)

#### Research and Development Cost Estimation—V. L. Lambert and H. F. Sackett (p. 8)

Four companies engaged in research and development were interviewed relative to their method of cost prediction. Company A, engaged in military equipment development, uses a complex operating organization to make estimates. Company B, engaged in commercial development, uses a summary unit cost, *i.e.*, number or predicted man years multiplied by a given number of dollars. Company C, commercial research and development, estimates on the basis of individual projects but holds department managers accountable only for total department costs. Company D, government research and development, uses the future operating structure as does Company A, but

seems to use a less complex structure. A further investigation into the use of summary unit costs is recommended.

#### Industry-University Cooperation in the Field of Research—R. I. Cole (p. 12)

#### Recruiting the Technically Creative—A Dual Responsibility—J. Kurshan (p. 14)

Effective recruiting requires teamwork between the personnel staff and the research and development organization; neither should shirk its responsibilities. Effective organization for the task of recruiting creative engineers and scientists will also improve the recruiting of engineers in general.

Creative individuals are especially sensitive to the nonmaterialistic aspects of their working environment. These include the nature of the work, reputation of the organization, opportunities for growth, quality of supervision, and living conditions.

With reference to the applicant, recruiting may be divided into six basic elements: attracting, screening, visiting, interviewing, decision-making, and follow-up. All involve participation by both line and staff parts of the organization, but the responsibilities of each group are distinct.

The personnel staff has the basic responsibilities of providing professional employment guidance and of handling administrative details. These duties include personnel research, employment planning, organizing the recruiting effort, training recruiters, screening applicants for further consideration, arranging laboratory or plant visits, interviewing from a nontechnical viewpoint, and coordinating the action on the decision, the offer, and the subsequent follow-up.

Technical management must set its own manpower requirements, both quantitatively and qualitatively, and make the final decision whether to hire a given individual. Nonsupervisory technical personnel can contribute effectively in attracting, recruiting, and interviewing applicants. Criteria are discussed for selecting people to do the recruiting tasks.

#### A Forward Look at Management Development—R. K. Greenleaf (p. 19)

#### Merit Rating and Productivity in an Industrial Research Laboratory: A Case Study—A. G. Grasberg (p. 31)

A merit rating plan based only on one criterion, over-all performance, was tested in one section of approximately sixty engineers/scientists in a large industrial research laboratory. The procedure used is discussed along with some general results. In addition, a measure of individual productivity was developed based on the average number of technical reports, publications, and patent disclosures written per year. A fairly good correlation was found between the merit ratings and the measure of productivity.

#### "Proselyting"—Good Business?—R. G. Swander (p. 37)

#### For Your Bookshelf (p. 40)

## Industrial Electronics

### VOL. IE-9, APRIL, 1959

#### Design of a High-Performance Instrument Servo for General Purpose Computation—K. V. Bailey and M. A. Ziniuk (p. 1)

A computing servo capable of accurate rate and position operation with extended frequency response relative to conventional computing servos is described. The usual tachometer generator with its penalizing inertia is avoided in this design. Rate feedback is provided instead by a precision capacitor which couples the follow-up potentiometer to the input summing junction. Design objectives, compromises, and performance data are reviewed.

### Electronics in Automatic Mail Cancelling and Facing—P. W. Barnhart (p. 14)

An automatic mail cancelling and facing machine has been developed which processes 500 letters per minute in a single pass. Letters are introduced into the complex, five-module machine system in random postage stamp orientations, high or low, leading or trailing. Stamp locations are individually detected by flying spot scanners. These scanners generate command signals for the selective actuation of the appropriate one of twelve cancelling heads and for the subsequent direction of the letter to one of four terminal letter collection stackers in which all letters are uniformly faced, that is, with stamps, and therefore addresses, oriented in the same manner.

### An Automatically Controlled Vertical Turret Lathe—A. O. Fitzner (p. 19)

This paper describes a punched-tape control system for a vertical turret lathe having a rail head and a side head. The system exercises complete control over turret indexing, tool feeds, table speed, tool dwell, etc. Coordination of the operations of the two heads is obtained by proper wording of the program. Calibrated adjusting potentiometers (automatically selected by the control) have been provided to allow the operator to compensate for tool deflections, and the like, on critical workpiece dimensions. Novel features such as single-cycle operation, single-head operation, and operation repeat have been incorporated in the control to facilitate the setup procedure.

Programming is simple and direct, with no equipment required other than a standard 8 channel Flexowriter.

### Minimizing Production Costs Through Modular Automatic Test Equipment—J. Tampico and H. B. Rose (p. 29)

The dominant consideration influencing the trend toward automatic test and checkout techniques for production-line test stations is cost reduction. Testing costs includes the costs of test equipment and spares, of operating and maintaining personnel, of floor space, and of down time. As test cost per unit produced is the basis for comparison, the test rate achieved per station is a major factor. An analysis of test station requirements for a mass-produced missile is used as an example. This analysis shows major reductions in test cost per unit as a result of automation for subassemblies which have fairly complex test requirements; whereas for simple subassemblies, there are no cost improvements and manual testing is quite satisfactory.

The necessity for maintaining high test rates with fewer, less skilled operators and with the corresponding reduction in test stations imposes high reliability requirements on automated equipment in order to preserve this economic advantage. Therefore, such features as self-checking, fail-safe design, and printed-out test results are essential.

The construction of automatic test sets from functional universal modules reduces first cost of equipment, permits rapid and economic servicing, and reduces obsolescence of equipment. Another advantage of the modular approach to production-line test set design is that test sets using only a low order of automation may utilize modules identical to those of the fully automated equipment. This extends the "break-even" point in the choice between automated and manual equipment towards increased application of the more sophisticated techniques.

**Data or Information—E. A. Keller (p. 35)**

## Information Theory

VOL. IT-5, No. 1, MARCH, 1959

Frontispiece—Joseph L. Dobb (p. 2)

Editorial—Joseph L. Dobb (p. 3)

### Correlation and Delay Line Attenuation—M. J. Jacobson (p. 4)

The effect of delay line attenuation on the output of a correlator is studied for the case where the attenuation in db to within a frequency independent loss varies linearly with delay and as the square root of frequency. It is shown how attenuation affects the output signal-noise ratio of the correlator, increasing it for some signal spectrum shapes and decreasing it for others. Upper bounds on the increase and decrease are computed and a sufficient condition on the input signal spectrum is established which, when satisfied, assures a degradation in signal-noise ratio. Also examined is the effect of a frequency independent gain inserted to compensate for the delay line attenuation. It is shown how this gain may be chosen in order to give uniform system output noise over all values of delay.

### On the First Probability of Detection by a Radar Receiver System—W. M. Stone, R. L. Brock, and K. J. Hammerle (p. 9)

Expressions for the detection probabilities associated with the output of filter-square law detector-filter radar receivers are presented for practical filter systems and with a slowly varying Rayleigh distributed signal amplitude.

### Full Decodable Code-Word Sets—M. P. Schutzenberger and R. S. Markus (p. 12)

This paper considers further how the decodability condition imposes restrictions on a set of code words. A generating function is defined that describes the composition of the code words. The relation between the generating function and a "full" set of code words is found. This relation shows that the sum of arbitrary probabilities associated with the words of a full set must be one. A full set of code words is one to which no code word can be added and still keep the set decodable. It is also shown that a full set is "completable." For a completable set of code words any string of symbols can be made into a sentence by adding a suitable prefix and a suffix.

### On a Property of Wiener Filters—Moshe Zakai (p. 15)

Let  $Y(\omega, \alpha)$  be the Wiener filter designed to yield an output which is the least-square approximation to  $s(t+\alpha)$  where  $s(t)$  is the desired signal input. Let  $Y_L(\omega)$  be the Wiener filter designed to yield an output which is the least-square approximation to some linear operation  $L$  on the desired signal input. The following simple relationship has been shown to hold between  $Y(\omega, \alpha)$  and  $Y_L(\omega)$ . If  $s(t)$  is the desired signal input and  $L_\alpha[s(t+\alpha)]$  is the desired output, where  $L$  is some linear operation with respect to  $\alpha$ , then  $Y_L(\omega) = L_\alpha[Y(\omega, \alpha^\alpha)]$ .

### Machine Recognition of Hand-Sent Morse Code—B. Gold (p. 17)

A transistorized special purpose digital computer called MAUDE (Morse AUtomatic DEcoder) has been designed, built and analyzed. This computer decodes a hand-sent Morse message, which is printed on a teletypewriter.

MAUDE'S decisions take place at a number of different levels, and its "knowledge" is not only that of relative durations of dots and dashes, but also of the Morse code and even of certain elementary properties of language.

MAUDE has successfully decoded between 90 per cent and 95 per cent of 184 operators. A successful decoding is one which results in a text which can be easily read by a man who knows the language.

It is felt that MAUDE can be a practical piece of equipment for a site with heavy traffic. Its performance will be inferior to that of a man until more sophisticated language rules, using at least a word vocabulary, are included.

### The Morse Distribution—M. Freimer, B. Gold, and A. L. Titter (p. 25)

A problem which arose during research in-

involved in designing a machine to translate hand-keyed Morse code into printed text may be stated as follows: Let  $X = \{x_i: i=1, 2, \dots, n\}$  be a sequence of independent random variables all of which have the same distribution. Assume that the probability that  $x_i = x_j, i \neq j$ , is zero. Let  $k$  be a positive integer  $\leq n$ , and consider all subsequences  $x_i, x_{i+1}, \dots, x_{i+k-1}$  of  $X$  consisting of  $k$  consecutive variables. Let us distinguish, with a check ( $\checkmark$ ), the largest member of each such subsequence. We have studied, and partially tabulated,  $p_n^k(r)$ , the probability that exactly  $r$  members of the sequence  $X$  are not checked. This paper contains most of the pertinent results.

Correspondence (p. 32)

Contributors (p. 35)

Roster of PGIT Members (p. 37)

## Instrumentation

VOL. I-8, No. 1, MARCH, 1959

Abstracts (p. 2)

### A High-Speed Low-Level Scanner—K. Ennslein (p. 3)

The scanner to be described in this paper has as its general specifications the following requirements: scan 500 pairs of strain-gauge inputs at a repetition rate of 50 pairs per second under control of a train of control pulses. A minimum dwell time of 5 msec is needed. The output of this scanner is to be connected to a high-speed digital voltmeter. This digitizer in turn feeds an IBM 727 tape unit. The data from the tapes are then to be processed by an IBM 704. Life, therefore, should be compatible with that of the associated equipment.

This scanner should then combine two features usually only available independently: its life should be such that at least 20,000,000 operations can be expected without readjustment and the scanning rate must be at least 50 points per second. In addition, the contact noise level must be extremely low due to the small amplitude of the strain-gauge signals. (There is, in this respect, one advantage: the strain-gauges are of rather low impedance.) The low-noise requirement is usually satisfied by means of very high quality relays. These are, however, inherently rather slow. The high-speed requirement could be satisfied either by electronic circuitry, which would be either prohibitively expensive for this noise level requirement, or by rotary sampling switches. The latter, while quiet and fast enough, are extremely restricted in their life span.

### Aperture Corrective Systems—J. Otterman (p. 8)

This paper is concerned with the averaging or smoothing effect common to a large variety of sensing devices, *i.e.*, devices that convert a signal in one form of energy into a signal in another form of energy. The distortion is known as an aperture effect since it appears in the instrumentation system that converts spot-by-spot light intensity in a picture into a continuous electric signal by aperture scanning. The correction for this effect can be achieved by incorporating a compensating electrical network in the instrumentation system.

The compensating circuits described in this paper represent a new approach to the problem of compensation. The transfer functions of the circuits suggested are inverse to those of the apertures, except for small departures introduced to insure stability or caused by losses in delay lines. The common property of all these circuits is that they incorporate a delay device in the feedback loop of an active element or a delay line mismatched at both ends. The impulse responses possess some repetitive character.

The compensating method offers special

advantages when some preferred forms of aperture, such as the bi-rectangular or exponential, can be used.

**Application of Quasi-Peak Detector to the Measurement of Probability Density Function**—K. Y. Ya'coub (p. 19)

An application of the quasi-peak detector, which is found in radio noise meters, is considered as a device for measuring probability density function. Two procedures are discussed: one for obtaining the probability density function at discrete values of the variable, the second for approximating the entire function by using Edgeworth series.

**Relative Voltmeter for "VHF/UHF Signal Generator Attenuator Calibration"**—B. O. Weinschel, G. U. Sorgor, and A. L. Hedrick (p. 22)

This paper describes a system for the calibration of a signal generator. The system includes a linear mixer, a standard 30-mc source, a piston attenuator and a detecting system. Changes in the signal generator output are matched against changes in the piston attenuator, which is used as a standard. It is capable of measuring insertion loss, or differences in signal level over a range of 62 db with an accuracy of 0.02 db per 10 db.

**IF Stabilization by Regenerative Frequency Conversion**—D. M. Makow (p. 32)

A feedback system is considered using three mixers M1, M2 and M3. An external input frequency  $f_0$  and an internally generated frequency  $f_b$  produce in M1 the difference  $f_0 - f_b = f_a$ , which is applied to M2 and M3. Also  $f_b$  is applied to M2, which then generates  $f_d = f_0$ . This frequency is in turn fed to M3, where  $f_b$  is obtained as a difference of  $f_d$  and  $f_a$ . The frequency  $f_b$  closes the loop providing the originally assumed inputs to M1 and M2. The output frequency of M2 is identical with  $f_0$ , while  $f_b$  is determined by the zero-phase shift condition in the oscillatory loop formed by the branches connecting M2 and M3.

In a converter where  $f_a$  is chosen much smaller than  $f_b$ ,  $f_a$  can be thought to be the IF and  $f_b$  the LO frequency. When certain design requirements are fulfilled, variations in the input frequency  $f_0$  will not greatly affect  $f_a$  or  $f_b$ . Also, improvement in the stability of  $f_b$  due to drifts in  $d$  and  $b$  can be shown to be possible, where  $d$  and  $b$  are the phase angles introduced to  $f_d$  and  $f_b$ .

Correspondence (p. 36)

## Vehicular Communications

VOL. VC-12, APRIL, 1959

*Ninth National Conference of the PGVC, December 4-5, 1958*

**Investigation of Antennas for Two-Way Mobile Communications in the VHF and UHF Regions**—E. F. Harris (p. 2)

Due to the nature of two-way mobile communications the normal requirement is for base station antennas, as well as the vehicular antennas, to have radiation characteristics which are omnidirectional in azimuth. It is therefore necessary first to investigate the spatial distribution of energy and then to investigate means for confining the radiation in the elevation plane so far as possible to that region of space which represents the path for communication. Scale model techniques have been employed using 1/20 scale automobile mounted on a turntable ground plane to yield useful radiation pattern information for the mobile unit. Data are given for the quarter-wave roof mounted whip antenna and a bumper mounted coaxial antenna, plus a colinear array. For the base station, as well as the mobile unit, it is pointed out that the majority of communication takes place within one degree of the horizon. Pattern shaping techniques and tower

and pole mounting effects are shown for various situations of side tower mounted arrays, both to allow for the tower consideration in those systems where side of tower mounting is a necessity and to enable the designer to take advantage of tower shaping situations where they can be used to increase gain in the desired directions. A newly developed three wavelength aperture colinear array for the vehicular unit for the 450 mc range is described. This unit is approximately 6 feet long over all and tests indicate a great improvement in systems performance over standard quarter-wave whip or dipole type antennas.

**Antenna-to-Mast Coupling in Communication**—M. W. Scheldorf (p. 5)

The use of vertically polarized antenna systems in the communication industry has been standard for quite a period of time. This standard results in a major part from the need for omnidirectional patterns with mobile stations and the fact that this solution is made in the simplest most direct manner by using a vertical conductor mounted over some metallic surface of the vehicle. In this application the vehicle surface becomes a very satisfactory horizontally disposed reflector, and the operating structure becomes fundamentally what is known as a ground plane antenna. That this choice of polarization places a handicap on the design of the station antenna is not a commonly well known fact. The problem in this connection arises from the inability to provide relatively large conducting ground plane surfaces at the top of tower structures used to mount these antennas at suitable heights for effective propagation. This paper is intended to discuss the problem in detail and to describe tests for measuring the effectiveness of solutions that are developed.

**Identifying 450 MC Interference Using a Tunable Receiver and a Panadaptor**—R. J. Klein (p. 13)

Many mobile systems operating in the 450-470 mc band are troubled by radio interference. Much of this interference is caused by intermodulation and spurious radiation from other systems operating in the same geographical area and frequency band. Due to the nature of intermodulation and spurious radiation interference, any intelligence originating from the interfering stations is obscured. This makes it impossible to identify them. If the stations causing the interference cannot be identified, it is difficult to take steps to eliminate the interference.

This paper describes how a 450-470 mc tunable receiver and Panadaptor can be used to identify interference. By scanning the band visually and aurally in 200 kc steps, the fundamental frequencies of the interfering stations can be found. Knowing these frequencies, the stations can be identified.

**Transmission Test of a Trial System of Telephone Service for Aircraft**—R. V. Crawford (p. 20)

This paper gives results of transmission tests on the air-to-ground telephone system which is presently in operation at 450 mc on a trial basis with land base stations located at Chicago and near Detroit. Received carrier level measurements, using a chart recorder, were made at both land stations prior to the start of commercial service on September 15, 1957. The effects of base station antenna height, aircraft elevation and distance from base station, nulling, capture in the overlap zone, and aircraft antenna patterns are discussed.

**Shadows of the Future in Vehicular Communication**—D. C. Pinkerton (p. 27)

This paper constitutes a discussion of certain changes in vehicular communication services and equipment that may be expected in the not too remote future. These changes are related, not only to technology, but also to the problems created by the slow but sure depletion

of the available radio frequency spectrum space.

Attention is directed to the great potential that exists for better utilization of the spectrum through improvements in the space (or volume) and time dimensions. It is pointed out that the effective use of air time through automation of routine or repetitive type communications promises to give, in effect, a many-fold increase in the available spectrum.

**A New Era in Communications Through Transistors**—W. J. Weisz (p. 35)

**Bias Considerations in Transistor Applications to Communications Circuitry**—A. G. Manke (p. 47)

**Transistorized Power Supplies for Higher Input Voltages**—L. S. Pearlman (p. 54)

**The Use of Miniature Tubes in Class C Circuits**—A. Dzik (p. 58)

Receiving-type tubes designed primarily for small-signal class A applications are being used in an increasing number of class C circuits in mobile transmitters and other revenue bearing industrial equipment requiring a high degree of reliability. Tube performance in such circuits is often limited by the effects of class C operation on the individual elements of the tubes. This paper analyzes such effects in the RCA 7054, a miniature sharp-cutoff pentode designed for use as a low-level frequency multiplier in mobile service, and describes circuit-design techniques which can be used to improve equipment reliability.

**Application of Vehicular Communications to Railroads**—L. R. Thomas (p. 62)

**Selective Signaling in the Bell System—Relay to Transistor**—W. G. Chaney (p. 67)

Selective Signaling is used in Domestic Public Land Mobile Systems by many telephone companies to individually call vehicles equipped for mobile telephone service. Problems of reliability, speed, simplicity and number capacity, rule out many possible systems. A mechanical device used successfully for the past years is now proving both bulky and costly so that newer devices are under consideration.

A description of the present day system is given in this paper inasmuch as any new equipment will be expected to perform in a compatible manner.

**A Transistorized Selector Device for Bell System Mobile Telephone Service**—J. R. Scantlin (p. 71)

**A Digital Selective Signaling System for Mobile Radio**—J. H. Green and J. Gordon (p. 74)

A system for remote operation of a particular receiver selected from any number of similar receivers, all using a very narrow bandwidth, is described. Two tones that are turned on or off according to an address-determining binary sequence are transmitted to a group of receivers, one of which will recognize the address, or call, and activate a control circuit. A small transistorized experimental model with low power consumption, was built to prove the system.

**New Ruggedized Hand-Held Microphone for Mobile Radio Equipment**—A. Brouns, G. Salisbury, and R. Troxel (p. 86)

The increased use of mobile radio equipment has resulted in an emphasis on intelligibility, reliability and a decrease in cost and weight.

The use of the controlled magnetic cartridge in place of carbon has done a very effective job of increasing intelligibility.

The use of plastic in the case of a hand held microphone has resulted in a reduction in both cost and weight. Using some of the new high impact plastics, the case demonstrates physical characteristics that exceed those of metal on almost every point. This plus other modifications in the mechanical and electrical design of the unit gives a microphone that is more reliable at extremes in operational environmental conditions that are incurred in mobile use.

# 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 papers, nor does it have reprints of the articles abstracted. Correspondence regarding these articles and requests for their procurement should be addressed to the individual publications, not to the IRE.

Acoustics and Audio Frequencies.....	1176
Antennas and Transmission Lines.....	1176
Automatic Computers.....	1177
Circuits and Circuit Elements.....	1177
General Physics.....	1179
Geophysical and Extraterrestrial Phenomena.....	1179
Location and Aids to Navigation.....	1182
Materials and Subsidiary Techniques.....	1182
Mathematics.....	1184
Measurements and Test Gear.....	1184
Other Applications of Radio and Electronics.....	1185
Propagation of Waves.....	1185
Reception.....	1186
Stations and Communication Systems.....	1186
Subsidiary Apparatus.....	1186
Television and Phototelegraphy.....	1187
Tubes and Thermionics.....	1187
Miscellaneous.....	1188

The number in heavy type at the upper left of each Abstract is its Universal Decimal Classification number. The number in heavy type at the top right is the serial number of the Abstract. DC numbers marked with a dagger (†) must be regarded as provisional.

## ACOUSTICS AND AUDIO FREQUENCIES

**534.213.4** **1422**  
Influence of the Thickness of the Walls of a Resonance Tube on the Velocity of Sound in Liquids—C. Sălceanu and M. Zăgănescu. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 812–814; September 15, 1958.)

**534.231-8-14** **1423**  
Theoretical Investigation of an Ultrasonic Field—N. Segard and J. Cassette. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 809–812; September 15, 1958.) The changes of pressure in a liquid caused by the vibrations of a piston-like quartz resonator are evaluated.

**534.414:534.84** **1424**  
Investigations of Nonlinear Helmholtz Resonators—F. Barthel. (*Frequenz*, vol. 12, pp. 72–82; March, 1958.) The causes of non-linearity of the absorption coefficient in resonators as a function of sound intensity are analyzed. Emphasizing these nonlinearities may be useful in improving room acoustics.

**534.6:681.8** **1425**  
The Use of the Acoustic Spectrograph and the Problem of the Score in Experimental Music—A. Moles and V. Ussachevsky. (*Ann. Télécommun.*, vol. 12, pp. 299–304; September, 1957.)

**534.75** **1426**  
Investigation of the Sensation of Loudness of Rhythmic Sounds—H. Niese. (*Hochfreq. und Elektroak.*, vol. 66, pp. 115–125; January, 1958.) Subjective tests on sinusoidal tones with sine-wave, rectangular-wave or pulse modulation show that loudness appears to increase at modulation frequencies below about 100 cps. Results are compared with calculations of this "inertia" effect in hearing, and conclusions are drawn for the design of objective loudness meters.

The Index to the Abstracts and References published in the PROC. IRE from February, 1958 through January, 1959 is published by the PROC. IRE, May, 1959, Part II. It is also published by *Electronic and Radio Engineer*, incorporating *Wireless Engineer*, and included in the March, 1959 issue of that journal. Included with the Index is a selected list of journals scanned for abstracting with publishers' addresses.

**534.79** **1427**  
Proposal for a Loudness Meter for an Acoustically True Indication of Sounds with Peaked Waveform in Acoustic Fields of Any Form—H. Niese. (*Hochfreq. und Elektroak.*, vol. 66, pp. 125–139; January, 1958.)

**534.84** **1428**  
On the Use of the Iteration Method in Room Acoustics—A. Moles. (*Ann. Télécommun.*, vol. 12, pp. 443–444; December, 1957.) A study of the progressive distortion of music or speech shows a separation of the relative constants of semantic and aesthetic information in each.

**621.395.61:621.3.082** **1429**  
Research on Probe Microphones—I. Barducci and V. Degano. (*Ann. Télécommun.*, vol. 12, pp. 436–442; December, 1957.) Optimum dimensions for a probe microphone have been obtained from a study of frequency response curves plotted for various dimensions of the probe tube and cavity. The construction of the probe, experimental response curves and a formula for calculating the frequency response are given.

**621.395.61.089.6** **1430**  
Various Methods of Absolute Calibration of Standard Microphones—R. Lehmann. (*Ann. Télécommun.*, vol. 12, pp. 393–408; November, 1957.) 30 references.

**621.395.612.4:534.773.2** **1431**  
Technical Fundamentals and New Aspects of the Construction of Miniature Magnetic Microphones—W. O. Holleufer. (*Elektrotech. Z., Edn A*, vol. 79, pp. 533–536; August 1, 1958.) The design and construction of microphones for hearing aids is discussed.

**621.395.616** **1432**  
The Electrostatic Uniaxial Microphone—H. F. Olson and J. Preston. (*J. Soc. Mot. Pic. Telev. Eng.*, vol. 67, pp. 750–753; November, 1958.) Description of a lightweight microphone comprising an es transducer combined with an acoustic network. It has a uniform frequency response and good directivity.

**621.395.623.7** **1433**  
The Response of Loudspeakers near the Principal Resonance—E. Paolini. (*Ann. Télécommun.*, vol. 12, pp. 387–391; November, 1957.) A theoretical study in which the loudspeaker is assumed to be fed by constant-current or constant-voltage signals whose frequency is compared with pressure along the loudspeaker axis.

**621.395.623.8:621.396.721** **1434**  
Sound Distribution at the Brussels Exhibition—Martin. (See 1691.)

## ANTENNAS AND TRANSMISSION LINES

**621.372.2:621.372.826** **1435**  
On the Propagation of Surface Waves over an Infinite Grounded Ferrite Slab—R. L. Pease. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 13–20; January, 1958. Abstract, *PROC. IRE*, vol. 46, p. 799; April, 1958.)

**621.372.8.002.2:621.357.6** **1436**  
The Electroforming of Waveguide Components—P. Andrews. (*Electronic Eng.*, vol. 31, pp. 150–152; March, 1959.) Details are given of a simple acid-copper electroforming process for use in laboratory or small-scale production. The bath, electrical supply and mandrel materials required are described.

**621.372.81.09:061.3** **1437**  
Waveguide Transmission—(*Wireless World* vol. 65, pp. 104–105; March, 1959.) A report is given of an IEE Convention on "Long-Distance Transmission by Waveguide" held in London, January 29–30, 1959. Problems associated with the use of the  $H_{01}$  mode in circular waveguides and with the single-wire surface-wave system are briefly discussed.

**621.372.826:537.226** **1438**  
Excitation of a Transverse Magnetic Surface Wave Propagated on a Dielectric Cylinder—C. Jauquet. (*Ann. Télécommun.*, vol. 12, pp. 217–233; June, 1957.) Detailed theoretical treatment (see 3031 of 1957) with a note of an experimental investigation of the surface-wave characteristics of a polystyrene rod.

**621.372.852.323:621.318.134** **1439**  
Resonance Isolator at 70 kMc/s—L. C. Kravitz and G. S. Heller. (*PROC. IRE*, vol. 47, p. 331; February, 1959.) Describes a precessional-resonance waveguide isolator using an oriented Ba ferrite having an effective internal magnetization of 17,000 G. With an additional external field of 8000 G the isolator gives an insertion loss of 1 db and a reverse loss of 12 db.

**621.396.67** **1440**  
Driving Point and Input Admittance of Linear Antennas—T. T. Wu and R. W. P. King. (*J. Appl. Phys.*, vol. 30, pp. 74–76; January, 1959.) "An infinity in the input admittance of linear antennas owing to the use of an idealized delta-function generator is investigated. It is shown that the infinity may be interpreted in terms of an infinite capacitance

between the two halves of the antenna. The conclusion is reached that conventionally used iterative procedures are not invalidated by difficulties with respect to the driving point."

621.396.67:621.315.1:621.396.11.029.45 1441  
Power-Line Aerial—Golden, Langmuir, Macmillan and Rusch. (See 1672.)

621.396.67.012.12 1442  
On the Fresnel Approximation—R. B. Barrar and C. H. Wilcox. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 43-48; January, 1958. Abstract, PROC. IRE, vol. 46, p. 799; April, 1958.)

621.396.674 1443  
Four Simultaneous Transmissions from One Aerial—F. D. Bolt. (*Electronic Eng.*, vol. 31, pp. 168-170; March, 1959.) The antenna is a single slotted cylindrical type which radiates an effective power of 120 kw omnidirectionally on 88.1, 90.3, 92.5 and 94.7 cps. The input voltage swr is not greater than 1.2.

621.396.674.3 1444  
Back-Scattering Cross-Section of a Centre-Loaded Cylindrical Antenna—Yueh-Ying Hu. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 140-148; January, 1958. Abstract, PROC. IRE, vol. 46, p. 800; April, 1958.)

621.396.677 1445  
The Prolate Spheroidal Antenna: Current and Impedance—C. P. Wells. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 125-128; January, 1958. Abstract, PROC. IRE, vol. 46, p. 800; April, 1958.)

621.396.677.32 1446  
End-Fire Echo Area of Long, Thin Bodies—L. Peters, Jr. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 133-139; January, 1958. Abstract, PROC. IRE, vol. 46, p. 800; April, 1958.)

621.396.677.833 1447  
An Omnidirectional Vertically Polarized Paraboloid Aerial—E. O. Willoughby and E. Heider. (*Proc. IRE (Australia)*, vol. 19, pp. 554-555; October, 1958.) A preliminary note giving a brief description of the antenna. At 28.7 cm  $\lambda$  the half-power beam width is 6.6°.

621.396.677.833 1448  
Wide-Angle Scanning with Microwave Double-Layer Pillboxes—W. Rotman. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 96-105; January, 1958. Abstract, PROC. IRE, vol. 46, p. 800; April, 1958.)

621.396.677.833.2 1449  
Surface-Wave Beacon Antennas—R. E. Plummer. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 105-114; January, 1958. Abstract, PROC. IRE, vol. 46, p. 800; April, 1958.)

621.396.677.85 1450  
Experimental Test on Some Microwave Configuration Lenses—P. F. Checcacci and V. Russo. (*Alta Frequenza*, vol. 27, pp. 92-107; April, 1958.) Report on tests carried out with two confection-refraction lenses and a confection doublet [see e.g. *J. Opt. Soc. Amer.*, vol. 45, pp. 621-624; August, 1955. (Torald di Franca)]. Radiation and wavefront diagrams are given.

621.396.679.1 1451  
The Distribution of Lightning Currents in the Earthing Systems of a Radio Mast—W. Griesinger, E. Popp and E. Schulz. (*Elektrotech. Z., Edn A*, vol. 79, pp. 526-529; August 1, 1958.) The protection of telephone cables anchored to radio masts is considered; this can

be achieved by enlarging the earthing system, particularly in mountainous regions, and by special screening of the telephone cable.

#### AUTOMATIC COMPUTERS

681.142 1452  
Digital Techniques for Small Computations—Y. Lundh. (*J. Brit. IRE*, vol. 19, pp. 37-44; January, 1959.) A special digital method of computing simple algebraic functions with one to four variables is described and analyzed. The method is particularly suited to division and root extraction.

681.142 1453  
Automatic Failure Recovery in a Digital Data Processing System—R. J. Doyle, R. A. Meyer and R. P. Pedowitz. (*IBM J. Res. Developm.*, vol. 3, pp. 2-12; January, 1959.) A program is described which automatically compensates for computer malfunctions so that recovery from errors may be affected with a negligible loss of operational time.

681.142 1454  
Representation, Grouping and Processing of Information in Automatic Data Processing—H. Zschechel. (*Elektrotech. Z., Edn A*, vol. 79, pp. 617-624; September 11, 1958.)

681.142 1455  
A Sensing System for Punched Cards or Continuous Punched Foil—S. Morleigh. (*Electronic Eng.*, vol. 31, pp. 140-141; March, 1959.)

681.142 1456  
Logic Synthesis of Some High-Speed Digital Comparators—M. Nesenbergs and V. O. Mowery. (*Bell. Sys. Tech. Jour.*, vol. 38, pp. 19-44; January, 1959.) Logical schemes are derived by Boolean algebra methods. Circuits satisfying the synthesis requirements and giving either exact or approximate proportional comparison are described.

681.142 1457  
Function Generator for Sines or Cosines—H. Schmid. (*Electronics*, vol. 32, pp. 48-50; January 23, 1959.) Operation of the transistorized unit described is based on the fact that the area under a sine curve varies as a cosine function. Basic components required are a linear pulse-width modulator, an ideal switch (two  $p-n-p$  transistors connected emitter-to-emitter), an integrator and a sine-wave source.

681.142:537.311.33:546.281.26 1458  
An Analogue Multiplier-Divider Using Thyrites—A. A. Maslov. (*Aytomalika i Telemekhanika*, vol. 18, pp. 336-348; April, 1957.) A new analytical method is proposed for determining quadratic elements using thyrite resistors. A description is given of a new multiplier with thyrite units which requires only two operational amplifiers. Experimental results are given.

681.142:621-526 1459  
Logic for a Digital Servo System—R. W. Ketchledge. (*Bell. Sys. Tech. Jour.*, vol. 38, pp. 1-17; January, 1959.) A crt beam in a photographic storage system is positioned by comparing a binary address with a digital indication of the present position. Digital servo logic circuitry is described for obtaining the sign and magnitude of the positional error.

681.142:621.318.5 1460  
An Electron Analyser of Contact Circuits—V. N. Rodin. (*Aytomalika i Telemekhanika*, vol. 18, pp. 437-443; May, 1957.) A description is given of an apparatus for rapid determination of possible combinations in the operation of a circuit comprising six relays. One of the possible

uses of the apparatus is the simplification of the existing circuits by removing the unnecessary contacts.

681.142:621.318.57 1461  
The Multipurpose Bias Device: Part 2—The Efficiency of Logical Elements—B. Dunham, D. Middleton, J. H. North, J. A. Sliter and J. W. Weltzien. (*IBM J. Res. Developm.*, vol. 3, pp. 46-53; January, 1959.) "The efficiency of a logical element can be equated with the set of subfunctions it realizes upon biasing or duplication of inputs. Various classes of elements are considered, and optimum or near-optimum examples are presented. Some related areas of study are suggested." Part 1: 38 of 1958 (Dunham).

681.142:621.318.57:538.221 1462  
Millimicrosecond Magnetic Gating and Storage Element—D. A. Meier. (*J. Appl. Phys.*, vol. 30, pp. 122-123; January, 1959.) The element consists of a round glass rod which is first chemically plated with a silver conductor and then electrodeposited with an Fe-Ni film. The element has an almost square saturation loop.

681.142:621.318.57:621.318.134 1463  
The Laddic—a Magnetic Device for Performing Logic—U. F. Gianola and T. H. Crowley. (*Bell. Sys. Tech. Jour.*, vol. 38, pp. 45-72; January, 1959.) The laddic is a ladder-like structure cut from rectangular-hysteresis-loop ferrite. All magnetic paths are flux-limited. By controlling the actual switching path any Boolean function of  $n$  variables can be produced.

681.142:621.395.625.3 1464  
Tape Recording System Speeds Data Processing—Way Dong Woo. (*Electronics*, vol. 32, pp. 56-68; February 6, 1959.) A PWM tape recording system is described with a capacity of 60,000 digits. The tape is three-inches wide and carries 36 channels. The transport mechanism is electronically controlled and can start, stop and reverse within a few milliseconds.

681.142:621.396.962.012 1465  
An Analogue Computer for Evaluating Radar Performance—C. C. Willhite and J. W. McIntyre. (*Bell. Lab. Rec.*, vol. 37, pp. 16-19; January, 1959.) Describes briefly the basic design of an analogue computer developed to evaluate the statistical information needed to assess radar tracking performance.

681.142:629.13 1466  
Application of Analogue Calculation to Flight Simulators—J. W. Swift. (*Rev. HF, Brussels*, vol. 4, pp. 1-10; 1958.) Computer elements for simulating the reactions resulting from the effects of gyration, pitch and roll are described.

#### CIRCUITS AND CIRCUIT ELEMENTS

621.314.2:621.372.4 1467  
The Transformer as a Two-Terminal Network—W. Klein. (*Arch. elekt. Übertragung*, vol. 12, pp. 133-137; March, 1958.) The equivalent two-terminal networks are derived for various types of transformer.

621.314.2:621.372.51 1468  
On Hybrid Transformers—H. O. Friedheim. (*A.T.E. J.* vol. 14, pp. 218-228; July, 1958.) Design formulas are derived from basic assumptions, and the properties and some common uses of these circuit elements are discussed.

621.314.2.073.3:621.318.124:621.374.3 1469  
Premagnetization of the Core of a Pulse Transformer by Means of Ferroxdure—H. G.

- Brujning and A. Rademakers. (*Philips Tech. Rev.*, vol. 19, pp. 28-37; July 27, 1957.) Premagnetization to improve the loading of the core of a transformer carrying unidirectional pulse currents is examined theoretically. A practical transformer is described.
- 621.314.5:621.314.7** 1470  
**High-Power Transistor D.C. Converters**—T. R. Pye. (*Electronic Radio Eng.*, vol. 36, pp. 96-105; March, 1958.) Circuits are considered for dc per dc or dc per ac converters using Ge and Si power transistors which can give power outputs of up to 100 watts.
- 621.318.57:621.318.435** 1471  
**Saturable-Transformer Switches**—B. D. Simmons. (*Electronic Radio Eng.*, vol. 36, pp. 82-90; March, 1959.) A switch using cores in a balanced saturable transformer is considered. An on-off impedance ratio adequate for switching magnetic recording heads for reading and writing operations can be obtained. Circuits are described for operation of a matrix selection system with a switching time of a few tens of microseconds.
- 621.319.4:537.227:621.396.662** 1472  
**Ferroelectrics Tune Electronic Circuits**—T. W. Butler, Jr. (*Electronics*, vol. 32, pp. 52-55; January 16, 1959.) Circuits are described in which varying bias is used to change the capacitance of ferroelectric capacitors used as tuning elements.
- 621.319.4.015.5** 1473  
**Direct-Voltage Instantaneous Breakdown of Oil-Impregnated Paper Capacitors as a Function of Area**—D. S. Girling. (*Elect. Commun.*, vol. 35, pp. 83-92; 1958.) The experimental results given show that the variation of mean breakdown voltage with increasing area is represented adequately by Milnor's equation. The coefficient of variation appears to be independent of area but decreases with increasing thickness of dielectric.
- 621.319.42:621.315.616.98** 1474  
**A Comparison of Wool Wax and Petroleum Jelly as Impregnants for Paper Capacitors**—J. S. Dryden and R. J. Meakins. (*Proc. IRE (Australia)*, vol. 19, pp. 551-535; October, 1958.) Results given show that wool wax impregnants give about 10 per cent higher capacitance without any serious loss in performance.
- 621.372.412.012.8** 1475  
**The Equivalent Circuit of Oscillating Piezoelectric Crystal Rods**—J. Tichý. (*Ann. Phys. (Lpz.)*, vol. 1, pp. 219-231; May 20, 1958.) Formulas are derived for calculating the circuit parameters for longitudinal, torsional and flexural oscillations taking account of internal damping and the air gap between electrodes and crystal.
- 621.372.45** 1476  
**Voltage- and Current-Controlled Negative-Resistance Two-Poles**—L. Pigiore. (*Alta Frequenza*, vol. 27, pp. 138-152; April, 1958.) The mechanism of voltage and current control in two-pole networks is discussed. Equations for obtaining the equivalent circuits from direct measurements at the network terminals are given.
- 621.372.5** 1477  
**The Geometrically Simplest Form of Representation of the General Lossy Quadripole**—J. de Buhr. (*Arch. elekt. Übertragung*, vol. 12, pp. 127-132; March, 1958.) The transformation discussed is based on a construction of four planes perpendicular to each other in pairs in a non-Euclidean system. It can be represented in a drawing by stereographic projection.
- 621.372.5** 1478  
**Coupling Coefficients of Ladder Networks with Maximally Flat Amplitude Response**—R. A. Waldron. (*J. Brit. IRE*, vol. 19, pp. 63-71; January, 1959.) The coefficients are expressed in terms of the dissipations in the input and output branches, using a generalized theory of Green's method of solution (370 of 1955).
- 621.372.5** 1479  
**H.F. Wide-Band Electronic Integrator Design**—H. Hodara. (*Electronic Ind.*, vol. 17, pp. 96-100; October, 1958.) Compensation for distortion of the integrated waveform due to valve output capacitances can be introduced by the addition of a capacitor in parallel with the integrating-network resistance. An expression is derived for calculating this capacitance for the basic RC integrator and the Miller integrator.
- 621.372.51** 1480  
**Impedance Converters**—F. Molo. (*Alta Frequenza*, vol. 27, pp. 120-137; April, 1958.) The quadripole relations and equivalent circuits of negative-impedance converters are summarized and applied to practical unbalanced converter circuits with valves and transistors.
- 621.372.55** 1481  
**A Laguerre Function Equalizer**—H. Mumford and E. J. Osborne. (*A.T.E. J.*, vol. 14, pp. 196-202; July, 1958.) A waveform corrector is described in which the input signal is successively differentiated and the differentials added to or subtracted from the input, the waveforms being given by Laguerre functions. Results are given for an experimental equalizer which need contain only passive elements.
- 621.373.4-523.6** 1482  
**Constant-Frequency Oscillators without Stabilized Voltage Supplies**—W. Wisotzky. (*Elektronik*, vol. 7, pp. 79-80; March, 1958.) Four valve oscillators controlled by thermistor bridge circuits are described.
- 621.373.421.13** 1483  
**The Design of Crystal-Controlled Three-Terminal Valve Oscillators**—F. Rockstuhl. (*Telefunken Ztg.*, vol. 31, pp. 50-58; April, 1958. English summary, pp. 68-69.) Design formulas for the Pierce-Miller and Pierce-Colpitts circuits are derived taking account of the effect of crystal resistance on frequency.
- 621.373.421.13** 1484  
**Stable, Low-Cost 1-Mc/s Oscillator**—J. F. Mercurio, Jr. (*Electronics*, vol. 32, pp. 50-51; February 6, 1959.) Description of a crystal-controlled transistorized oscillator with frequency stability within one part in  $10^9$  per day at normal room temperatures. Total weight of oven and oscillator is 2.5 lb.
- 621.373.421.13:621.396.666:621.396.11** 1485  
**Frequency Stepper for Radio Propagation Tests**—E. H. Hugenholz, A. Seljak and A. Towle. (*Electronics*, vol. 32, pp. 44-46; January 23, 1959.) Full circuit details are given of a stepped-frequency exciter unit controlling a pulse transmitter and receiver at widely separated locations. The range 25.05-48.95 mc is covered in 100-kc steps at 1-sec intervals. 100-kc and 1-mc crystals are used as reference-frequency sources, and frequency control is achieved by means of a beam-deflection tube. Transmitter and receiver circuits are briefly described: the transmitter radiates 25- $\mu$ s pulses with repetition frequency 30 per sec or 15 per sec and output power 10-20 kw.
- 621.373.421.14:621.385.029.64** 1486  
**A Barkhausen-Kurz Oscillator at Centimetre Wavelengths**—E. M. Boone, M. Ueno-
- hara and D. T. Davis (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 196-205; July, 1958. Abstract, PROC. IRE, vol. 46, pp. 1890-1891; November, 1958.)
- 621.373.421.14:621.396.962.23** 1487  
**Dual-Cavity Microwave Discriminator**—M. Rudin, R. E. Shafer and B. W. Baker. (*Electronics*, vol. 32, pp. 74, 76; January 16, 1959.) Intended for stabilizing reference klystron oscillators used in Doppler-shift radar systems. Tuning range is  $\pm 50$  mc thermal drift 1 part in  $10^7$  per  $^\circ$ F, and discriminator sensitivity 1.5 v per mc.
- 621.373.431** 1488  
**The Generation of Extremely Steep Pulse Edges in Multistage Nonlinear Amplifiers**—G. Kohn. (*Arch. elekt. Übertragung*, vol. 12, pp. 109-118; March, 1958.) The pulse generator described incorporates three transformer-coupled amplifier stages for sharpening the multivibrator pulses; it can supply 100-v 50- $\mu$ s pulses with a rise time  $< 3$   $\mu$ s into a 60- $\Omega$  load.
- 621.373.52:621.376.3** 1489  
**Transistorized F.M. Oscillator**—P. W. Wood. (*Electronics*, vol. 32, p. 64; January 30, 1959.)
- 621.375.1.018.424** 1490  
**The Amplification of Very Wide Frequency Bands**—K. H. Fischer and G. Daisenberger. (*Elektrotech. Z., Edn A.*, vol. 79, pp. 625-632; September 11, 1958.) Various techniques for amplification in bandwidths exceeding 100 mc are discussed and the suitability of cascade and distributed amplifiers is investigated. Direct and indirect methods of delay-line equalization are outlined and details are given of a seven-stage distributed amplifier with 10-db gain over a 125-mc band.
- 621.375.121** 1491  
**Wide-Band Amplifier Design Data**—R. H. Engelmann. (*Electronics*, vol. 32, pp. 43-49; February 6, 1959.) Design data are tabulated for three basic types of coupling, RC, shunt-peaked and series-peaked. Gain is calculated using a series of approximations taking account of the gain reduction factor for a number of stages.
- 621.375.121.2** 1492  
**Extended-Range Distributed-Amplifier Design**—J. J. Eichholz, C. F. Nelson and G. T. Weiss. (*Rev. Sci. Instr.*, vol. 30, pp. 1-6; January, 1959.) "A distributed amplifier employing a straight-wire transmission line is described. The amplifier has a gain of  $15 \pm 3$  db from 10 to 400 mc, and input and output impedances of 50  $\Omega$ . Detailed electrical and mechanical design considerations are included."
- 621.375.132.3.029.3** 1493  
**An Inexpensive Ultralinear Output Stage**—I. F. Barditch. (*Electronic Ind.*, vol. 17, p. 89; October, 1958.) A circuit is described which uses two cathode followers in place of a tapped output transformer.
- 621.375.23:621-52** 1494  
**Operational Amplifier without a Stabilized Power Supply**—V. M. Evseev. (*Aytomatika i Telemekhanika*, vol. 18, pp. 427-436; May, 1957.) A new operational amplifier is proposed in which two parallel circuits for low and high frequencies are used, with a summing amplifier and two feedback circuits. The theory of the amplifier is discussed and some experimental results are given. For a mains voltage variation of  $\pm 10$  v, the output voltage of the amplifier varies by 16 mv with a drift rate of 45  $\mu$ v sec.
- 621.375.3** 1495  
**On the Theory of a Half-Wave Magnetic**

**Amplifiers**—R. A. Lipman and I. B. Negnevitskii. (*Ayatomalika i Telemekhanika*, vol. 18, pp. 349–370 and 449–465; April and May, 1957.) A detailed theoretical analysis is given of an amplifier proposed by Ramey (see 3503 and 3507 of 1953). Experimental results confirm the theoretical conclusions.

**621.375.4** 1496  
**Differential Amplifier Features D.C. Stability**—W. T. Matzen and J. R. Biard. (*Electronics*, vol. 32, pp. 60–62; January 16, 1959.) Drift in transistorized amplifiers is analyzed and circuits for reducing drift are described for a particular amplifier. A 12-h test showed the equivalent input drift to be less than 60  $\mu$ v from 70° to 80°C.

**621.375.4(083.57)** 1497  
**Nomographs for Designing Transistor Narrow-Band Amplifiers**—L. M. Krugman. (*Electronic Ind.*, vol. 17, pp. 78–81; October, 1958.)

**621.375.4.024** 1498  
**Transistor Amplifiers for D.C. Signals**—H. Keimbhadjian. (*Mullard Tech. Commun.*, vol. 4, pp. 162–172; December, 1958.) Basic direct-coupled and chopper-type amplifiers are examined and compared.

**621.375.4.029.3:534.6** 1499  
**A Low-Noise Transistor Amplifier for Acoustic Measurements**—E. R. Hauri. (*Tech. Mill. PTT*, vol. 36, pp. 142–144; April 1, 1958.) The portable equipment described has a gain of about 60 db and a frequency range 30 cps–15 kc with noise output 3.5 db above the theoretical minimum.

**621.375.9:538.569.4.029.6** 1500  
**The Maser: a New Form of Microwave Oscillator**—C. M. Cade. (*J. Telev. Soc.*, vol. 8, pp. 509–511; October–December, 1958.) A short review of the principles of operation of molecular-beam and solid-state masers.

**621.375.9:538.569.4.029.6** 1501  
**Small-Signal Analysis of Molecular-Beam Masers**—J. C. Helmer. (*J. Appl. Phys.*, vol. 30, pp. 118–120; January, 1959.) A perturbation calculation gives the resonance polarization of a molecule in an electric field as a function of time. An analysis is then made of maser operation with a divergent univelocity beam.

**621.375.9:538.569.4.029.6** 1502  
**Three-Level Solid-State Maser**—S. M. Bergmann. (*J. Appl. Phys.*, vol. 30, pp. 35–36; January, 1959.) The maximum values of the real and imaginary components of the paramagnetic susceptibility of a three-level solid-state maser are calculated. General expressions for the gain and bandwidth of a traveling-wave maser are given.

**621.375.9.029.64:621.3.011.23** 1503  
**Low Noise in Solid-State Parametric Amplifiers at Microwave Frequencies**—W. E. Danielson. (*J. Appl. Phys.*, vol. 30, pp. 8–15; January, 1959.) “The principles of parametric amplification are described through the use of simple low-frequency electrical circuits and their mechanical analogs. Extension of these principles to low-noise amplification at microwave frequencies is made in a qualitative way which emphasizes the physical processes involved. Amplification is only achieved when energy is transferred from a microwave oscillator or pump to the signal frequency, and it is shown why such a transfer may take place in circuits exhibiting variable capacitance or inductance but not in circuits where only a resistance is varied. Major noise sources are discussed and the special role of the image or idler frequency is noted. Finally, experimental data on four dif-

ferent types of parametric amplifier (three using semiconductor diodes and one using ferrite) are given.”

**621.376.3:621.3.018.78** 1504  
**Calculation of Harmonic Distortion in Sinusoidal Frequency Modulation. Cases of a Waveguide of Any Cross-Section and an Amplification State with a Tuned Transformer**—L. Robin. (*Ann. Télécommun.*, vol. 12, pp. 415–418; December, 1957.) See also 1282 of 1953 (Robin).

#### GENERAL PHYSICS

**537.212:537.291** 1505  
**Potential of an Electrostatic Field and Trajectories of Charged Particles**—G. Sacerdoti and H. Toschi. (*Alta Frequenza*, vol. 27, pp. 108–119; April, 1958.) A procedure is described for calculating the electric potential in the proximity of the path of a charged particle.

**537.222:621.314.63** 1506  
**Solutions for the Static Junction**—L. Gold. (*J. Electronics Control*, vol. 5, pp. 427–431; November, 1958.) Discussion of certain classes of solution of the nonlinear differential equation for the potential distribution near a *p-n* junction. An inverse series solution is developed which is appropriate for diverse impurity distributions.

**537.311.1** 1507  
**A New Method for the Evaluation of Electric Conductivity in Metals**—S. F. Edwards. (*Phil. Mag.*, vol. 3, pp. 1020–1031; September, 1958.) A method is developed for evaluating the closed formal expressions for electrical conductivity which have recently been developed. See e.g. 1105 of 1958 (Kohn and Luttinger) and 2361 of 1958 (Greenwood).

**537.534.8** 1508  
**Secondary Positive Ion Emission from Metal Surfaces**—R. C. Bradley. (*J. Appl. Phys.*, vol. 30, pp. 1–8; January, 1959.) Secondary positive ions ejected from surfaces of Mo, Ta, and Pt under bombardment by inert gas ions of low energy (<1000 eV) have been studied in high vacuum using a mass spectrometer with a 6 inch radius of curvature and a 60° sector field.

**537.56** 1509  
**Current/Voltage Behaviour in a Plasma**—L. Gold. (*J. Electronics Control*, vol. 5, pp. 432–434; November, 1958.) An exact solution for the current/voltage relation in a planar diode filled with a uniform plasma.

**537.56:537.122** 1510  
**Electron Recombination and Negative Bands in Afterglowing Active Nitrogen**—H. H. Brömer and V. Stille. (*Optik*, vol. 15, pp. 382–388; June, 1958.) The emission of negative bands in the auroral afterglow of nitrogen is interpreted as being due to excitation of molecular ions; this is supported by recent measurements of the decay of light intensity of these bands.

**537.56:538.56** 1511  
**A Note on the Confinement of a Plasma by R.F. Fields**—E. S. Weibel. (*J. Electronics Control*, vol. 5, pp. 435–538; November, 1958.) A rigorous derivation of certain results previously obtained by Boot *et al.* (3411 of 1958) by inferior methods.

**537.56:538.561** 1512  
**Microwave Emission from High-Temperature Plasmas**—D. B. Beard. (*Phys. Rev. Lett.*, vol. 2, pp. 81–82; Feb. 1, 1959.) Energy losses from cyclotron radiation are concluded to be negligible for plasma temperatures less than 100 keV and probably for those greater than

100 keV. Emission at the fundamental frequency and first few harmonics is detectable, and measurement of the relative intensities offers a means of determining the plasma temperature.

**537.56:538.566** 1513  
**Excitation of Oscillations in a Plasma Layer**—M. Sumi. (*Phys. Rev. Lett.*, vol. 2, pp. 37–39; January 15, 1959.) Development of a linearized theory of plasma oscillations excited by an electron beam injected into a uniform plasma.

**538.3:535.13** 1514  
**Laws of Magnetism and of Static Electricity in Any Given Frame of Reference**—H. Arzelies and J. Henry. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 815–817; September 15, 1958.) A study of Maxwell's equations in terms of the general theory of relativity is made as a preliminary to attempting to solve practical electrical problems.

**538.566:535.42** 1515  
**On the Diffraction of Electromagnetic Waves at Infinitely Thin Ideally Conducting Flat Screens**—G. A. Grinberg and Yu. V. Pimenov. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2326–2339; October, 1957.) A new method is described based on the solution of two independent equations. As an example, the case of diffraction of a wave at an infinitely thin conducting disk is discussed.

**538.566:535.42** 1516  
**Electromagnetic Diffraction by Dielectric Strips**—D. C. Stickler. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 148–151; January, 1958. Abstract, *PROC. IRE*, vol. 46, p. 800; April, 1958.)

**538.566:535.43** 1517  
**An Analytical Study of Scattering by Thin Dielectric Rings**—L. L. Philipson. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 3–8; January, 1958. Abstract, *PROC. IRE*, vol. 46, pp. 798–799; April, 1958.)

**538.566:535.43** 1518  
**A Statistical Model for Forward Scattering of Waves off a Rough Surface**—L. M. Spetner. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 88–94; January, 1958. Abstract, *PROC. IRE*, vol. 46, p. 800; April, 1958.)

**538.566:535.43** 1519  
**Light Scattering by Small Particles [Book Review]**—H. C. van de Hulst. Publishers: John Wiley and Sons, New York, N. Y., and Chapman and Hall, London, Eng., 1957, 470 pp., 96s. (*Nature, London*, vol. 182, pp. 1470–1471; November 29, 1958.)

#### GEOPHYSICAL AND EXTRATERRESTRIAL PHENOMENA

**523.164+621.396.93** 1520  
**Propagation of Electromagnetic Waves, Radio Location and Radio Astronomy**—Roessler. (See 1581.)

**523.164:523.75** 1521  
**Ionization of the Upper Atmosphere by Low-Energy Charged Particles from a Solar Flare**—H. Leimbach and C. G. Reid. (*Phys. Rev. Lett.*, vol. 2, pp. 61–63; January 15, 1959.) Discussion of recordings made at Thule, Greenland, and Barrow and College, Alaska, of cosmic-noise absorption at 27.6 mc at the time of a class-3 solar flare on July 29, 1958.

**523.164.3** 1522  
**Radio Observations of the Planet Jupiter**—K. L. Franklin and B. F. Burke. (*J. Geophys.*

Res., vol. 63, pp. 807-824; December, 1958.) Details are given of observations outlined previously (*Astrophys. J.*, vol. 61, p. 177, May, 1956; and *Carnegie Inst. of Wash. Year Book*, no. 55, pp. 74-76, 1955/1956) together with unpublished information from earlier observations in June, 1954. Correlation of the observations with rotation, polarization, and frequency characteristics is discussed. Most of the radio noise bursts were strongly circularly polarized, and the spectrum does not appear to be continuous over the frequencies observed. Critical reflections by an ionosphere around Jupiter are discounted.

**523.165:523.74** 1523  
**Cosmic-Ray Intensity Variations during Two Solar Cycles**—S. E. Forbush. (*J. Geophys. Res.*, vol. 63, pp. 651-669; December, 1958.) From ionization-chamber data the decrease of cosmic-ray intensity from its maxima (near sunspot minima) is shown to lag a year or more behind the increase of solar activity following sunspot minima. The variability of daily mean values was particularly great in 1957.

**523.165:550.38** 1524  
**Cosmic Rays in the Earth's Magnetic Field**—P. Rothwell. (*Phil. Mag.*, vol. 3, pp. 961-970; September, 1958.) Discrepancies between center dipole predictions using the Störmer equation and experimental observations of cosmic-ray intensities and cutoff phenomena are attributed to differences between the earth's real field and the dipole approximation to it, rather than to distortion of the earth's outer magnetic field by ionized interplanetary matter. An empirical expression for the cutoff phenomena is deduced which is in good agreement with experimental results over a wide range of latitude and longitude.

**523.3:621.396.9** 1525  
**Comparison of Measured and Computed Values of the Rapid Fading Rate of Ultra-High-Frequency Signals Reflected from the Moon**—S. J. Fricker, R. P. Ingalls, W. C. Mason and M. L. Stone. (*Nature (London)*, vol. 182, pp. 1438-1439; November 22, 1958.) Observations made in U.S.A. in August, 1957, during a bistatic moon reflection experiment carried out at 412.85 mc are reported. They show good agreement with values calculated from the effective libration rate of the moon.

**523.5:551.510.535** 1526  
**Meteors in the Ionosphere**—L. A. Manning and V. R. Eshleman. (PROC. IRE, vol. 47, pp. 186-199; February, 1959.) The nature of radio reflections from the trails of ionization produced by meteors is reviewed and the principal use of meteors as research tools are described. Some further directions in which meteor research may be pursued are suggested.

**523.5:621.396.11:551.510.535** 1527  
**Concerning Booker's Theory of Meteoric Reflections**—L. A. Manning and V. R. Eshleman. (*J. Geophys. Res.*, vol. 63, pp. 737-739; December, 1958.) The assumption of small-scale ionospheric turbulence [see e.g. 1417 of 1957 (Booker and Cohen)] conflicts with many known facts about meteor echoes, such as the persistence of aspect sensitivity for as long as ten seconds and the known variation with time of the angular correlation function.

**523.53:621.396.11** 1528  
**Electromagnetic Scattering by Low-Density Meteor Trails**—Brysk. (See 1665.)

**523.746** 1529  
**The Zurich Sunspot Number and its Variations for 1700-1957**—E. J. Chernosky and M. P. Hagan. (*J. Geophys. Res.*, vol. 63, pp. 775-788; December, 1958.) Monthly and auroral mean values of relative sunspot num-

ber are tabulated for 1749-1957. Estimates of annual values by Wolf for the period 1700-1759 are also given, together with derived data.

**523.75:523.72** 1530  
**Radiation and Particle Precipitation upon the Earth from Solar Flares**—L. Biermann and R. Lüst. (PROC. IRE, vol. 47, pp. 209-210; February, 1959.) A brief survey in which the energies of these emissions are estimated and their effects on the ionosphere discussed.

**550.380.2** 1531  
**Centenary of Melbourne-Toolangi Magnetic Observatory**—J. C. Dooley. (*J. Geophys. Res.*, vol. 63, pp. 731-735; December, 1958.) Brief history of the observatory since 1958.

**550.385** 1532  
**Possible Causes of Geomagnetic Fluctuations Having a Six-Second Period**—F. B. Daniels. (*Nature (London)*, vol. 182, p. 1439; November 22, 1958.) The fluctuations described by Duffus *et al.* (3441 of 1958) may be due to small oscillations of the earth's crust ("microseisms") or, more probably, atmospheric oscillations with the same period ("microbaroms").

**550.385:551.510.535** 1533  
**The Diurnal Variation of Irregular Geomagnetic Fluctuations: Part 2**—S. B. Nicholson and O. R. Wulf. (*J. Geophys. Res.*, vol. 63, pp. 803-806; December, 1958.) "Using the eight daily K numbers (three-hour-range indexes) for six observatories in moderately low latitudes and fairly well distributed in longitude, a universal-time component of the daily variation of irregular geomagnetic fluctuations having an apparently significant amplitude in the yearly average has been found in the data for the nine years 1940-1948. It is suggested that a portion of magnetic disturbance may be produced by dynamo action in the ionosphere, and give rise to this yearly-average universal-time component." For Part 1, see 1403 of 1956.

**550.385.2:525.624** 1534  
**Lunar Geomagnetic Tides at Kodaikanal**—K. S. Raja Rao and K. R. Sivaraman. (*J. Geophys. Res.*, vol. 63, pp. 727-730; December, 1958.) Tschu's method (636 of 1950) has been applied to the calculation of the lunar semi-diurnal variation of the horizontal intensity of the earth's magnetic field at Kodaikanal. The solar diurnal variation is also determined.

**550.385.4** 1535  
**Time Constants in the Geomagnetic Storm Effect**—C. O. Hines and L. R. O. Storey. (*J. Geophys. Res.*, vol. 63, pp. 671-682; December, 1958.) An estimate is made of the delay between formation of a ring current round the earth and the occurrence of a geomagnetic storm at ground level. Parker (1422 of 1957) has estimated the delay as several months; but by taking proper account of the fluidity of the medium and the presence of the main geomagnetic field the delay is found to be small enough to be consonant with the ring-current theory of geomagnetic storms.

**550.385.4** 1536  
**Inadequacy of Ring-Current Theory for the Main Phase of a Geomagnetic Storm**—E. N. Parker. (*J. Geophys. Res.*, vol. 63, pp. 683-689; December, 1958.) There appears to be insufficient dissipation to allow significant diffusion, during a magnetic storm, of the magnetic fields in the vicinity of the earth. A ring-current field can therefore only increase the horizontal component of the geomagnetic field and cannot account for the main phase of a geomagnetic storm.

**550.385.4** 1537  
**Studies on Geomagnetic Storm in Relation to Geomagnetic Pulsation**—Y. Kato and T.

Watanabe. (*J. Geophys. Res.*, vol. 63, pp. 741-756; December, 1958.) For the 27-day recurring geomagnetic storm (M type), the "cone of avoidance" model of Pecker and Roberts (1970 of 1955) is supported by both geomagnetic and cosmic-ray data. For the sporadically occurring storm (S type) a shock front in interplanetary space may be responsible for the sudden commencement. An exhaustive study should be made to determine whether the sudden commencement is only associated with the S type. The storm-time variation of the geomagnetic pulsation should also be measured during world-wide storms to aid the study of the shock wave.

**550.389.2** 1538  
**The International Geophysical Year**—L. V. Berkner. (PROC. IRE, vol. 47, pp. 133-136; February, 1959.) The organization which made possible the large international co-operative research effort is described and arrangements for collecting and storing the observational data are outlined.

**550.389.2** 1539  
**The Day-to-Day Coordination of I.G.Y. Observations**—A. H. Shapley. (PROC. IRE, vol. 47, pp. 323-327; February, 1959.) In addition to regular observational programs, the IGY included special series of observations made during outstanding solar and geophysical events. The organization of these series and the necessary ancillary communication facilities are described briefly.

**550.389.2** 1540  
**The Earth and its Environment**—S. Chapman. (PROC. IRE, vol. 47, pp. 137-141; February, 1959.) A general outline of the many fields of investigation included in the IGY program of observations. Particular attention is given to the earth's atmosphere and the flow of heat through it from the hot interplanetary gas.

**550.389.2:629.19** 1541  
**Motion of a Satellite around an Unsymmetrical Central Body**—R. R. Newton. (*J. Appl. Phys.*, vol. 30, pp. 115-117; January, 1959.) If a central body does not have an inversion center, the eccentricity of a satellite orbit varies periodically, with low frequency and large amplitude. Details are worked out for a central body composed of a large spherical mass and a small point mass contained somewhere within it.

**530.389.2:629.19** 1542  
**Change of Inclination of a Satellite Orbit**—C. H. Bosanquet. (*Nature (London)*, vol. 182, p. 1533; November 29, 1958.) Comment on 792 of March (Merson and King-Hele).

**550.389:2:629.19** 1543  
**Irregularities of Satellite Drag and Diurnal Variations in Density of the Air**—G. V. Groves. (*Nature (London)*, vol. 182, pp. 1533-1534; November 29, 1958.) The theory is applied to satellites 1957 $\beta$ , 1958 $\delta$ 1 and 1958 $\delta$ 2.

**550.389.2:629.19** 1544  
**Moon Rocket Tracking**—(*Nature (London)*, vol. 182, p. 1413; November 22, 1958.) A 108-mc interferometer situated at Lasham, Hampshire, was used to track the third American moon rocket on November 8, 1958. A satisfactory record was obtained from 0747 to 0801 GMT.

**550.389.2:629.19** 1545  
**Progress of the Russian Earth Satellite Sputnik III (1958 $\delta$ )**—D. G. King-Hele. (*Nature (London)*, vol. 182, pp. 1409-1411; November 22, 1958.) Orbital data, based chiefly on optical observations, are given for the instrument-carrying vehicle 1958  $\delta$ 2 and its rocket 1958  $\delta$ 1.

- 550.389.2:629.19 1546  
**The "Explorers"**—W. von Braun. (*Engineer (London)*, vol. 206, pp. 372–375; September 5, 1958.) Extracts from a paper presented at the congress of the International Astronautical Federation, Amsterdam, 1958. Details are given of the rocket firing procedure and satellite instrumentation as well as the tracking system used during the launching of the Explorer satellites.
- 550.389.2:629.19 1547  
**Instrumenting the Explorer I Satellite**—H. L. Richter, Jr., W. Pilkington, J. P. Eyraud, W. S. Shipley and L. W. Randolph. (*Electronics*, vol. 32, pp. 39–43; February 6, 1959.) A description is given of the 10-mw PM and 60-mw AM transistorized transmitters operating on 108 and 108.3 mc respectively. Total power consumption was 5 mw obtained from four mercury-type cells. Information was transmitted on skin temperature, nose cone temperature, meteorite counts and impacts, and cosmic rays. Details are also given of the test program for the completed apparatus.
- 550.389.2:629.19:551.510.535 1548  
**Earth-Satellite Observations of the Ionosphere**—W. W. Berning. (*Proc. IRE*, vol. 47, pp. 280–288; February, 1959.) The various methods of deriving integrated electron content below a satellite are critically reviewed. Although the methods are theoretically capable of giving information of the diurnal, seasonal and latitude variations of content, interpretation of the measurements is difficult because all these variations are present simultaneously; hence some assumptions about one or two of the variables must be made in order that the third one may be studied.
- 550.389.2:629.19:551.510.535 1549  
**Investigation of the Ionosphere Using Signals from Earth Satellites**—E. Woyk (Chvojková). (*Nature (London)*, vol. 182, pp. 1362–1363; November 15, 1958.) Possible effects of the ionosphere on radio signals from artificial earth satellites are discussed with reference to theoretical results noted earlier (3085 of 1958).
- 550.389.2:629.19:551.510.535 1550  
**Exploration of the Upper Atmosphere with the Help of the Third Soviet Sputnik**—V. I. Krassovsky. (*Proc. IRE*, vol. 47, pp. 289–296; February, 1959.) The instruments carried are described and some of the preliminary results of observations are given. Up to 250 km the principal ionospheric constituent is nitric oxide; above this height ions of atomic oxygen and nitrogen become more important. Electron density falls off slowly above the *F* region maximum. Positive-ion density at a height of 250 km is about  $5 \times 10^{10}$  ions/cm<sup>3</sup>, i.e. about half the corresponding electron density. The number of collisions with micrometeorites at heights of 150–300 km is 50/m<sup>2</sup>/sec.
- 551.51 1551  
**The Constitution and Composition of the Upper Atmosphere**—M. Nicolet. (*Proc. IRE*, vol. 47, pp. 142–147; February, 1959.) A survey of the observational data and the theoretical problems involved. Dissociation of molecular O<sub>2</sub> and N<sub>2</sub>, diffusion in the thermosphere, and the conduction of heat through the atmosphere are discussed.
- 551.510.535 1552  
**The Distribution of Electrons in the Ionosphere**—J. O. Thomas. (*Proc. IRE*, vol. 47, pp. 162–175; February, 1959.) The methods of deriving *N*(*h*) profiles are reviewed with particular emphasis on those which include the effects of the earth's magnetic field. Attention is drawn to the use of the extraordinary ray for estimating the electron density below the
- peak of the *E* layer. Some results are given of typical diurnal, seasonal, solar-cycle and latitude variations in the electron-density distribution. 92 references.
- 551.510.535 1553  
**The Normal D Region of the Ionosphere**—J. J. Gibbons and A. H. Waynick. (*Proc. IRE*, vol. 47, pp. 160–161; February, 1959.) A discussion of low-frequency sounding of the lower ionosphere indicating the radio transmission characteristics to be satisfied. It is suggested that recent measurements of parameters, such as reaction rates, will provide an improved model of the *D* layer.
- 551.510.535 1554  
**The Normal E Region of the Ionosphere**—E. V. Appleton. (*Proc. IRE*, vol. 47, pp. 155–159; February, 1959.) An outline of the structure and behavior of the *E* layer with particular emphasis on how it differs from an ideal Chapman layer.
- 551.510.535 1555  
**The Normal F Region of the Ionosphere**—D. F. Martyn. (*Proc. IRE*, vol. 47, pp. 147–155; February, 1959.) A review which includes comments on the physical conditions of the layer as determined from rocket and satellite observations, Bradbury's hypothesis, "spread-*F*" and radio-star scintillations.
- 551.510.535 1556  
**Motions in the Ionosphere**—C. O. Hines. (*Proc. IRE*, vol. 47, pp. 176–186; February, 1959.) The theoretical factors involved in the interpretation of observational data are discussed. Particular attention is given to tidal motions and the need for theoretical advances in studies of drifts, scintillations, auroral forms, turbulence and large-scale travelling disturbances.
- 551.510.535 1557  
**Horizontal Drift Measurements in the Ionosphere near the Equator**—N. J. Skinner, J. Hope and R. W. Wright. (*Nature (London)*, vol. 182, pp. 1363–1365; November 15, 1958.) Drift of the *E* and *F* layers has been measured regularly at Ibadan, Nigeria, by the spaced-receiver method. The westward daytime drift velocity in the *F* layer is about 105 m/s.
- 551.510.535 1558  
**The Interpretation of Night-Time Low-Frequency Ionograms**—J. M. Watts. (*J. Geophys. Res.*, vol. 63, pp. 717–726; December, 1958.) Two types of retardation of the extraordinary ray on passing through a region whose critical frequency is below the gyro-frequency are discussed. One is observed near the critical frequency of the ordinary ray and the other near the gyro-frequency. Information about the total number and distribution of electrons below the *F* layer can be deduced from them.
- 551.510.535 1559  
**Rocket Observations of the Ionosphere**—H. Friedmann. (*Proc. IRE*, vol. 47, pp. 272–280; February, 1959.) A review of the electron-density profiles obtained at Fort Churchill and at White Sands. A comparison of similar data obtained in Russia up to 200 km is made. The effects of polar blackouts and of flares are discussed with particular reference to the nature of the solar ionizing radiation. Comments are also made on the radiation which can be detected at night, interplanetary hydrogen and the electron density in interplanetary space.
- 551.510.535 1560  
**Summer-Day Auroral-Zone Atmospheric-Structure Measurements from 100 to 210 Kilometres**—R. Horowitz and H. E. LaGow. (*J. Geophys. Res.*, vol. 63, pp. 757–773; December, 1958.) Density and pressure determined during a rocket flight at 1600 CST, July 29, 1957, at Manitoba, Canada, are compared with earlier measurements.
- 551.510.535(98) 1561  
**Polar E<sub>s</sub>**—R. Penndorf and S. C. Coroniti—(*J. Geophys. Res.*, vol. 63, pp. 789–802; December, 1958.) Published ionosonde data are analyzed over the period 1954–1957 and show three distinct types of E<sub>s</sub> for north polar stations, each type occurring in a particular geomagnetic zone. A strong correlation between E<sub>s</sub> and magnetic activity existing along the auroral belt points to the influence of corpuscular radiation.
- 551.510.535:523.165 1562  
**Abnormal Ionization in the Lower Ionosphere Associated with Cosmic-Ray Flux Enhancements**—D. K. Bailey. (*Proc. IRE*, vol. 47, pp. 255–266; February, 1959.) Two separate abnormalities in ionization were recognizable during the great solar event of February 23, 1956: the "early effects" observed at the time of the sudden cosmic-ray enhancement, and the "late effects" which reached a maximum an appreciable time later. A difference in composition between streams of solar particles of cosmic-ray energies and ordinary cosmic rays is postulated to explain the former effect, while the latter is explained in terms of ionization at heights of 30–110 km produced by the passage or stopping of solar particles, predominantly protons. The absence of auroral and magnetic effects is consistent with these explanations. A provisional evaluation is also made of the coefficient of collisional attachment of electrons from negative ions and the negative-ion/electron ratio and effective recombination coefficient at night between 30 and 40 km.
- 551.510.535:550.385.4 1563  
**The F Region during Magnetic Storms**—K. I. Maeda and T. Sata. (*Proc. IRE*, vol. 47, pp. 232–239; February, 1959.) Magnetic storms are investigated using the variation of *f*<sub>o</sub>*F*<sub>2</sub> and *h*<sub>p</sub>*F*<sub>2</sub> over a wide range of latitude. It is shown that ionization drift theory in association with the dynamo theory is satisfactory in explaining the observed variations.
- 551.510.535:621.3.087.4 1564  
**Automatic Sweep-Frequency Ionosphere Recorder, Model C-4**—J. N. Brown. (*Proc. IRE*, vol. 47, pp. 296–300; February, 1959.) A description of an improved equipment for use during the IGY. In addition to the "virtual-height" frequency records, time-lapse motion pictures are taken.
- 551.510.535:621.3.087.4 1565  
**The I.G.Y. Three-Frequency Back-Scatter Sounder**—A. M. Peterson, R. D. Egan and D. S. Pratt. (*Proc. IRE*, vol. 47, pp. 300–314; February, 1959.) The equipment consists of three transmitters operating at 12, 18 and 30 mc, each coupled to its own horizontal Yagi antenna on a single rotating pole, and by a TR switch to a receiver whose tuning is interlocked with the transmitter. A camera records the plan-position displays for the three frequencies together with a short-range display. Preliminary results from 13 stations in polar, temperate and equatorial regions are discussed. Phenomena reported include sporadic-*E* movements, tilted-*F*-layer propagation, large irregularities in the *F* layer and echoes from field-aligned ionization.
- 551.510.535:621.087.4:523.164 1566  
**The Riometer—A Device for the Continuous Measurement of Ionospheric Absorption**—C. G. Little and H. Leimbach. (*Proc. IRE*, vol. 47, pp. 315–320; February, 1959.) A description of an instrument for the measurement of

ionospheric absorption at high latitudes during the IGY, using the cosmic-noise method (see 1152 of 1958). The circuit details of this self-balancing equipment are given and its advantages over the system of total-power measurement are discussed.

551.593.9 1567

**The Night Airglow**—F. E. Roach. (Proc. IRE, vol. 47, pp. 267-271; February, 1959.) "A phenomenological description of the night airglow is presented, reviewing the historical background, and what is known about height, temporal and spatial variations in intensity, and movements. The very important relationship to aurora and evidence for latitude-seasonal effects are examined."

551.594.5 1568

**Auroral Phenomena**—E. N. Parker. (Proc. IRE, vol. 47, pp. 239-244; February, 1959.) Present knowledge of the formation of the aurora is summarized. Mechanisms of particle acceleration are considered, and it is emphasized that there is no consistent theory of auroral phenomena. 37 references.

551.594.5 1569

**A Man-Made or Artificial Aurora**—A. L. Cullington. (*Nature (London)*, vol. 182, pp. 1365-1366; November 15, 1958.) An auroral display with which were associated radio fade-outs and a magnetic storm was observed at Apia (13°48'S., 171°46'W) on August 1, 1958. These phenomena were probably due to the explosion of a nuclear bomb over Johnston Island on the same date.

551.594.5:550.385 1570

**Auroral Ionization and Magnetic Disturbances**—B. Nichols. (Proc. IRE, vol. 47, pp. 245-254; February, 1959.) The literature on radio studies of the aurora, the relation between aurora and magnetic storms, and the movements of the ionosphere and auroral forms is coordinated (56 references). An average density of  $5 \times 10^6$  electrons/cm<sup>3</sup> is sufficient to explain normal radar echoes. Magnetic disturbances are closely related to ionization and luminosity of the aurora. Magnetic variations are also associated with increases in the speeds of motion of the ionization.

551.594.5:621.396.11 1571

**Radio Echoes from Auroral Ionization Detected at Relatively Low Geomagnetic Latitudes**—Leadabrand and Peterson. (See 1670.)

551.594.5:621.396.11 1572

**The Geometry of Auroral Communications**—Leadabrand and Vabroff. (See 1671.)

551.594.5:621.396.96 1573

**U.H.F. Auroral Radar Observations**—B. C. Blevis. (*J. Geophys. Res.*, vol. 63, pp. 867-868; December, 1958.) Describes auroral Doppler spectra obtained at Ottawa with bistatic radars operating at 488 and 944 mc 10-kw CW klystron amplifiers with 28- and 60-foot parabolic reflectors were used, with transmitter-receiver separations of 19 and 100 km. Echoes indicating cross sections of several thousand square meters were obtained on occasions at 944 mc.

551.594.5:621.396.96 1574

**A Low-Power V.H.F. Radar for Auroral Research**—R. S. Leonard. (Proc. IRE, vol. 47, pp. 320-322; February, 1959.)

551.594.6 1575

**The Very-Low-Frequency Emissions Generated in the Earth's Exosphere**—R. M. Gallet. (Proc. IRE, vol. 47, pp. 211-231; February, 1959.) Types of noise other than whistlers have been examined using high resolution spectrograms which show the frequency dis-

tribution as a function of time. Most of these noises are excited in the exosphere by streams or bunches of high-speed ionized particles trapped in the earth's magnetic field. The excitation mechanism is similar to that of a traveling-wave valve. Most of the observations require that particle velocities be about 10,000 km. A model in which the ratio of electron density to magnetic field strength is almost constant along a line of force in the exosphere seems to be indicated by several types of noise.

551.594.6 1576

**Atmospheric Whistlers**—R. A. Helliwell and M. G. Morgan. (Proc. IRE, vol. 47, pp. 200-208; February, 1959.) The historical background of the discovery of whistlers and the development of an adequate theoretical exploration of their behavior are reviewed. The program of study during the IGY is described and some examples of results obtained are presented and discussed.

551.594.6 1577

**Path Combinations in Whistler Echoes**—M. G. Morgan, H. W. Curtis and W. C. Johnson. (Proc. IRE, vol. 47, pp. 328-329; February, 1959.) A note describing groups of whistler echoes recorded at Unalaska (50° geomagnetic latitude) and explaining them in terms of a combination of alternative propagation paths and multiple trips over each path.

551.594.6 1578

**A Note on Whistler Propagation in Regions of Very Low Electron Density**—O. K. Garriott. (*J. Geophys. Res.*, vol. 63, pp. 862-865; December, 1958.) Calculations made on a model ionosphere indicate that nose whistlers may exist, although attenuation may prevent their observation.

551.594.6:523.75 1579

**A 27-kc/s Sudden Enhancement of Atmospherics Anomaly**—W. A. Feibelman. (*J. Geophys. Res.*, vol. 63, p. 866; December, 1958.) On May 1, 1958, two "anomalous" solar flares caused a sharp drop in the 27-kc signal level recorded at Pittsburgh, Pa., instead of the usual enhancement. Enhancements were recorded elsewhere.

551.594.6:621.396.821 1580

**Study of the Atmospheric Radio Noise at 27 and 100 kc/s at Delhi**—D. K. Sachdev. (*J. Sci. Indust. Res.*, vol. 17A, pp. 262-270; July, 1958.) Preliminary survey of observations extending over a year, with particular reference to sudden enhancement effects.

#### LOCATION AND AIDS TO NAVIGATION

621.396.93+523.164 1581

**Propagation of Electromagnetic Waves, Radio Location and Radio Astronomy**—E. Roessler. (*Electrotech. Z., Edn A*, vol. 79, pp. 737-740; October 1, 1958.) Review of recent developments. 49 references.

621.396.93:621.396.677.83 1582

**The Application of Metallic Reflectors for Purposes of Location**—G. Megla. (*Hochfreq. und Elekroak.* vol. 66, pp. 106-115; January, 1958.) The use of stationary and rotating plane mirrors for radio beacons and other navigational aids is described. For English version see IRE 1957 WESCON CONVENTION RECORD, vol. 1, Part 10, pp. 29-40.

621.396.934 1583

**Radio Interferometers Track Airborne Vehicles**—M. W. Miles. (*Electronic Ind.*, vol. 17, pp. 94-95, 151; October 1958.) Basic problems in interferometer systems are examined.

621.396.96 1584

**Miniature X-Band Radar Has High Resolution**—C. D. Hardin and J. Salerno. (*Electronics*, vol. 32, pp. 48-51; January 20, 1959.) The unit operates with a pulse repetition frequency of 100 kc and peak power 150 w. The IF is 400 mc with a bandwidth of 100 mc and gain 36 db in four stages. A low-range altimeter with a full-scale reading of 30 feet and calibrated at 6-inch intervals has been constructed.

621.396.96:551.594.5 1585

**A Low-Power V.H.F. Radar for Auroral Research**—R. S. Leonard. (Proc. IRE, vol. 47, pp. 320-322; February, 1959.)

621.396.96.083.7 1586

**Radar Data Transmission**—T. E. Schilizzi. (*Proc. IRE (Australia)*, vol. 19, pp. 467-480; September, 1958.) Practical methods of achieving a remote display over wide-band and narrow-band circuits are reviewed and an experimental system using wide-band radio links is outlined. Image-converter and bright-image storage tubes for large-screen projection equipment are described. See also 2770 of 1957 (Dixon and Thomas).

621.396.962.012:681.142 1587

**An Analogue Computer for Evaluating Radar Performance**—Willhite and McIntyre. (See 1465.)

621.396.963 1588

**Dynamic Compression for Radar Receivers**—D. Levine (*Electronic Ind.*, vol. 17, pp. 102-106 and 82-84; October and November, 1958.) Various types of receiver are examined in terms of crt brightness for direct observation and for recording on film. Graphs are given to assist in the selection of the required dynamic compression curve for the system in which the receiver is to be used.

621.396.967 1589

**The Rotterdam Harbour Radar System**—B. H. G. Prins and J. M. G. Seppen. (*Philips Telecommun. Rev.*, vol. 20, pp. 16-30; September, 1958.) A chain of seven radar stations working in the 3-cm band provides accurate positional information to ships with the aid of leading lines and cursor lines which are superimposed electronically on display screens.

621.396.967:621.372.413:621.317.799 1590

**A Radar Echo Box with Remote Control**—G. van Gelder and E. Scholten. (*Philips Telecommun. Rev.*, vol. 20, pp. 33-38; September, 1958.) A description of the construction of a high-Q resonant cavity for the 8.5- or 10-cm band and its tuning mechanism. It can be used for the measurement of the over-all performance of a radar installation and for obtaining the power spectrum of a magnetron pulse.

621.396.967.029.65 1591

**An 8-mm, High-Definition Radar Set**—J. M. G. Seppen and J. Verstraten. (*Philips Telecommun. Rev.*, vol. 20, pp. 5-15; September 1958.) The design of the equipment is discussed and general circuit features and photographs of the radar display are given.

621.396.967.2 1592

**Development of the DEW Line**—G. R. Frantz. (*Bell. Lab. Rec.*, vol. 37, pp. 2-6; January, 1959.) A general outline of the radar warning system is given.

#### MATERIALS AND SUBSIDIARY TECHNIQUES

535.215:546.482.21 1593

**Photochemical Effects and the Effect of Oxygen on Photoconducting Cadmium Sulfide Crystals**—J. Woods. (*J. Electronics Control*, vol. 5, pp. 417-426; November, 1958.) The

photochemical effects and the effects of heat treatment in oxygen on a number of differently activated crystals are described and discussed.

**535.215:546.482.21** 1594  
**Growth of Photoconductivity of Single Crystals of Cadmium Sulphide from Irradiation in an Atomic Pile**—M. Martineau. (*Compt. rend. Acad. Sci. Paris*, vol. 247, pp. 639–643; August 11, 1958.)

**535.37** 1595  
**Luminescence and Luminescent Materials**—D. G. Anderson. (*J. Electronics Control*, vol. 5, pp. 457–470; November, 1958.) "A general review of luminescent materials is given, with some emphasis on the more recent applications. The theoretical aspects of the absorption and emission processes are discussed briefly."

**535.37:546.472.21** 1596  
**Electron Traps in Zinc-Sulphide Phosphors**—W. Hoogenstraaten. (*Philips Res. Rep.*, vol. 13, pp. 515–693; December, 1958.) An attempt is made to clarify the chemical and physical nature of electron traps in ZnS phosphors, including those based on mixed crystals of ZnS with CdS or ZnSe.

**535.376** 1597  
**Electroluminescence and Image Intensification**—G. Diemer, H. A. Klasens and P. Zalm. (*Philips Tech. Rev.*, vol. 19, pp. 1–11; July 27, 1957.) The mechanism of electroluminescence and the application of the effect in solid-state image intensifiers is discussed.

**535.376:539.234** 1598  
**Electroluminescent Thin Films**—W. A. Thornton. (*J. Appl. Phys.*, vol. 31, pp. 123–124; January, 1959.) Thin luminescent films of ZnS-Cu, Cl and ZnS-Cu,Mn,Cl have been prepared by a two-step evaporation/firing process similar to that described by Feldman and O'Hara (3494 of 1957) but using finished electroluminescent powder phosphors as the starting material.

**537.226/.227:546.431.824-31** 1599  
**On Growing BaTiO<sub>3</sub> Single Crystals**—A. M. Cherepanov. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2280–2284; October, 1957.) Triangular crystals with a hypotenuse up to 15 mm long are obtained after heating 1 BaCl<sub>2</sub>·0.53 BaCO<sub>3</sub>·0.26 TiO<sub>2</sub> to a temperature of 1300–1450°C. These crystals have lattice characteristics  $a = 3.986 \text{ \AA}$ ,  $c = 4.026 \text{ \AA}$ , and a Curie point at 122°–123°C.

**537.226:546.212-16** 1600  
**Dielectric Properties of Ice Crystals**—(*Helv. Phys. Acta*, vol. 30, pp. 553–610; December 30, 1957. In German, with English summaries.)

Part 1—Dynamic Theories of Dielectric Constants—A. Steinemann and H. Grätznicher (pp. 553–580.)

Part 2—Dielectric Investigations on Ice Crystals with Occluded Impurity Atoms—A. Steinemann (pp. 581–610.)

**537.226:546.431.824-31** 1601  
**Dielectric Properties of Barium Titanate at High Frequencies and Low Temperatures**—H. Rabenhorst and J. Melcherčik. (*Ann. Phys. (Lpz.)*, vol. 1, pp. 261–263; May 20, 1958.) Dielectric constant and loss factor have been measured in the temperature range  $-180^\circ$  to  $+20^\circ\text{C}$  at a frequency of 9.1 kmc.

**537.226:621.319.2** 1602  
**Contribution to the Phenomenological Theory of Electron**—A. N. Cubkin. (*Zh. Tekh. Fiz.*, vol. 27, pp. 1954–1968; September, 1957.) Five dielectrics, carnauba wax, naphthalene,

steatite, nylon and magnesium titanate, were prepared under fields ranging from 5 to 25 kv per cm and observed for as long as three months. An expression is derived for the lifetime of these electrets and experimental and calculated results are tabulated.

**537.226:621.319.2** 1603  
**The Anomalous Stability of New Inorganic Polycrystalline Electrets**—A. N. Gubkin and G. I. Skanavi. (*Zh. Tekh. Fiz.*, vol. 27, pp. 1969–1970; September, 1957.) Experiments show that the properties of CaTiO<sub>3</sub> electrets of specific conductivity  $10^{-12}$ – $10^{-14} \Omega^{-1} \text{ cm}^{-1}$  are similar when kept open-circuited or short-circuited by metallic foils. See also 3164 of 1957.

**537.311.33** 1604  
**Contribution to the Theory of Semiconductors with Excited Impurity Zone**—M. I. Klinger and Yu. I. Zozulya. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2285–2290; October, 1957.) The variations of conductivity, Hall effect and thermo-EMF as functions of temperature in  $n$ -type semiconductors are discussed.

**537.311.33:546.23** 1605  
**Self-Diffusion in Selenium**—B. I. Boltaks and B. T. Plachenor. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2229–2231; October, 1957.) Results of an investigation of the diffusion of impurities in crystalline and amorphous Se are given. The difference between the diffusion constants for the two types increases with increasing temperature.

**537.311.33:546.23:536.2** 1606  
**The Influence of Bromine Additives on the Thermal Conductivity of Selenium**—G. B. Abdullaev and A. A. Bashshaliev. (*Zh. Tekh. Fiz.*, vol. 27, pp. 1971–1975; September, 1957.) Amorphous and crystalline Se samples with Br content varying from 0.008 to 0.5 per cent were tested at 27.5°C; their thermal conductivity coefficients were respectively  $3.08 \times 10^{-3}$  and  $7.02 \times 10^{-3} \text{ cal/deg. cm sec}$ .

**537.311.33:546.24:538.214** 1607  
**Magnetic Susceptibility in Semiconductors of the Tellurium Type**—Yu. A. Firsov. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2212–2228; October, 1957.) A mathematical analysis of zone junctions. A method is proposed for the investigation of the zone structure based on the angular dependence of susceptibility in strong magnetic fields.

**537.311.33:546.28** 1608  
**Lattice Absorption Bands in Silicon**—F. A. Johnson. (*Proc. Phys. Soc.*, vol. 73, pp. 265–272; February 1, 1959.) Detailed measurements are described of the absorption spectrum of pure vacuum-grown Si crystals, for wavelengths  $2\mu$ – $30\mu$  and at temperatures of 20, 77, 290 and 365°K. Some new absorption bands were found. The positions and temperature dependence of all the principal absorption peaks are explained in terms of multiple phonon interactions.

**537.311.33:546.28** 1609  
**The Effect of Heat Treatment on the Breakdown Characteristics of Silicon  $p$ - $n$  Junctions**—A. R. Plummer. (*J. Electronics Control*, vol. 5, pp. 405–416; November, 1958.) Soft breakdown, a more gradual breakdown than the normal avalanche characteristic, can be due to unsatisfactory bulk properties, which can be induced by heat treatment. An explanation for the effect is suggested.

**537.311.33:546.281.26:537.533** 1610  
**Electron Emission from Breakdown Regions in SiC  $p$ - $n$  Junctions**—L. Patrick and W. J. Choyke. (*Phys. Rev. Lett.*, vol. 2, pp. 48–50; January 15, 1959.) The maximum emission was found to range from  $10^{12}\text{A}$  to  $10^8\text{A}$ , the

emission depending strongly on sample preparation.

**537.311.33:546.289** 1611  
**A Note on the Theory of Diffusion of Copper in Germanium**—M. D. Sturge. (*Proc. Phys. Soc.*, vol. 73, pp. 297–306; February 1, 1959.)

**537.311.33:546.289** 1612  
**The Diffusion of Boron in Germanium**—M. D. Sturge. (*Proc. Phys. Soc.*, vol. 73, pp. 320–322; February 1, 1959.)

**537.311.33:546.289** 1613  
**Electrical Properties of Germanium Doped with Zinc**—S. G. Kalashnikov, E. Yu. L'vova and V. V. Ostrobodova. (*Zh. Tekh. Fiz.*, vol. 27, pp. 1925–1930; September, 1957.) Zn has little effect on recombination velocity and appears to be suitable for doping Ge to give a low specific resistance and long electron lifetime.

**537.311.33:546.289** 1614  
**The Influence of Elements of Groups III and V on the Recombination Velocity of Electrons and Holes in Germanium**—V. G. Alekseeva, S. G. Kalashnikov, L. P. Kalnach, I. V. Karpova and A. I. Morozov. (*Zh. Tekh. Fiz.*, vol. 27, pp. 1931–1939; September, 1957.) Shorter electron and hole lifetimes were obtained with Bi and Tl than with Sb and Ga. Elements possessing a small distribution coefficient such as Bi, Tl, Cu, Ni and Fe strongly accelerated the electron-hole recombination but elements such as Sb and Ga with a larger coefficient proved much less active. Co, Mn and Au seemed also to accelerate this recombination in Ge.

**537.311.33:546.289** 1615  
**Valley-Orbit Splitting of Arsenic Donor State in Germanium**—G. Weinreich, W. S. Boyle, H. G. White and K. F. Rodgers. (*Phys. Rev. Lett.*, vol. 2, pp. 96–98; February 1, 1959.) Investigation of the far-infrared excitation spectrum of bound electrons in an elastically strained sample verifies that of the two states into which the donor ground state is split, the triplet is the lower one.

**537.311.33:546.289** 1616  
**Edge Breakdown of  $p$ - $n$  Junctions in Germanium**—B. M. Vul and A. P. Shotov. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2189–2194; October, 1957.) Investigation shows that the permittivity of the medium influences appreciably the breakdown voltage of  $n^+p$  junctions but affects only slightly the  $p^+n$  and diffusion junctions. This can be explained by assuming a negative charge on the Ge surface which has the effect of altering the thickness of the junction near the surface.

**537.311.33:546.289** 1617  
**Ambipolar Transport of Carrier Density Fluctuations in Germanium**—J. E. Hill and K. M. van Vliet. (*Physica*, vol. 24, pp. 709–720; September, 1958.) A theoretical study is made of the spectrum of generation-recombination noise at relatively high field strengths. Experimental results for near-intrinsic crystals confirm that if the ambipolar drift velocity is sufficiently large the noise spectrum is quite different from that at low field strengths.

**537.311.33:546.289** 1618  
**The Diffusion Constant, Mobility and Lifetime of Minority Carriers in Germanium Containing Parallel Arrays of Dislocations**—J. B. Arthur, A. F. Gibson, J. W. Granville and E. G. S. Paige. (*Phil. Mag.*, vol. 3, pp. 940–949; September, 1958.) Measurements of the diffusion constant  $D$  show that for  $n$ -type Ge  $D$  is anisotropic; for  $p$ -type Ge  $D$  is isotropic, and the carrier lifetime anisotropic. Measurements of drift mobility of holes in  $n$ -type Ge at high

electric fields show this to be anisotropic with respect to the dislocation array, no comparable effect occurring in  $p$ -type material. A qualitative interpretation of the data is given. See also 494 of 1958 (Bell and Hogarth).

537.311.33:546.289 1619

**An Interpretation of Certain Transport Properties in Germanium Containing Parallel Arrays of Edge Dislocations**—A. F. Gibson and E. G. S. Paige. (*Phil. Mag.*, vol. 3, pp. 950–960; September, 1958.) Using a model similar to that considered by Read (457 of 1955), a quantitative interpretation is given of the anisotropic effects observed by Arthur *et al.* (1618 above). The diameter of the space-charge cylinder surrounding the dislocations ( $1.6 \times 10^{-4}$  cm) and the fraction of time (about one half) that an injected hole spends within the space-charge region are deduced from the analysis.

537.311.33:546.289:537.32 1620

**Thermoconductivity of Germanium**—E. D. Devyatkovka and I. A. Smirnov. (*Zh. Tekh. Fiz.*, vol. 27, pp. 1944–1949; September, 1957.) The thermoconductivity of eight  $p$ -type and  $n$ -type samples was examined in the temperature range  $80^{\circ}$ – $300^{\circ}$ K. Below  $200^{\circ}$ K the thermoconductivity for similar types of Ge depended on the current-carrier concentration. The  $p$ -type Ge showed a greater thermoconductivity than the  $n$ -type.

537.311.33:546.289:538.632 1621

**The Temperature Dependence of the Hall Coefficient in Semiconductors with Constant Carrier Concentration**—E. I. Kaplunova and K. B. Tolpygo. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2246–2251; October, 1957.) Mathematical treatment with reference to Ge. Results are shown graphically.

537.311.33:546.482.21 1622

**Space-Change Limited Currents in Insulating Materials**—G. T. Wright. (*Nature*, (London), vol. 182, pp. 1296–1297; November 8, 1958.) The steady  $I/V$  characteristic of a plate of CdS about  $3 \times 10^{-3}$  cm thick provided with a cathode of In metal and an anode of colloidal graphite is shown. In the reverse direction the crystal was, as expected, an insulator.

537.311.33:546.561-31 1623

**Diamagnetic Zeeman Effect and the Exciton Structure in Cuprous Oxide Crystals**—E. F. Gross and B. P. Zakharchenya. (*Zh. Tekh. Fiz.*, vol. 27, pp. 1940–1943; September, 1957.) See also 1499 of 1957.

537.311.33:546.682.19 1624

**Preparation of Indium Arsenide**—R. H. Harada and A. J. Strauss. (*J. Appl. Phys.*, vol. 30, p. 121; January, 1959.)

537.311.33:546.682.86 1625

**Electrical Conduction in  $n$ -Type InSb between  $20^{\circ}$ K and  $300^{\circ}$ K**—E. H. Putley. (*Proc. Phys. Soc.*, vol. 73, pp. 280–290; February 1, 1959.) Measurements are reported of the electrical conductivity and Hall coefficient for samples of InSb containing from  $5 \times 10^{13}$  to  $10^{18}$  conduction electrons per  $\text{cm}^3$ . The intrinsic carrier concentration is calculated from the results obtained above  $180^{\circ}$ K. The Hall mobility results are explained as due to a combination of acoustic lattice and ionized-impurity scattering.

537.311.33:621.314.7 1626

**Multiplication of Minority-Carrier Current in the Nonideal  $p$ - $n$  Junction**—V. I. Stafecv. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2195–2211; October, 1957.) The possibility of increasing minority-carrier current by means of a non-ideal  $p$ - $n$  junction is examined. Alloy transistors

of  $p$ - $n$ - $p$  and  $n$ - $p$ - $n$  type with  $\alpha > 1$  and a region of negative resistance in the collector circuit are considered. The disruptive voltage can be adjusted from a few volts to 150–250 v and current after breakdown may reach 150–300 ma. The time of passage from one state to another is 0.1–0.2  $\mu$ s. See also 1893 of 1956 (Schenkel and Statz).

538.22:538.569.4 1627

**Paramagnetic Spectra of Substituted Sapphires: Part 1—Ruby**—E. O. Schulz-DuBois. (*Bell Sys. Tech. J.*, vol. 38, pp. 271–290; January, 1959.) Resonance properties of  $\text{Cr}^{+++}$  ions in  $\text{Al}_2\text{O}_3$  have been investigated theoretically and experimentally for three-level solid-state maser application. Energy levels and associated eigenvectors are presented as functions of applied magnetic field, and transition probabilities at certain orientations are discussed. Paramagnetic spectra for signal frequencies between 5 and 24 mc are shown.

538.22:538.569.4 1628

**Paramagnetic Resonance Spectrum of  $\text{Cr}^{+++}$  in Emerald**—J. E. Gusic, M. Peter and E. O. Schulz-DuBois. (*Bell Sys. Tech. J.*, vol. 38, pp. 291–296; January, 1959.) Spectra observed in the  $X$ ,  $K$ , and  $M$  bands gave the spectroscopic splitting factors for this material. The large value of zero-field splitting observed suggests an application as a solid-state maser for higher microwave frequencies.

538.22:538.569.4 1629

**Paramagnetic Resonance in the 10000-Mc/s Band of Europium and Gadolinium subjected to a Cubic Crystalline Field**—C. Rytter. (*Helv. Phys. Acta*, vol. 30, pp. 353–373; October 15, 1957. In French.)

538.22:538.569.4 1630

**Theory of Paramagnetic Resonance of Europium and Gadolinium subjected to a Cubic Crystalline Field**—R. Lacroix. (*Helv. Phys. Acta*, vol. 30, pp. 374–394; October 15, 1957. In French.)

538.221 1631

**Some Magnetic Properties of Dilute Ferromagnetic Alloys: Part 2**—B. W. Lohian, A. C. Robinson and W. Sucksmith. (*Phil. Mag.*, vol. 3, pp. 999–1012; September, 1958.) Part 1: 3667 of 1955 (Bate *et al.*).

538.221:621.318.134:538.614 1632

**Faraday Effect and Birefringence in Ferrites at Microwave Frequencies**—F. Mayer. (*Ann. Télécommun.*, vol. 12, pp. 279–288, 305–332, and 334–342; July–October, 1957.)

538.221:621.318.134:548.0 1633

**Some Properties of Mixed Gadolinium-Erbium and Gadolinium-Yttrium**—G. Villers, J. Loriers and C. Claudel. (*Compt. rend. Acad. Sci., Paris*, vol. 247, pp. 710–713; August 25, 1958.)

538.221:621.318.134:621.318.57 1634

**Observations of Rotational Switching in Ferrites**—W. L. Shevel, Jr. (*IBM J. Res. Developm.*, vol. 3, pp. 93–95; January, 1959.) Experimental data on switching times as a function of applied field for square-loop ferrites, suggest three mechanisms of flux reversal, each being dominant over a certain region of the curve of inverse switching time against applied field.

621.315.61:621.318.4 1635

**Special-Purpose Magnet-Wire Insulation**—G. Sideris. (*Electronics*, vol. 32, pp. 60–61; February 13, 1959.) A brief review of wires available in U.S.

## MATHEMATICS

512:621.316.5 1636

**Ternary Switching Algebra**—E. Mühlendorf. (*Arch. elekt. Übertragung*, vol. 12, pp. 138–148; March, 1958.) A switching algebra is developed for use in systems with a ternary code.

517.949.8:681.142 1637

**The  $Z$  Transformation**—H. A. Helm. (*Bell Sys. Tech. J.*, vol. 38, pp. 177–196; January, 1959.) "The Stieltjes integral is used to develop a rigorous derivation of the  $z$  transform. Sufficient properties of the transformation are included to form a reasonably complete basis for the operational solution of constant-coefficient, linear, finite-difference equations."

519.251.7:53.088 1638

**Analysis of Errors in the Determination of the Mean Value of a Random Quantity and of its Mean Square Error due to the Finite Time of Observation**—A. E. Kharybin. (*Ayatomalika i Telemekhanika*, vol. 18, pp. 304–314; April, 1957.) Formulas are derived and nomograms are plotted by means of which it is possible a) to estimate whether the interval for the observation of a random quantity is sufficient, and b) to find the assumed mean-square errors in the determination of the mean value and its dispersion for a typical case.

519.2 1639

**The Advanced Theory of Statistics. Vol. 1: Distribution Theory.** [Book Review]—M. G. Kendall and A. Stuart. Publishers: Griffin, London, Eng., 1958, 433 pp., 84s. (*Nature (London)*, vol. 182, p. 1470; November 29, 1958.)

## MEASUREMENTS AND TEST GEAR

621.3.018.41(083.74) 1640

**Standard-Frequency Transmissions**—(*Electronic Radio Eng.*, vol. 36, pp. 117–118; March, 1959.) Details are given of a new method of adjusting the MSF standard-frequency transmission with reference to the second of Ephemeris Time (ET), to be operated from March, 1959. The MSF frequency corrections published monthly in *Electronic Radio Eng.* will be given to  $\pm 1$  part in  $10^{10}$ , and the form of presentation will be changed.

621.3.018.41(083.74):538.569.4 1641

**Theory of the Cavity Microwave Spectrometer and Molecular Frequency Standard**—Y. Beers. (*Rev. Sci. Instr.*, vol. 30, pp. 9–16; January, 1959.) The theory predicts that a considerable improvement in signal/noise ratio and sensitivity can be obtained in the cavity spectrometer. A fractional stability of about  $4 \times 10^{-12}$  can be achieved in a frequency standard using the 3.3 ammonia line.

621.317.3:621.372.412 1642

**Quartz Crystals Require Testing for Spurious Response**—A. N. Silverstein. (*Electronic Ind.*, vol. 17, pp. 85–88; October, 1958.) The test equipment described is a modified form of the low-frequency spectrum analyzer reported by McDuffie (202 of 1956). It is used for detecting interfering spurious responses in quartz crystals operating in the frequency range 20–60 mc.

621.317.361+621.317.373 1643

**A Pulse-Beat Method for Accurate Frequency and Phase Measurements and Phase-Locked Frequency Changing**—G. Becker. (*Frequenz*, vol. 12, pp. 82–90; March, 1958.) A method is described for accurate short-term comparison of two standard frequencies, particularly if their relations are rational, even if they differ considerably in magnitude. A differentiation method for accurate determination of

phase coincidence, and the principle of phase-locking a frequency changer are outlined.

**621.317.42:538.569.4** 1644  
**Measurement of Magnetic Fields by Nuclear Resonance**—G. C. Lowe. (*Electronic Eng.*, vol. 31, pp. 138-140; March, 1959.) A simple feedback circuit is described for the accurate measurement of magnetic fields over a wide range with little readjustment of circuit elements.

**621.317.6:621.3.001.4** 1645  
**Sorting Components by Measuring Waveforms**—B. Agusta. (*Electronics*, vol. 32, pp. 56-59; February 13, 1959.) Waveform characteristics of circuit components are obtained by sampling at discrete time intervals. A pulse whose amplitude is proportional to that of the waveform at the sampling time is produced and a digital output reading is obtained.

**621.317.7:621.314.7** 1646  
**Measurement of Junction-Transistor Equivalent-Circuit Parameters**—J. J. Sparkes. (*A.T.E.J.*, vol. 14, pp. 176-187; July, 1958.)

**621.317.7:621.314.7** 1647  
**A Transistor D.C.-A.C. Beta Tester**—T. P. Sylvan. (*Electronic Ind.*, vol. 17, pp. 90-92; October, 1958.) Details are given of a tester for measuring both the dc and ac common-emitter current gain and the collector-to-emitter leakage current.

**621.317.733:[621.314.63+621.314.7]** 1648  
**Capacitance Bridges for Semiconductor Measurements**—N. F. Blackburne. (*A.T.E.J.*, vol. 14, pp. 166-175; July, 1958.) Two bridges are described. The first is an admittance bridge for measurement of the effective parallel resistance and capacitance of *p-n* junctions under forward, reverse and zero bias in the frequency range 1 kc-1 mc. The second bridge measures collector capacitance and extrinsic base resistance of junction transistors at frequencies up to 10 mc. Measurements can be made at elevated temperatures.

**621.317.734** 1649  
**On the Accuracy of the Ohmmeter**—R. Braae. (*Engineer (London)*, vol. 206, pp. 563-565; October 10, 1958.) The classical theory of errors is used to analyze the capabilities and limitations of measuring equipment, particularly of ohmmeters.

**621.317.75** 1650  
**The Recording and Collocation of Waveforms: Part 1**—R. J. D. Reeves. (*Electronic Eng.*, vol. 31, pp. 130-137; March, 1959.) The basic principles and history of stroboscopic oscillographs are given. Modern methods of strobe pulse generation, synchronization and time delay are discussed.

**621.317.77** 1651  
**Measurement of Phase and Amplitude at Low Frequencies**—R. J. A. Paul and M. H. McFadden. (*Electronic Eng.*, vol. 31, pp. 142-149; March, 1959.) Description of a 1f oscillator and phasemeter for testing feedback control systems. The range extends from below 0.01 cs to at least 10 kc.

**621.317.794:537.324** 1652  
**Thermocouples and Bolometers for Radiation Measurements in the Infrared Spectral Region**—G. Grave and W. Heimann. (*Elektronik*, vol. 7, pp. 65-69; March, 1958.) The construction of sensitive radiation measuring equipment is described with brief details of apparatus for medical applications.

**621.317.794:537.324** 1653  
**Noise in Radiation Thermocouples**—V. Schley and F. Hoffmann. (*Optik*, vol. 15, pp.

358-371; June, 1958.) Measurements indicate that solder joints to the couple may be a source of thermal noise.

#### OTHER APPLICATIONS OF RADIO AND ELECTRONICS

**535.376.07** 1654  
**Solid-State Panels for Display or Storage**—R. K. Jurgen. (*Electronics*, vol. 32, pp. 46-47; January 30, 1959.) The use of photorectifier matrices with packing densities of 256 per inch<sup>2</sup> for character generation and storage is described.

**621-52:621.395.625.3** 1655  
**Magnetic Drum Provides Analogue Time Delay**—H. L. Daniels and D. K. Sampson. (*Electronics*, vol. 32, pp. 44-47; February 6, 1959.) The use of a magnetic drum to provide delays of between 5 and 20 s is described. A recording range of  $\pm 50$  v and an accuracy within 0.1 per cent for frequencies between 0 and 1 cs are achieved.

**621-52:681.142** 1656  
**Digital System Positions Shafts over Phone Line**—R. B. Palmiter. (*Electronics*, vol. 32, pp. 62-66; February 13, 1959.) Description of a carrier system by which data describing the position of three master shafts as well as all necessary synchronizing and auxiliary information are transmitted over a telephone line for controlling slave shaft positions.

**621.384.6:621.319.339** 1657  
**An Accelerator with an 800-kV Cascade Generator**—K. Simonyi. (*Acta Tech. Acad. Sci. Hung.*, vol. 19, pp. 353-362; 1958. In German.) Description of an installation of the Central Physical Research Institute, Budapest, with details of the rectifier-valve heating and ion focusing arrangements.

**621.396.969:533.6.011.7** 1658  
**Radio Observations of Hypersonic Shock Waves**—J. S. Hey, J. T. Pinson and P. G. Smith. (*Nature (London)*, vol. 182, pp. 1220-1221; November 1, 1958.) Application of the Doppler technique described earlier (3625 of 1957).

**621.398:621.376.5** 1659  
**A Static Transmitting Device for Pulse-Frequency Telemetry Systems**—A. M. Pshenichnikov. (*Ayatomatika i Telemekhanika*, vol. 18, pp. 444-448; May, 1957.) Description of a magnetic-modulator/multivibrator circuit with compensating feedback network. The frequency of the multivibrator output is a linear function of the measured quantity. Advantages over existing types of transmitting device are noted.

#### PROPAGATION OF WAVES

**621.396.11** 1660  
**On the Spectrum of a Passive Scalar Mixed by Turbulence**—A. D. Wheelon. (*J. Geophys. Res.*, vol. 63, pp. 849-850; December, 1958.) A discussion of 1527 and 1533 of 1958 (Bolgiano). Previous analysis (2881 of 1957) is extended to show that removal of spectral content by convective transfer is more important than diffusion damping throughout the inertial range, whatever the spectrum may be. See also 1661 below and *ibid.*, pp. 854-855.

**621.396.11** 1661  
**On the Role of Convective Transfer in Turbulent Mixing**—R. Bolgiano, Jr. (*J. Geophys. Res.*, vol. 63, pp. 851-853; December, 1958.) The extended theory given by Wheelon (1660 above) is not necessarily valid, since additional mechanisms of transfer (rather than dissipation) are introduced. The differences between the two theories are due to conceptu-

ally distinct methods of treating the convective transfer process.

**621.396.11** 1662  
**Propagation of Electromagnetic Pulses Around the Earth**—B. R. Levy and J. B. Keller. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 56-61; January, 1958. Abstract, *PROC. IRE*, vol. 46, p. 799; April, 1958.)

**621.396.11** 1663  
**Scattering of Electromagnetic Waves in Beyond-the-Horizon Radio Transmission**—D. I. Paul. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 61-65; January, 1958. Abstract, *PROC. IRE*, vol. 46, p. 799; April, 1958.)

**621.396.11** 1664  
**Downcoming Radio Waves**—J. R. Wait. (*Electronic Radio Eng.*, vol. 36, pp. 106-107; March, 1959.) A method of measuring the angle of arrival, azimuth and polarization of a downcoming radio wave using a crossed-loop direction finder with a four-element Adcock antenna.

**621.396.11:523.53** 1665  
**Electromagnetic Scattering by Low-Density Meteor Trails**—H. Brysk. (*J. Geophys. Res.*, vol. 63, pp. 693-716; December, 1958.) A uniform underdense line of electrons is considered and the model is then extended to include diffusion and to take account of the fast initial movement of the particles before they are slowed to thermal velocities. The effect of antenna gain is discussed. Various limiting cases are examined and expressions are given for specular and nonspecular scattering.

**621.396.11:551.510.52** 1666  
**Wavelength Dependence in Transhorizon Propagation**—R. Bolgiano, Jr. (*PROC. IRE*, vol. 47, pp. 331-332; February, 1959.) An experiment to determine the wavelength dependence of the scattering coefficient in tropospheric scatter propagation is described. Scaled antennas at 417 mc and 2290 mc were used, and a wide range found for the wavelength dependence. The variation is considered to be statistically significant and to correspond to a variation in the meteorological processes responsible for the scattering.

**621.396.11:551.510.535** 1667  
**On the Reciprocity of H.F. Ionospheric Transmission**—M. Balsler, W. B. Smith and E. Warren. (*J. Geophys. Res.*, vol. 63, pp. 859-861; December, 1958.) Describes preliminary simultaneous two-way tests with pulses over a 1685-km path in Canada, using the same antenna at each terminal for both receiving and transmitting. Sample records show nonreciprocal effects in the fading, even though antennas insensitive to Faraday rotation were used.

**621.396.11:551.510.535** 1668  
**Polarization Fading over an Oblique-Incidence Path**—D. A. Hedlund and L. C. Edwards. (*IRE TRANS. ON ANTENNAS AND PROPAGATION*, vol. AP-6, pp. 21-25; January, 1958. Abstract, *PROC. IRE*, vol. 46, p. 799; April, 1958.)

**621.396.11:551.510.535:537.56** 1669  
**Equivalence Theorems of Wave Absorption in Plasma**—K. Rawer and K. Suchy. (*Ann. Phys., Lpz.*, vol. 1, pp. 255-260; May 20, 1958.) Using wave and ray theory and neglecting the earth's magnetic field it is shown that Martyn's theorem of ionospheric absorption is valid only if attenuation is mainly due to collisions and is proportional to the number of collisions.

**621.396.11:551.594.5** 1670  
**Radio Echoes from Auroral Ionization Detected at Relatively Low Geomagnetic Latitudes**

tudes—R. L. Leadabrand and A. M. Peterson. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 65-79; January, 1958. Abstract, PROC. IRE, vol. 46, p. 799; April, 1958.)

621.396.11:551.594.5 1671  
**The Geometry of Auroral Communications**  
 —R. L. Leadabrand and I. Yabroff. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 80-87; January, 1958. Abstract, PROC. IRE, vol. 46, pp. 799-800; April, 1958.)

621.396.11.029.45:621.396.67:621.315.1 1672  
**Power-Line Aerial**—R. M. Golden, R. V. Langmuir, R. S. Macmillan and W. V. T. Rusch. (*Electronic Radio Eng.*, vol. 36, pp. 116; March, 1959.) A note on propagation experiments carried out in California using an 8-mile section of medium-voltage single-phase power transmission line as a radiator at 8.4 kc. Parallel-resonant circuits tuned to the operating frequency were used to isolate the line, and the power supplied to the aerial terminals was 10 kw. Measurements of extremely strong ionospheric reflections have been made at near-vertical incidence, and satisfactory signals have been detected 250 km away.

621.396.11.029.64 1673  
**Radio Attenuation at 11 kMc and Some Implications affecting Relay System Engineering**—S. D. Hathaway and H. W. Evans. (*Bell Sys. Tech. J.*, vol. 38, pp. 73-97; January, 1959.) Measurements have been made of attenuation and rainfall over paths of length 27 and 12 miles. The results are compared with theoretical expectations of attenuation, and the effect of attenuation on usable path length is discussed.

#### RECEPTION

621.376.23:621.396.822 1674  
**The Response of Nonlinear Devices to Band-Limited High-Frequency Signals and Noise**—N. W. W. Smith. (*J. Electronics Control*, vol. 5, pp. 385-401; November, 1958.) "The envelope detection model is used to describe the response of nonlinear devices e.g. detectors and mixers) to band-limited high-frequency signals. Results obtained for discrete sinusoidal signals are shown to be applicable to the continuous spectra of high-frequency noise without recourse to probability theory."

621.396.3 1675  
**A Modern Frequency-Shift Telegraph Receiver**—E. J. Allen. (*Electronic Eng.*, vol. 31, pp. 161-164; March, 1959.) In this system the "mark" and "space" signals are separated. Since diversity reception is employed, two mark and two "space" signals are available, each pair of which are combined in a ratio squarer. The resultants are converted to double-current signals and added. The system will function with only one of the four signals present.

621.396.62+621.397.62]:658.5 1676  
**Some Uses of Statistical Methods in the Manufacture of Radio and Television Receivers**—A. I. Godfrey. (*J. Brit. IRE*, vol. 19, pp. 15-28; January, 1959.)

621.396.621 1677  
**Receivers with Zero Intermediate Frequency**—J. C. Greene and J. F. Lyons. (PROC. IRE, vol. 47, pp. 335-336; February, 1959.) Measurements made on a frequency-sweep receiver covering the range 500-950 mc at a sweep rate of 10 cps and using a low-pass video amplifier in place of a normal IF stage, indicated an average noise figure of 8 db.

621.396.662 1678  
**Some Aspects of Permeability Tuning**—W. D. Meevzen. (*J. Brit. IRE*, vol. 19, pp.

47-60; January, 1959.) Reprint, See 3622 of 1958.

621.396.662:534.76 1679  
**F.M. Tuner Adapter for Multiplexed Stereo**  
 —L. Feldman. (*Electronics*, vol. 32, pp. 52-54; February 6, 1959.) Description of a 4-valve converter unit for the reception of a multiplex-type two-channel stereophonic transmission.

621.396.821:551.495.6 1680  
**Study of the Atmospheric Radio Noise at 27 and 100 kc/s at Delhi**—Sachdev. (See 1580.)

621.396.822 1681  
**The Spectrum of Limited Gaussian Noise**—F. L. H. M. Stumpers. (*Philips Res. Rep.*, vol. 13, pp. 509-514; December, 1958.) "The energy spectrum of the output of a nonlinear device with an input of Gaussian noise can be expressed directly in terms of the Laplace transform of its characteristic and of convolutions of the input bandpass form. The expression is derived and the result applied to several forms of limiter characteristic. The same method can be applied if a carrier signal is also present."

621.396.822 1682  
**Some Aspects of Noise in Communications and Servo Systems**—R. M. Huey. (*Proc. IRE (Australia)*, vol. 19, pp. 486-493; September, 1958.) Methods of system design involving studies of the system response to random signals, as opposed to sinusoidal signals, are discussed.

#### STATIONS AND COMMUNICATION SYSTEMS

621.391 1683  
**On the Mathematical Theory of Error-Correcting Codes**—H. S. Shapiro and D. L. Slotnick. (*IBM J. Res. Developm.*, vol. 3, pp. 25-34; January, 1959.)

621.396.3:621.376 1684  
**An Experimental Modulation-Demodulation Scheme for High-Speed Data Transmission**—E. Hopner. (*IBM J. Res. Developm.*, vol. 3, pp. 74-84; January, 1959.) A description is given of a low-cost system which was designed to determine speed and reliability limitations on transmitting binary data over private telephone lines.

621.396.324 1685  
**An Electronic Coder and Decoder for Teletypewriter Signals**—J. Das. (*Electronic Eng.*, vol. 31, pp. 156-160; March, 1959.)

621.396.41:621.396.62 1686  
**The Development of Radio Telegraphy Systems on Short Waves and its Influence on Receiver Techniques**—R. Tastenoy. (*Rev. IIF Brussels*, vol. 4, pp. 11-23; 1958.) A review of the principal multiplex systems and of the problems of selectivity and stability in receivers associated with these systems. Diversity reception is briefly discussed.

621.396.5 1687  
**Radio Link Installations for the Transmission of Telephony and Television in the 4-kMc/s and 2-kMc/s Bands**—K. Christ, O. Laaff and K. Schmid. (*Elektrotech. Z., Edn A.*, vol. 79, pp. 687-693; October 1, 1958.) The equipment of two FM systems is described. The 4-kmc system has a capacity of 600 telephony channels or one 625-line television channel, and the 2-kmc installation has 120 carrier-frequency channels; both conform to CCITT and CCIR recommendations.

621.396.5.029.62 1688  
**Radio Transmission into Buildings at 35 and 150 Mc/s**—L. P. Rice. (*Bell. Sys. Tech. J.*, vol. 38, pp. 197-210; January, 1959.) The em field

on the ground floor of a building was found to be 20-25 db less than the median field in city streets at the same distance from the transmitter. It is shown that the effective coverage range in buildings is greater for the higher frequency.

621.396.65:621.396.933 1689  
**Radio Links for the Control of Aeronautical Air-Ground-Air Equipment**—W. S. McGuire. (*Proc. IRE (Australia)*, vol. 19, pp. 541-550; October, 1958.) A multichannel FM system between Melbourne and outlying stations for the control of aircraft communication equipment is described. Frequency bands used are 160 mc and 450 mc; equipment is duplicated, and automatic change-over switching and test facilities are provided.

621.396.712 1690  
**Common-Channel Common-Programme Operation of Medium-Wave Broadcasting Stations**—S. F. Brownless. (*Proc. IRE, Australia*, vol. 19, pp. 529-541; October, 1958.) Two experimental systems are described, one using a low-power booster transmitter, the other using two manually synchronized crystal-controlled high-power transmitters. Conclusions are drawn from the test results reported. 30 references.

621.396.721:621.395.623.8 1691  
**Sound Distribution at the Brussels Exhibition**—A. V. J. Martin. (*Audio Eng.*, vol. 43, pp. 26-28, 80; February, 1959.) Five 250-w FM transmitters were used to supply about 450 sound sources, each consisting of one receiver feeding four loudspeakers. The receivers distributed over the exhibition grounds were in continuous operation for the six months' duration of the World Fair 1958. Details of circuitry and special installations are given.

621.396.74+621.397.7]:(434.7) 1692  
**The Network of Broadcast and Television Transmitters in Württemberg**—H. Rupp. (*Electrotech. Z., Edn A.*, vol. 79, pp. 670-673; October 1, 1958.)

#### SUBSIDIARY APPARATUS

621-526:621.314.7-555.621 1693  
**A Simple Temperature-Control System for Transistors**—H. Kemhadjian. (*Mullard Tech. Commun.*, vol. 4, pp. 186-190; December, 1958.) A description of a closed-loop servomechanism in which a power transistor, which supplies the current for a heating coil, is controlled by an error signal produced by a temperature-sensing AF transistor.

621-526:621.396.6 1694  
**Loop controls Scatter Power to Offset Fading**—L. P. Yeh. (*Electronics*, vol. 32, pp. 60-62; January 30, 1959.) A closed-loop servo system is described whereby the transmitted power of an UHF tropospheric scatter link is reduced under conditions of good reception and increased under fading conditions. A statistical analysis of the system is given.

621.3.087.6 1695  
**Converting Recorders to Rectilinear Outputs**—N. D. Diamantides. (*Electronic Ind.*, vol. 17, pp. 82-84; October, 1958.) A computer circuit generating variable transport time delays is described for transforming pen recordings from arcs to straight lines. See also 1914 of 1957 (Massa and Massa).

621.311.6:621.317.7.029.6 1696  
**Stabilized Variable-Frequency A.C. Instrument Calibration Source**—C. A. Master and W. L. Mandrell. (*Rev. Sci. Instr.*, vol. 30, pp. 38-40; January, 1959.) A network is described which renders a 20 cps-20 kc amplifier insensitive to variations of signal input, power supply and load.

- 621.311.62:621.316.722.1 1697  
**High-Stability Mains-Operated Valve Heater Supply**—C. T. Murray and R. E. Aitchison. (*Proc. IRE, Australia*, vol. 19, pp. 494-495; September, 1958.) Details are given of a circuit using a high-gain tuned amplifier with feedback controlled by a thermistor bridge.
- 621.314.57:621.314.7 1698  
**A Four-Transistor D.C. Converter Circuit for Use with Relatively High-Voltage Supplies**—W. L. Stephenson. (*Mullard Tech. Commun.*, vol. 4, pp. 191-192; December, 1958.) An experimental bridge circuit is described with a power output of 40 w and conversion efficiency of 90 per cent.
- TELEVISION AND PHOTOTELEGRAPHY**
- 621.397.5 1699  
**Television Techniques**—E. Schwartz. (*Elektrotech. Z., Edn A*, vol. 79, pp. 743-746; October 1, 1958.) Review of recent developments. 69 references.
- 621.397.5:535.623 1700  
**A Simple and Compatible System for Colour Television**—L. Chrétien and R. Aschen. (*TSF et TV*, vol. 33, pp. 323-324; October, 1957.) The system outlined is based on transmitting the luminance signal in a bandwidth of 5.5 mc and the chrominance signals on two sub-carriers of different frequency, all within the 14-mc channel normally allotted for 819-line monochrome transmissions.
- 621.397.5:535.623:535.417 1701  
**Applications of the Interference of Light in Thin Films**—P. M. van Alphen. (*Philips Tech. Rev.*, vol. 19, pp. 59-67; August 24, 1957.) The theory of optical interference and its application to dichroic mirrors for color television is briefly discussed. The construction and production of the mirrors is described.
- 621.397.5:621.395.625.3 1702  
**Eight Papers on Video Tape and Recording**—(*J. Soc. Mot. Pic. Telev. Eng.*, vol. 67, pp. 721-745; November, 1958.) Problems in recording and editing video signals in the Ampex system are described and discussed. The mechanical design of the tape transport system and head is not described. The electronic system for recording and playback and the processing of the recovered signal to restore the synchronization and gating pulses is described together with the use of magnetic powders to enable splices to be located accurately. The problem of the development of high-quality tape for use with rotating heads and the importance of head alignment are discussed.
- 621.397.611.2:535.215 1703  
**Semiconducting Materials in Vidicon-Type Television Pickup Tubes**—V. A. Babits. (*J. Telev. Soc.*, vol. 8, pp. 498-502; October-December, 1958.) A summary of the target materials used in the vidicon-type tube, and the physical principles underlying their performance. Possible increases in sensitivity and temperature range are suggested as a result of proper choice of material for the various layers.
- 621.397.621.2 1704  
**The Current Characteristic of TV Picture Tubes**—W. F. Niklas. (*J. Telev. Soc.*, vol. 8, pp. 512-515; October-December, 1958.) The value of the exponent in the cathode current characteristic of cathode ray tubes increases from 2.5 to 3.5 with increasing current values for space-charge-limited emission.
- 621.397.621.2:535.623 1705  
**The Faraday Cell in Colour Television**—R. W. Wells. (*J. Telev. Soc.*, vol. 8, pp. 503-506; October-December, 1958.) A summary of the advantages and limitations of a projection cathode ray tube with a Faraday cell and a cellophane filter.
- 621.397.7 1706  
**Isle of Wight I.T.A. Station**—(*Brit. Commun. Electronics*, vol. 5, pp. 850-851; November, 1958.) An outline is given of the transmitting equipment at Chillerton Down and the microwave radio-relay system linking London and Southampton. For a similar account see *Engineer (London)*, vol. 206, pp. 340-341; August 29, 1958.
- 621.397.7(4) 1707  
**European Television Stations**—(*Wireless World*, vol. 65, pp. 109-116; March, 1959.) A comprehensive survey of European networks including summaries of services available in each country and reproductions of test cards. World television standards and European channel allocations in bands I and III are tabulated. Transmitters, links, and conversion centers for Eurovision links are indicated on a map.
- TUBES AND THERMIONICS**
- 621.314.63:546.289 1708  
**Blocking Junction Process in Planar Germanium Diodes Type DG-Ts**—Yu. K. Barsukov. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2252-2261; October, 1957.) An investigation of the first stage of the blocking process, corresponding to limitation of the reverse current through the diode by the resistance of the external circuit. The results for five diodes are tabulated and presented graphically.
- 621.314.63.002 1709  
**The Manufacture of Germanium Diodes**—C. F. Hühn. (*Telefunken Ztg.*, vol. 31, pp. 11-20; April, 1958. English summary, p. 66.) A description of the complete manufacturing process.
- 621.314.632 1710  
**High-Frequency Gallium Arsenide Point-Contact Rectifiers**—W. M. Sharpless. (*Bell Sys. Tech. J.*, vol. 38, pp. 259-269; January, 1959.) Recent work is described using single crystals as frequency converters at frequencies as high as 60 kmc and as switching diodes for switching times of the order of  $10^{-10}$ s. Satisfactory operation is obtained over a considerable range of temperature.
- 621.314.632:621.316.722.1 1711  
**Applications for Zener Diodes**—G. Porter. (*Electronic Ind.*, vol. 17, pp. 108, 110; October, 1958.) Tables are given of the manufacturers of various types of Zener diode and of its applications.
- 621.314.7 1712  
**New Transistor Works at Cryogenic Temperatures**—S. Weber. (*Electronics*, vol. 32, pp. 34-35; January 23, 1959.) A "grain-boundary" transistor capable of operating at temperatures as low as 2°K is described. Its operation is based on the characteristics of the boundary formed between two crystal lattice structures of different grain orientation.
- 621.314.7 1713  
**Some Criteria for the Thermal Stability of Transistors**—F. Weitzsch. (*Frequenz*, vol. 12, pp. 65-71; March, 1958.) The conditions giving rise to thermal instability are analyzed and formulas are derived to determine whether a given circuit requires stabilization.
- 621.314.7 1714  
**An Analysis of Base Resistance for Alloy-Junction Transistors**—A. J. Wahl. (*IRE TRANS. ON ELECTRONIC DEVICES*, vol. ED-5, pp. 131-139; July, 1958. Abstract, *PROC. IRE*, vol. 46, p. 1889; November, 1958.)
- 621.314.7 1715  
**Physical Mechanisms Leading to Deterioration of Transistor Life**—G. C. Messenger. (*IRE TRANS. ON ELECTRONIC DEVICES*, vol. ED-5, pp. 147-151; July, 1958. Abstract, *PROC. IRE*, vol. 46, p. 1890; November, 1958.)
- 621.314.7:621.317.7 1716  
**Measurement of Junction-Transistor Equivalent-Circuit Parameters**—J. J. Sparkes. (*A.T.E. J.*, vol. 14, pp. 176-187; July, 1958.)
- 621.314.7:621.317.7 1717  
**A Transistor D.C.-A.C. Beta Tester**—Sylvan. (See 1647.)
- 621.314.7.012.8 1718  
**Transistor Equivalent Circuit**—D. A. Green. (*Electronic Radio Eng.*, vol. 36, pp. 108-114; March, 1959.) A circuit is derived which is valid at all frequencies at which alloyed-junction transistors give useful gain. The circuit may be used to calculate the performance of hf amplifiers.
- 621.314.7.012.8 1719  
**The Junction Transistor as a Network Element at Low Frequencies: Part I—Characteristics and  $h$  Parameters**—J. P. Beijersbergen, M. Beun and J. te Winkel. (*Philips Tech. Rev.*, vol. 19, pp. 15-27; July 7, 1957.)
- 621.314.7.012.8:530.17 1720  
**The Thermal Equivalent Circuit of a Transistor**—P. R. Strickland. (*IBM J. Res. Developm.*, vol. 3, pp. 35-45; January, 1959.) It is shown that an exact electrical analogue can be given for the thermal system between the collector junction and the constant-temperature environment of a transistor. An experimental method is given for obtaining the parameters of the equivalent circuit, and its application in circuit design is discussed.
- 621.383:546.289:535.61-1 1721  
**The Frequency Characteristic of a Germanium Infrared Diode Modulator**—Yu. I. Ukhanov. (*Zh. Tekh. Fiz.*, vol. 27, pp. 1950-1953; September, 1957.) Infrared rays passing through a Ge diode modulator were received on a PbS photocell. The modulation coefficient was found to be constant in the frequency range 20 cps-20 kc.
- 621.383.27 1722  
**Stroboscopic Operation of Photomultiplier Tubes**—C. F. Hendee and W. B. Brown. (*Philips Tech. Rev.*, vol. 19, pp. 50-58; August 24, 1957.) A technique is described for analyzing a periodically repeated low-intensity flash which is synchronized with a phase-delayed pulse to a photomultiplier tube. The tube can be triggered by a pulse as short as  $10^{-7}$  sec to act as a light shutter, and the tube gain is much higher than under dc operation. Circuit details and experimental results are given.
- 621.383.5 1723  
**Grain-Boundary Photovoltaic Cell**—R. K. Mueller and R. L. Jacobson. (*J. Appl. Phys.*, vol. 30, pp. 121-122; January, 1959.) A photocell is described which is sensitive to light-spot movement in two dimensions. It comprises a rectangular rod of n-type Ge containing a grain boundary perpendicular to its main axis. Ohmic contacts are applied at the ends of the sample and In contacts are alloyed on opposite sides covering the grain boundary.
- 621.385:621.374.32 1724  
**The Trochotron**—N. P. R. Sherry. (*Brit. Commun. Electronics*, vol. 5, pp. 842-843; November, 1958.) A description of the mode of operation of a hot-cathode high-vacuum valve with a pulse counting rate in the region

of  $10^6$  per second. See also 3079 of 1954 (Björkman and Lindberg).

- 621.385.029.6 1725  
**On the Application of Harmonic Vibrations of Electrons for the Generation of Ultra High Frequencies**—P. A. Borodovskii. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2353–2355; October, 1957.) A brief description of results obtained with an experimental valve in which variations of the accelerating voltage from 150 to 1000 v produce frequency changes from 310 to 600 mc. See 3398 of 1954 (Alfvén and Romell).
- 621.385.029.6 1726  
**Effect of Electron Lenses on Beam Noise**—R. C. Knechtli. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 84–88; April, 1958. Abstract, Proc. IRE, vol. 46, p. 1440; July, 1958.)
- 621.385.029.6 1727  
**New Mechanism of Noise Reduction in Electron Beams**—M. R. Currie and D. C. Forster. (*J. Appl. Phys.*, vol. 30, pp. 94–103; January, 1959.) Experiments have demonstrated that the basic noise quantities of an electron beam can be transformed and optimized by varying the field configuration and flow characteristics in the multivelocity region near the cathode. A detailed investigation of the conditions for minimum beam noisiness is presented.
- 621.385.029.6 1728  
**Reduction of Electron-Beam Noisiness by means of a Low-Potential Drift Region**—A. W. Shaw, A. E. Siegman, and D. A. Watkins. (Proc. IRE, vol. 47, pp. 334–335; February, 1959.)
- 621.385.029.6 1729  
**Ballistic Analysis of a Two-Cavity Finite-Beam Klystron**—S. E. Webber. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 98–108; April, 1958. Abstract, Proc. IRE, vol. 46, p. 1440; July, 1958.)
- 621.385.029.6 1730  
**Principle of Operation and Working Data for the Power Reflex Klystron TK7**—F. Möhring. (*Tech. Mitt. PTT*, vol. 36, pp. 145–149; April 1, 1958.) The klystron has a maximum hf output  $>5$  w for a 35-mc bandwidth in the 4-kmc range.
- 621.385.029.6 1731  
**Propagation in a Crossed-Field Periodic Structure**—A. Kiel, M. Scotto and P. Parzen. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 76–84; April, 1958. Abstract, Proc. IRE, vol. 46, p. 1440; July, 1958.)
- 621.385.029.6 1732  
**Characteristics of Travelling-Wave Tubes with Periodic Circuits**—R. W. Gould. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 186–195; July, 1958. Abstract, Proc. IRE, vol. 46, p. 1890; November, 1958.)
- 621.385.029.6 1733  
**A Hybrid-Type Travelling-Wave Tube for High-Power Pulsed Amplification**—E. J. Nalos. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 161–166; July, 1958. Abstract, Proc. IRE, vol. 46, p. 1890; November, 1958.)
- 621.385.029.6 1734  
**Conditions for Minimum Noise Generation in Backward-Wave Amplifiers**—M. R. Currie and D. C. Forster. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 88–98; April, 1958. Abstract, Proc. IRE, vol. 46, p. 1444; July, 1958.)
- 621.385.029.6:537.533 1735  
**Space-Charge-Wave Excitation in Solid-Cylindrical Brillouin Beams**—W. W. Rigrod and J. R. Pierce. (*Bell Sys. Tech. J.*, vol. 38, pp. 99–118; January, 1959.) The voltage and current modulation of ideal cylindrical electron beams in Brillouin flow and in beams of zero magnetic field are studied by means of Laplace transforms. The problems of field modulation by an annular gap in a drift tube and noise excitation of a beam are discussed.
- 621.385.029.6:537.533 1736  
**Space-Charge Wave Harmonics and Noise Propagation in Rotating Electron Beams**—W. W. Rigrod. (*Bell Sys. Tech. J.*, vol. 38, pp. 119–139; January, 1959.) The higher-order space-charge waves on solid-cylindrical electron beams are considered. The properties of such waves are derived from slow-wave small-signal analysis.
- 621.385.029.6:621.316.726.078 1737  
**Nonvacuum Devices control Klystrons**—M. C. Harp. (*Electronics*, vol. 32, pp. 68–70; February 13, 1959.) Close frequency control of transmitter and receiver local-oscillator klystrons in a 6-kc microwave link is achieved by the use of magnetic amplifiers and transistors.
- 621.385.029.6:621.372.2+621.372.8 1738  
**Electron Waves in Periodic Structures**—L. A. Vainshtein. (*Zh. Tekh. Fiz.*, vol. 27, pp. 2340–2352; October, 1957.) Relations derived earlier (338 of 1957) for retarding systems are generalized, and an expression is obtained for calculating the linear properties of travelling-wave and backward-wave valves.
- 621.385.029.6:621.376.3 1739  
**Linearization of the Frequency-Modulation Characteristic of a Reflex Klystron**—E. Schuon and H. J. Butterweck. (*Arch. elekt. Übertragung*, vol. 12, pp. 99–108; March, 1958.) The characteristic can be straightened and the frequency swing increased by applying a coupled-resonator load with band-filter characteristics.
- 621.385.029.64/.65 1740  
**Two Backward-Wave Oscillator Tubes for the 29,000 to 74,000 Megacycle Frequency Range**—D. J. Blattner and F. Sterzer. (*RCR Rev.*, vol. 19, pp. 584–597; December, 1958.) The characteristics and performance of two experimental backward-wave oscillator valves are described. Their wide frequency range and milliwatt-level power output make them suitable for signal-generator and local-oscillator applications.
- 621.385.029.64 1741  
**The Reflex Klystron as a Negative-Resistance-Type Amplifier**—C. F. Quate, R. Kompfner, and D. A. Chisholm. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 173–179; July, 1958. Abstract, Proc. IRE, vol. 46, p. 1890; November, 1958.)
- 621.385.029.64 1742  
**A 20- to 40-kMc/s Backward-Wave Oscillator**—R. W. Grew, D. A. Dunn, S. W. McLaughlin and R. P. Lagerstrom. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 152–156; July, 1958. Abstract, Proc. IRE, vol. 46, p. 1890; November, 1958.)
- 621.385.029.65/.66 1743  
**Dielectric Slow-Wave Structures for the Generation of Power at Millimetre and Submillimetre Wavelengths**—R. H. Patnell, P. D. Coleman, and R. C. Becker. (IRE TRANS. ON ELECTRON DEVICES, vol. ED-5, pp. 167–173; July, 1958. Abstract, Proc. IRE, vol. 46, p. 1890; November, 1958.)
- 621.385.852 1744  
**Cathode-Ray Storage Tubes for Direct Viewing**—A. S. Kramer. (*Electronics*, vol. 32, pp. 40–41; January 23, 1959.) Physical and electrical characteristics, performance data, and applications of 17 types of tube are tabulated.
- 621.385.832 1745  
**Survey of Half-Tone Image Storage Tubes**—H. G. Lubszynski. (*Nachrichtentech. Z.*, vol. 11, pp. 115–124; March, 1958.) See 3362 of 1957.
- 621.385.832 1746  
**Characteristics and Applications of the Iatron Storage Tube**—D. W. Davis. (*Commun. and Electronics*, pp. 47–53; March, 1957. *Elect. Commun.*, pp. 93–102; 1958.)
- 621.385.832 1747  
**Cathode-Ray Tube Adds Third Dimension**—E. L. Withey. (*Electronics*, vol. 31, pp. 81–83; May 23, 1958.) Details are given of the "peritron," an experimental crt in which the fluorescent screen, of 18 cm diameter, is displaced harmonically along the Z-axis using a drive motor and crank assembly. A sinusoidal signal relating to the instantaneous position of the screen and derived by means of a magnet mounted in the push-rod and a fixed coil, is used to provide beam gating and focusing correction. Applications in air-traffic control systems are discussed.
- 621.385.832.087.6 1748  
**An Electron-Beam Tube for 'Writing' Directly on Special Paper**—(*Elektronik*, vol. 7, p. 48; February, 1958.) The prototype tube described has numerous short lengths of wire incorporated in its glass screen with the bare wire ends protruding slightly from the screen to form a raster. The electron beam impinging on the inner end of the wires will leave a record of its trace on current-sensitive paper (Type L, teledeltos) placed in contact with the tube screen.

#### MISCELLANEOUS

- 061.4:621.3 1749  
**Physical Society's Exhibition [1959]**—(*Wireless World*, vol. 65, pp. 126–134; March, 1959.) Brief descriptions are given of selected exhibits.
- 621.3-71 1750  
**Prediction of Temperatures in Forced-Convection-Cooled Electronic Equipment**—L. Friedl. (IRE TRANS. ON COMPONENT PARTS, vol. CP-5, pp. 102–107. Abstract, Proc. IRE, vol. 46, p. 1553; August, 1958.)