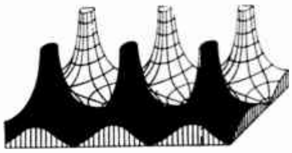
November, 1959
Vol 47 No. 11

Proceedings of the IRE

Poles and Zeros



Piled Higher and Deeper. In earlier days these words were often a facetious, and occasionally a derogatory reference to the Doctor of Philosophy degree. The value of graduate study has not always been understood in electrical engineering, and we remember the day when our major industries refused a salary differential for holders of graduate degrees, and when their personnel interviewers advised new graduates to get "practical experience" in preference to continuation of their studies.

Such treatment of the holder of the M.S. or the Ph.D. is not now found in electrical engineering, where graduate degrees are well understood as desirable items of academic preparation for tomorrow's leaders in applied science. It is unfortunate, however, that this view of advanced education is not more uniformly shared throughout the engineering field. We use the word unfortunate because we fail to see how preparation at the bachelor's level can continue to be sufficient to solve the problems of tomorrow involving new materials, new energy sources, new mathematics, and new systems useful to man.

Last year the engineering colleges of the U.S.A. granted 35,322 Bachelor's degrees, 5788 Master's degrees, and 647 Doctorates. Of the total, electrical engineering led the other fields with 27.1 per cent of the Bachelor's degrees, 27.4 per cent of the Master's degrees, and 22.3 per cent at the Doctorate level. Mechanical engineering granted 25.6 per cent at the first level, but only 16.7 per cent and 11.8 per cent at the Master's and Doctor's levels. Civil engineering was third at 14.5, 14.2, and 10.2 per cent respectively. During the academic year 10,779 students were studying for advanced degrees in electrical engineering. The field of mechanical engineering, with an almost equal number of B.S. graduates, was able to attract only 5492 to advanced study.

Is the failure to attract greater numbers to graduate study related to industry attitudes and pressures in some fields? Are some facets of industry still unbelieving that graduate study is such an important part of the battle for our civilization? Or that graduate study is a gain and only a deferral, and not a loss to the hiring process? Added to such questions concerning industry attitudes must be some consideration of faculty attitudes as well.

The promotion of graduate study is or should be a sales function of our college staffs, one of great importance to them because only through the graduate college can they hope to match their own teaching numbers to the pedagogic load of the next decade. Unfortunately, not all teachers are willing to accept the necessity of doing more than teaching, that by their own example and enthusiasm they will have tremendous influence on our next generation.

To electrical teachers the graduate years are usually looked upon as an opportunity to reiterate the fundamental truths, and to build upon them a broadened and deepened theoretical structure. Increased mathematical rigor is always a part of such programs—in fact, some schools with more limited mathematical resources are hard put to fulfill the mathematical demands of electrical graduate programs.

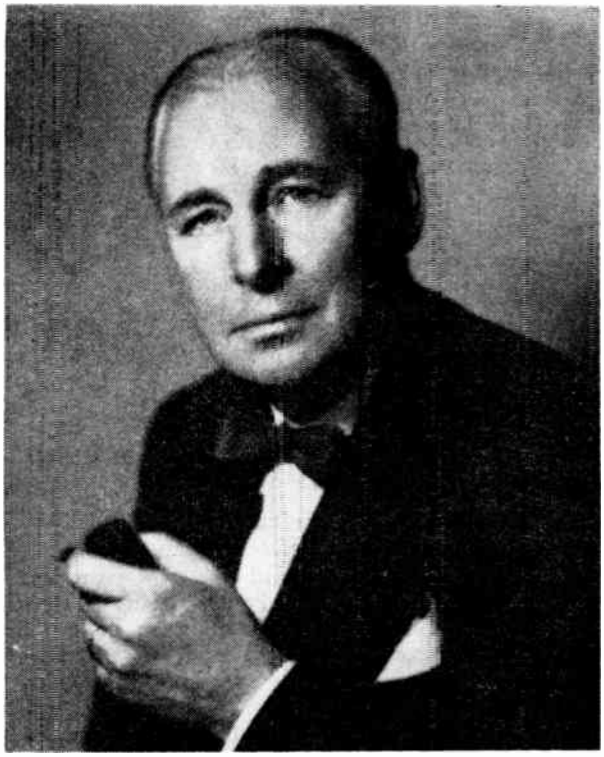
On the other hand, some of the nonelectrical areas regard the graduate years as time to be used in further developing a student's abilities in various specialties, or as time for some of the applications which could not be fitted into the basic four years' work. Both approaches may be proper for the needs of the respective employers, but we cannot help speculating on which program will best supply broad scientific leadership, as well as the teachers for the next generations.

Other differences are to be found in the viewpoint on the propriety of a research thesis, and the importance placed on the foreign language requirements. To this editor a degree without a thesis leaves the student less than complete, and the foreign language ability admits that those of other lands may also have ideas.

In our opinion last year's total of 647 doctorates lacks in both number and distribution if the requirements of our civilization are to be fulfilled in teaching, research, and advanced development. A concerted selling campaign, augmented by industrial financial support, must be developed in the colleges to the end that every boy standing in the upper half of his graduating class shall have graduate study presented to him in terms of its academic, professional, and economic advantages. Let us divert a bit of the undergraduate numbers campaign to the graduate scene as well, or our colleges will have the seniors teaching the freshmen. THIS has happened!

Scientific Marriage. Biomedicine, as a field of application of electronics, is recognized in this special issue of PROCEEDINGS. Proposed initially by T. A. Hunter of the IRE Editorial Board, and planned and put together by John W. Moore of the National Institutes of Health with the support of the Professional Group on Medical Electronics, the issue provides another illustration of the way in which electronics contributes to a sister science. The electronic potential of the field has long been recognized, but growth has been slowed by lack of people with training in *both* medicine or biology and electronics. That this situation is improving is demonstrated by the appearance of five IRE Fellows as authors in this issue.

A Saddlepoint. The Board of Directors and the lawyers having finished their work, the new Bylaws appear following page 2053 of this issue. Now we will really find those missing commas!—J.D.R.



Lloyd V. Berkner

Director, 1959

Lloyd V. Berkner (A'26–M'34–SM'43–F'47) was born February 1, 1905, in Milwaukee, Wis. He received the B.S.E.E. degree in 1927 from the University of Minnesota, which in 1952 honored him with the Distinguished Alumni Award, and from 1933–35 studied physics at the George Washington University. He holds honorary Doctorate degrees from Brooklyn Polytechnic Institute, Uppsala University in Sweden, University of Calcutta in India, Dartmouth College, University of Notre Dame, Columbia University, and the University of Edinburgh in Scotland.

While still an undergraduate, he was engineer-in-charge at radio station WLB-WGMS in Minnesota. For one year after graduation he worked as an electrical engineer for the Airways Division of the U. S. Bureau of Lighthouses. He was an engineer with the first Byrd Expedition to the Antarctic in 1928–30, and was awarded the U. S. Special Congressional Gold Medal, the Silver Medal of the Aeronautical Institute, and the Gold Medal of the City of New York for his services. For three years thereafter he was on the staff of the National Bureau of Standards. From 1933–1941 he was a physicist with the Department of Terrestrial Magnetism of the Carnegie Institution of Washington, and during 1940–1941 he was a consultant to the National Defense Committee.

An aviator in the Naval Reserve since 1926, Dr. Berkner was called to active duty as head of the Radar Section, Bureau of Aeronautics, in 1941. He directed the Bureau's Electronics Materiel Branch from 1943–1945, and served on the *U.S.S. Enterprise* in 1945. He has held the rank of Rear Admiral, USNR, since 1955.

During 1946–1947 he was Executive Secretary of the Research and Development Board and remained a consultant to the Board until 1951. He was head of the Section on Exploratory Geophysics of the Atmosphere, Department of Terrestrial Magnetism, Carnegie Institution, from 1947–1951. Since 1951 he has been President of Associated Universities, Inc., New York, New York, an educational institution which operates such research facilities as the Brookhaven National Laboratory under contract with the Atomic Energy Commission and the National Radio Astronomy Observatory under contact with the National Science Foundation.

Dr. Berkner has held numerous offices and advisory positions in government, industry, and education. In the State Department he served as Special Assistant to the Secretary of State, and Director of the Foreign Military Assistance Program in 1949, and Chairman of the International Science Steering Committee which produced the report "Science and Foreign Relations." Recently he was instrumental on national and international committees for the International Geophysical Year. He is presently a member of the Board of several industrial organizations. He was formerly a member of the President's Science Advisory Committee, and at the expiration of his term became consultant to the Committee.

He received the Science Award of the Washington Academy of Sciences in 1941; Commendation Ribbon of the Secretary of the Navy in 1944; Honorary Officer, Order of the British Empire in 1945; U. S. Legion of Merit in 1946; and Alumni Recognition Award of Acacia Fraternity in 1954.

He is Chairman of the Space Science Board of the National Academy of Sciences, President of the International Scientific Radio Union, past President of the International Council of Scientific Unions, and a former member of the Executive Committee of the International Union of Geodesy and Geophysics. He is a member of the National Academy of Sciences and of the American Philosophical Society, President of the American Geophysical Union, and a Fellow of the American Academy of Arts and Sciences, the American Institute of Electrical Engineers, the American Physical Society, the Arctic Institute of North America, and the New York Academy of Sciences. He is a Foreign Fellow of the Royal Swedish Academy of Sciences and holds membership in numerous other professional and honorary societies in the U. S. and abroad.

Guest Editorial*

JOHN W. MOORE†, SENIOR MEMBER, IRE

THE purpose of this special issue of the PROCEEDINGS is to provide its readers with some interesting, informative, and perhaps provocative examples of various weddings of electronic arts and concepts to some of the life sciences. This collection of articles is not intended to delineate "Biomedical Electronics" but rather to illustrate the breadth of the field of interest of the Professional Group on Medical Electronics which, by constitutional definition, is "the study of biological and medical systems."

The cover design suggests that electronic tools and concepts are useful at the levels of cells, organs, and the whole man. The experimental articles are arranged roughly in this order, but starting at the very subcellular level of molecular reactions. It is regretted that illness of one prospective author prevented him from writing a paper on the measurement of human responses and performance in space.

Not only have many biological measurements been made more convenient and accurate, but also new avenues of experimentation have been opened by the speed, versatility, and precision of modern electronics. It is hoped that this is not a unidirectional flow and that the difficult electronic problems faced in the life sciences will stimulate the development of instrumentation which will be generally useful. As an example of this type of feedback, a number of persons working in electrophysiology participated in the development of a rather broad-band electrometer preamplifier. This is described in the article by Gesteland, Howland, Lettvin, and Pitts entitled "Comments on Microelectrodes."

It was encouraging to the guest editor to find, in the process of compiling this issue, an increasing number of individuals with competence in both the electronic and biological areas who could speak and think in both languages. In many of the articles of this issue the electronic-biological subject complex is treated from an over-all point of view. With the increasing number of complicated biomedical problems, more of this "systems engineering" approach will be required to avoid undesirable and unintended interaction between the electronic and biological systems.

It is very often true that the direct application of electronics in medical diagnosis or therapy is quite difficult, and may be impossible until appropriate basic research has been done to elucidate the problem so that a clear statement of the electronic requirements can be made. Therefore, this issue gives considerable emphasis to basic biological research.

The recruitment and training of personnel to work in the biomedical instrumentation area is probably the most pressing problem to be faced by the Professional Group on Medical Electronics. Because of this a supplementary group of articles has been included, one of which tells of the origin and early history of the PGME. A report of the National Academy of Science—National Research Council is reproduced in which the problem is carefully stated and some recommendations are made. Two additional papers give comments on the problem and suggestions as to how it might be tackled.

* Original manuscript received by the IRE, September 14, 1959.

† Biophysics Lab., Natl. Institute of Neurological Diseases, Natl. Institutes of Health, Bethesda, Md.

An Analog Computer to Simulate Systems of Coupled Bimolecular Reactions*

E. F. MACNICHOL, JR.†, MEMBER, IRE

Summary—**An analog computer has been constructed to simulate, as nearly as possible, the flux of material in systems of coupled chemical reactions. Concentrations of various reactants, intermediates, and products are represented by the potentials at the outputs of electronic integrators. Rates of turnover of materials are represented by charges flowing to and from the integrators. The charges are caused to circulate by means of a "pump" mechanism that transfers charge at a rate proportional to the triple product of three voltages, two of which are derived from integrators and represent the concentrations of reactants. The third represents a rate constant. One voltage controls the frequency of an oscillator, the second, the duration of a triangular waveform, which is triggered by the oscillator, and the third, its rate of rise. By suitable interconnection of a number of integrators and pumps, a wide variety of reaction schemes can be simulated.**

Introduction

THE metabolic processes in living organisms are known to be composed of successions of individual reactions each of which is catalyzed by a specific enzyme. These steps are coupled to one another and to various side reactions. In order to be able to pick out the true reaction sequence in a given metabolic pathway from a number of possible alternative systems, it is desirable to compare the kinetic data obtained on a given biochemical system with those predicted by the various alternative schemes. The formal mathematics required to predict the kinetic behavior of such systems consist of a system of first-order second-degree differential equations expressing the rate of change of concentration of the various reactants and products as a function of time. To find out how the concentration of a given product changes with time, these equations must be solved simultaneously, a procedure which becomes very time-consuming and laborious as the number of reaction steps increases. While accurate solutions can be obtained on a large-scale general-purpose digital computer and the results can be plotted to display the kinetic data, it was felt that a simple analog machine that could be used in the laboratory to give approximate solutions would be a real help in deciding between various reaction schemes. Once the most likely system had been found, accurate analysis could then be made on a digital computer. Some years ago Chance[1] and his colleagues initiated such a program, and have had considerable success in comparing the kinetic data obtained by his rapid recording spectrophotometric technique, oxygen polarography, and other methods with the results predicted by a rather large analog computer. The device described in this paper was designed to do a job similar to the Chance computer's but with a maximum of simplicity and economy. Although it was constructed some time ago (1951–1953), it is worthwhile to describe it here because of the somewhat unusual multiplying circuit which makes use of sawtooth waveforms.

The Problems and Methods of Solution

The basic process to be considered is the bimolecular reaction in which substances A and B combine to form a product C. By the law of mass action, the rate at which the product is formed is proportional to the concentration of the reactants. Thus the simple bimolecular reaction $A+B\rightarrow C$, the rate of formation of the product $dC/dt=\alpha_1 AB$.[2] If the reaction is reversible, that is, if C splits up into A and B, the increase in C can be expressed in terms of the difference between the two opposing reactions. $dC/dt=\alpha_1 AB-\beta_1 C$ where α_1 and β_1 are the rate constants of the two reactions. In a biochemical system, several of these reactions can operate in sequence, as illustrated by the following schema:

$$\text{source} \rightarrow A + B \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} \overset{\overset{D}{+}}{C} \underset{\beta_2}{\overset{\alpha_2}{\rightleftharpoons}} E \rightarrow Z \rightarrow \text{sink}.$$

The reaction may also be coupled in some as yet unknown way to another reaction sequence $E\rightarrow F$ in which the rate at which E is converted to F is determined by the concentration of the product C.

It is also possible to represent energetically coupled reaction systems in which $K\rightarrow L$ molecule for molecule as $N\rightarrow P$, although K and N are intermediates in different reaction sequences.

In still other systems it is possible to have intermediates which are changed during one step of a reaction and subsequently regenerated so that they are again free to combine with a reactant.

In order to build a model that was sufficiently flexible to represent these and other types of systems, it was decided to represent the material fluxes in the chemical system by the flow of electric charges onto or off of capacitors in electronic integrators. The concentration of each substance is then represented by the voltage across

* Original manuscript received by the IRE, September 1, 1959. This work was done under Contract Nonr-248(11) between The Johns Hopkins University and the Office of Naval Research.

† Thomas C. Jenkins Lab. of Biophysics, The John Hopkins University, Baltimore, Md.

[1] B. Chance, D. S. Greenstein, J. Higgins, and C. C. Yang, "The mechanism of catalase action, part II," *Arch. Biochem. Biophys.*, vol. 37, pp. 322–339; June, 1952.

[2] In the chemical equations, A, B, C, etc., will be used to designate chemical substances. In the mathematical equations they will represent the concentrations of these substances.

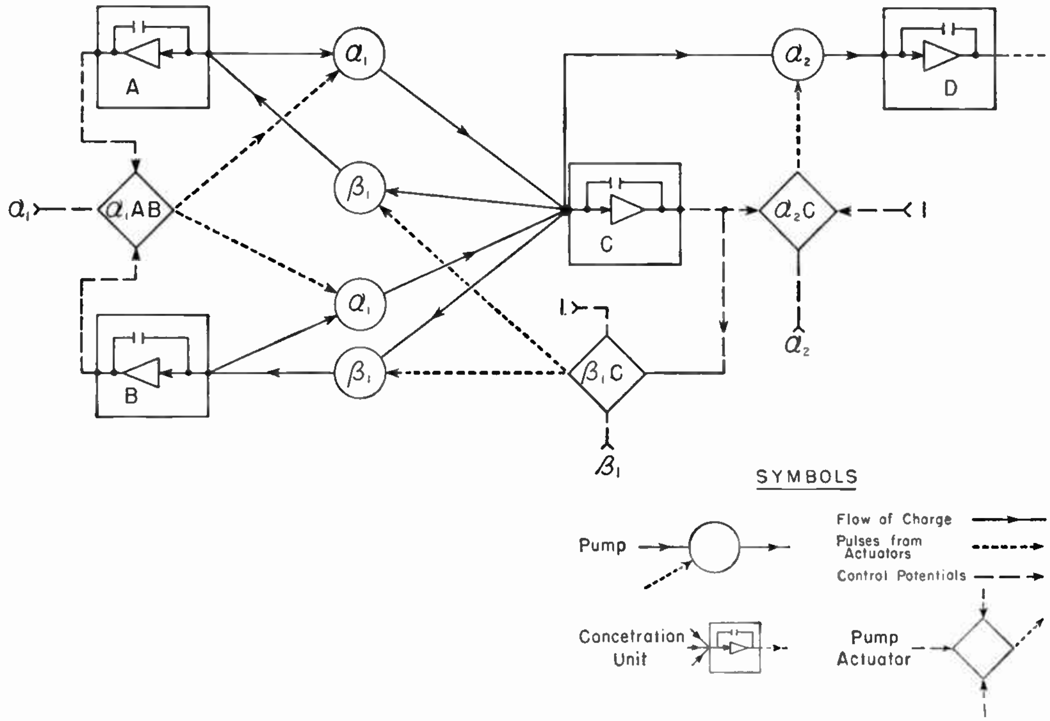


Fig. 1—Block diagram of analog of reaction.

$$A + B \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} C \overset{\alpha_2}{\rightarrow} D.$$

a capacitor in the corresponding integrator. This is done in the following manner: the total quantity of a substance is represented as a charge q on a capacitor. The concentration is given by the potential $v = q/c$. Thus the capacitance c plays a role analogous to the volume of the chemical system. The currents that charge and discharge the capacitors represent the flux of materials through the system

$$i = \frac{dq}{dt} = \frac{cdv}{dt}$$

and the concentration of a substance at time T is

$$v = \frac{1}{c} \int_0^T idt.$$

To complete the system, "pumps" are required to produce a flow of charge from capacitors representing reactants to other capacitors representing products. This charge must flow at a rate determined by the triple product of two concentrations (voltages across capacitors) and a rate constant (fixed potential). Thus, a pump circuit is required for each reaction step (two are required for a reversible step) and an electronic integrator (concentration unit) is required for each participating substance (reactant, intermediate, or product). Fig. 1 shows how the components are interconnected for a single bimolecular reversible step.

The charge pump consists of two units: an actuator, which is controlled by voltages proportional to the rate constant and to the concentrations of the reactants, and the pump unit itself, which transfers charge in response to pulses from the actuator. Thus three integrators are required to represent substances A, B, and C. Two pump actuators represent the forward and backward reaction rates and four pumps are operated by them. Two pumps are under control of the forward actuator and transfer charge from A to C and from B to C; the other pair are under control of the reverse actuator and transfer charge back from C to A and B.

In a typical sequence of reactions, C might react with D and produce products E and F. For this step two more integrators, two more actuators, and four more pumps would be required. In general, for n substances in the system, n integrators, and about n actuators and $2n$ pumps will be required, although fewer will be required when steps are monomolecular and irreversible and more will be needed when there is complex branching and coupling between reaction sequences.

The integrators are conventional phase-inverting operational amplifiers with a capacitance connected between input and output. Thus, negative charge flowing into the input circuit at zero potential produces a positive voltage at the output terminals proportional to the integral of the current according to the well-known principle of the Miller integrator.[3]

[3] B. Chance, *et al.*, "Waveforms," M. I. T. Rad. Lab. Ser., McGraw-Hill Book Co., Inc., New York, N. Y., vol. 19, pp. 31, 195; 1949.

The pump actuators produce triangular pulses of amplitude proportional to the product of two variables, and frequency proportional to a third.

The pump units are diode step counters which transfer a pulse of charge proportional to the amplitude of each actuating pulse. Since the current is proportional to the quantity of charge transferred by each pulse, and to the number of pulses per second, it is proportional to the product of the three input quantities.

Fig. 2 indicates in more detail how a single bimolecular step is represented. The rate constant α_1 is set in as a fixed potential to control the frequency of the voltage-controlled oscillator in the pump actuator. The voltage V_A representing the concentration of reactant A controls the duration of the gate pulse from the rectangle generator and is derived from integrator A. Similarly, the voltage V_B controls the slope of the triangular waveform from the triangle generator. The output wave from the triangle generator applied to pump No. 1 causes charge to be transferred from the B integrator to the C integrator, thus increasing V_C which represents the concentration of the product and decreasing V_B.

Since V_A must decrease at the same rate that V_C increases and V_B decreases, there are two possible arrangements. Either charge from A and B can be pumped to C and the capacitor in the C integrator can be made twice as large as those in the A and B units, or the output of pump No. 2 can be thrown away since it is equal to that of pump No. 1 under all conditions. The second alternative, which is indicated in Fig. 2, is usually the more convenient since it simplifies scaling problems.

Although only very simple situations are indicated in this paper, it is evident that given a sufficient number of integrators, actuators, and pumps, it is possible to represent reaction systems of arbitrary complexity. It is also possible to represent autocatalytic reactions by connecting one of the pump actuator inputs to an integrator toward which charge is being circulated so that pumping becomes more rapid as the concentration increases.

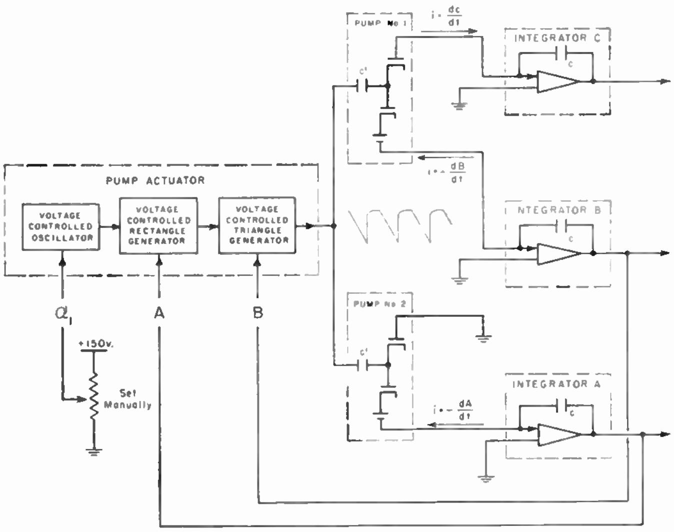


Fig. 2—Block diagram of circuit representing reaction step $A+B \xrightarrow{\alpha_1} AB=C$ by differential equation

$$\frac{dC}{dt} = \alpha_1 AB = \frac{-dA}{dt} = \frac{-dB}{dt}.$$

Circuit Details

The entire system is operated from conventional regulated power supplies that furnish +300 volts and −250 volts with an accuracy of about ±2 per cent. Most of the operating current is taken directly from these supplies. Precision voltages of +150 volts and −150 volts are obtained by further regulation of the output voltages from these supplies. These voltages are held to about ±0.1 per cent and are used to supply the potentiometers for setting initial conditions and for stabilization of the waveforms generated in the pump actuators.

The waveforms produced by the pump actuators are shown in Fig. 3, and a schematic diagram is shown in Fig. 4. An astable modification of the screen-suppressor coupled phantastron[4] is used to produce a recurrent sawtooth of fixed amplitude. The Miller feedback sawtooth generator comprises pentode V_1 and cathode follower V_{2a}, which is used to speed up recharge of the timing condenser. The linear run-down of the sawtooth waveform lasts until the pentode "bottoms" at the knee of its characteristic curve increasing screen current, which cuts off the plate current by making the suppressor grid negative. The plate voltage rises rapidly, thus charging the timing capacitor until the plate of diode V_{3a} becomes positive of its cathode and bringing the suppressor grid of V_1 positive so that plate current again flows and the linear run-down recommences. The rate at which the sawtooth waveform is repeated is proportional to the slope of the sawtooth wave, which is in turn proportional to the potential V_a to which the timing resistor is returned and inversely proportional to the RC product of the timing circuit. Thus, a frequency proportional to V_a is obtained.

The rectangular switching waveform at the screen of V_1 is amplified by V_{2a} and is used to shock excite a small inductance giving a positive trigger pulse when screen current increases at the end of the run-down.

This pulse is used to trigger the screen-suppressor coupled phantastron V_4, V_{5a} through diode V_{3b}. In the quiescent state, the screen grid of V_4 is drawing heavy current so that the suppressor is negative and plate current is cut off. The grid voltage of cathode follower V_{5a} is determined by the voltage V_B applied through diode V_{6a}. This determines the starting voltage of the sawtooth waveform generated across the timing capacitor.

[4] E. F. MacNichol, Jr. and J. A. H. Jacobs, "Electronic device for the measurement of reciprocal time intervals," *Rev. Sci. Instr.*, vol. 26, pp. 1176–1180; December, 1955.

The trigger pulse applied through V_{3b} causes the suppressor grid to go positive, thus initiating plate current and reducing screen current so that the run-down state of the phantastron is initiated. The linear run-down lasts until diode V_6 conducts when the grid of V_{5a} reaches ground potential. The timing resistor is returned to a fixed voltage. The input signal, V_B, determines the total length of the sawtooth and, hence, its duration.

The positive rectangular wave at the suppressor of V_4 is applied through cathode follower V_{5b} to the suppressor grid of V_8. This tube, together with the double cathode follower V_9, V_{10}, forms a triangular waveform generator. In the quiescent condition, plate current in V_8 is cut off so that plate voltage is maintained at +150 volts through diode V_{7b}. While the suppressor is brought to zero volts during the gating signal, a linear negative-going sawtooth is produced at the output circuit by means of Miller feedback. The slope of the sawtooth is determined by V_C and the length of time the sawtooth runs by the duration of the gating waveform. Thus, the amplitude of the sawtooth is determined by the product of slope and duration which are, in turn, proportional to V_B and V_C. This waveform actuates the pumps which, as shown in Fig. 1, consist of diode step counters with a 3-volt battery to prevent reverse conduction. The time constant produced by the output impedance of the double cathode follower V_9, V_{10} and the pump capacitor C' is sufficiently short so that C' charges to within 1 per cent of the peak value of the triangular wave. Thus, the charge transferred between integrators per cycle is proportional only to the amplitude of the wave.

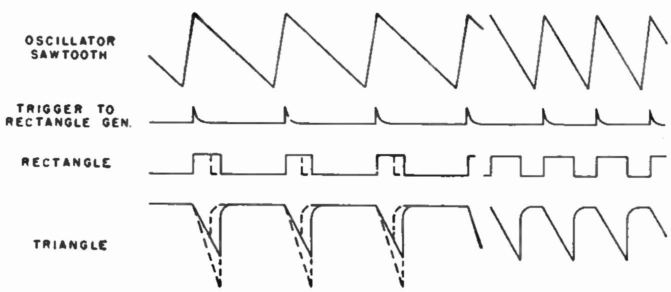


Fig. 3—Waveforms in pump actuator.

The integrator is shown in Fig. 5 and consists of a Miller feedback circuit using an operational amplifier having a grid current of about 10^{-10} amperes, a voltage gain of about 3000 and a range of output voltage limited at 0 volts and +150 volts (concentrations are never negative). The design is quite conventional.

In order to avoid floating power supplies for setting the initial conditions, the amplifier is converted into a unity gain feedback amplifier during the charging of the integrating capacitor. The method by which this is accomplished is indicated in the simplified circuit of Fig. 6.

To set initial conditions, a relay operates in each of the integrators. It transfers one side of the integrating capacitor from the input grid of the operational amplifier to ground. Simultaneously, the grid is connected to the common point of two precision resistors R which are connected in series between the output of the operational amplifier and the initial conditions potentiometer. The output of the amplifier is therefore required to

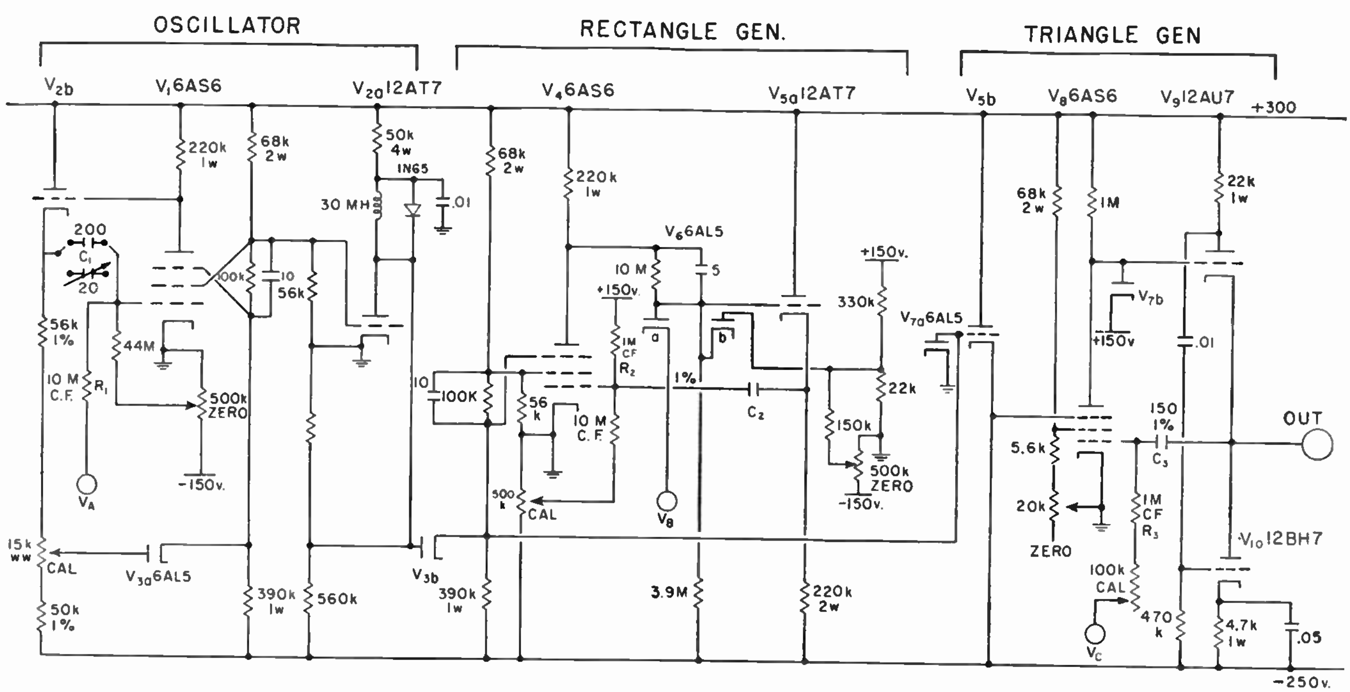


Fig. 4—Schematic diagram of pump actuator.

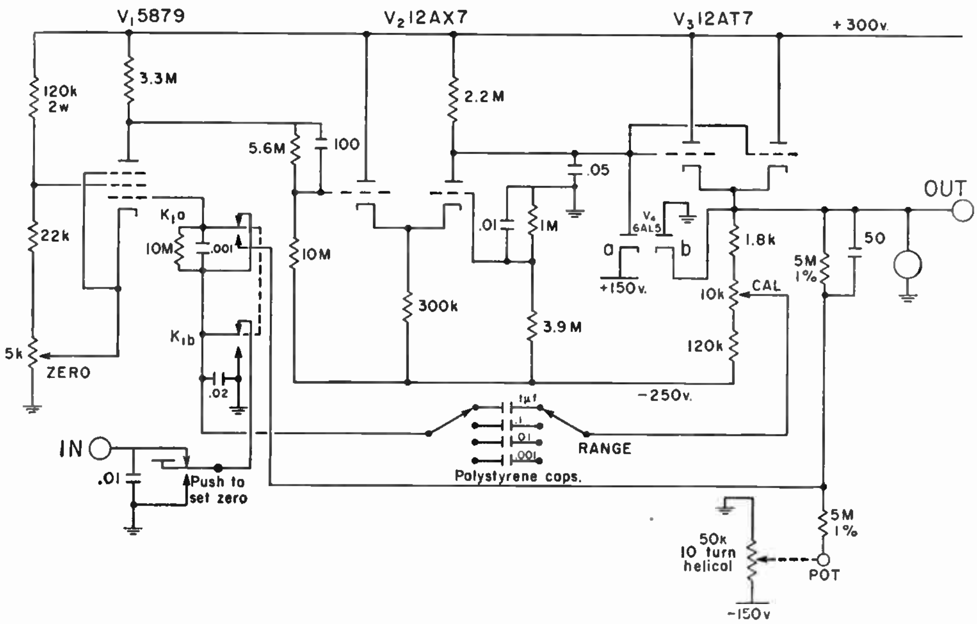


Fig. 5—Schematic diagram of integrator.

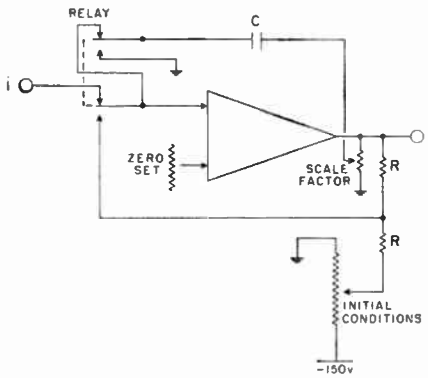


Fig. 6—Circuits for normalizing and setting initial conditions in integrator.

assume an equal and opposite voltage to that set in on the potentiometer. This charges the integrating capacitor to the correct initial potential.

Because it is difficult to obtain integrating capacitors matched to sufficient accuracy, an adjustment of the "effective" capacitance is provided. The capacitors are padded to slightly higher capacitance than the desired value and the signal applied to the capacitor is attenuated by an adjustable amount. Thus, the output voltage change is greater for a given value of capacitance than it would be if the capacitor were connected directly to the output.

Results

A system comprising six integrators, five pump actuators, and a suitable number of pumps was constructed. Data were displayed on a multichannel direct-writing recorder (Sanborn). It was possible to obtain currents from the pump units that were proportional to the triple product of the three input quantities within an accuracy of about 5 per cent of full scale within a 20-fold variation of each of the input quantities. After careful adjustment, it was possible to hold steady-state conditions for sufficient time to explore the effects of transient changes in rate constants or sudden changes in the concentration of one of the reactants.

Acknowledgment

The author wishes to express his thanks to Drs. H. K. Hartline, F. Brink, and J. P. Hervey for their helpful encouragement and criticism, to E. N. Shipley for constructing the device, and to Dr. C. C. Yang for getting it to work.

Electron Transfer in Biological Systems*

BRITTON CHANCE†, FELLOW, IRE

***Summary*—The direct approach to a study of the dynamics of the essential intermediates in life processes is afforded by sensitive optical techniques that accurately measure and record the absorbancy of the iron proteins (cytochromes) in living cells, tissues, and particles derived therefrom. This paper outlines the nature of physical phenomena measurable in the biological systems and emphasizes current thinking on the nature of electron transfer between the proteins which involves oxidation and reduction of their iron-containing active centers with the simultaneous conservation of energy required for driving essential biological processes. The methods for these measurements are reviewed and spectrophotometric techniques at room and liquid nitrogen temperatures, and new developments such as microspectrophotometry of the cytochromes in portions of the living cell are emphasized. Data evaluation and representation of electron transfer processes and metabolic control sequences by analog and digital computers are described and particular reference is made to the operation of metabolic controls in ascites tumor cells.**

Introduction

THE basic reactions upon which one of the most fundamental processes of life in man depends are

$$Fe^{+++} + e \rightarrow Fe^{++} \quad (1)$$

$$Fe^{++} - e \rightarrow Fe^{+++}, \quad (2)$$

a pair of electron transfer reactions in which ferric iron is reduced to the ferrous form and reoxidized to the ferric form [1]–[5]. The source of the electrons is food stuffs, such as glucose. An ultimate "sink" for the electrons is the oxygen molecule which is made available to the living cell by the blood stream and is, itself, eventually reduced to water by the following reaction. One of the steps may be

$$Fe^{++} + O_2 \rightarrow Fe^{+++} + O_2^{-}. \quad (3)$$

These reactions provide much of the energy needed to form the "common currency" of energy exchange of the body, adenosine triphosphate (ATP) [6]. This substance is an energy source for mechanical work, as in muscular contraction; for the transmission of electrical impulses in nerves; for the accumulation of ions; for the synthesis of new materials for growth; and even for the production of light in living systems. Two current problems of modern biophysics and biochemistry are to determine how the electrons are transferred, and how a portion of the energy available from the electron transfer reaction can be converted into the common currency of energy exchange, ATP.

In this paper we will first summarize the current state of knowledge of electron transfer and energy conservation processes in biological systems; then take up physical and electronic techniques that are suitable for their study; and finally, consider the extent to which analog and digital computers can be used in the representation of such phenomena.

The Physical and Chemical Properties of the System

Components of the Electron Transfer System

The iron atom involved in the electron transfer reactions of (1) and (2) is attached first to a small porphyrin group having a molecular weight of ~600 and then to a large protein whose molecular weight is in the range of 10^4–10^5 [7].

The absorption bands of these iron-proteins (hemoproteins) are distinctive in the visible and the ultraviolet regions of the spectrum. In the visible region are two bands, α and β (Fig. 1), and in the ultraviolet region

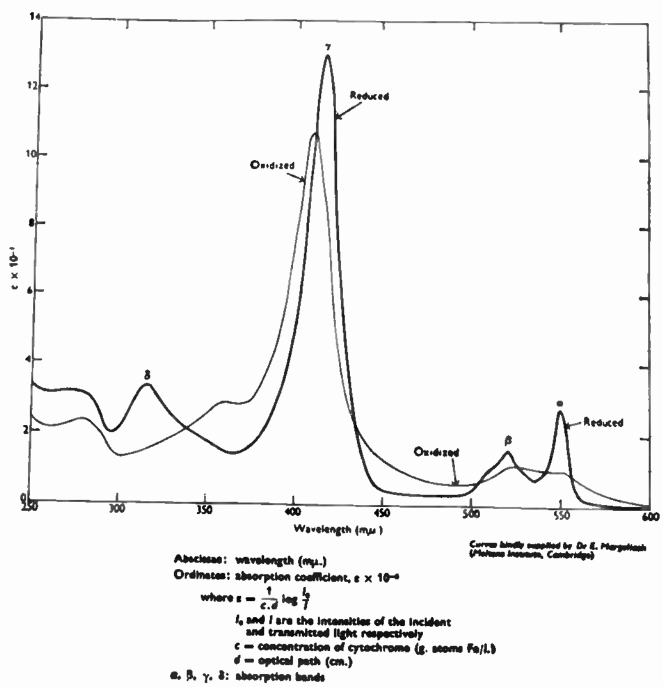


Fig. 1—The spectrum of the oxidized and reduced forms of a highly purified sample of cytochrome *c* (0.45 per cent iron). The appearance of sharp absorption bands (α, β, γ, δ) upon reduction of the iron of cytochrome *c* affords a basis for sensitive spectrophotometric methods for measuring the extent of its reduction. Absorbancy differences between the oxidized and reduced forms are maximal in the region of the γ band (0–71). (Reprinted with permission of the *British Medical Bulletin* [4].)

* Original manuscript received by the IRE, September 1, 1959. This research has been supported in part by grants from the National Science Foundation, the Office of Naval Research, the U. S. Public Health Service, and the American Cancer Society, and the Muscular Dystrophy Association. The computations were supported by the University of Pennsylvania Computer Center and, in part, by the National Science Foundation. The author wishes to express his thanks to these supporting organizations and also to the members of the Johnson Foundation staff who have contributed to this research.

† Director, Johnson Research Foundation, University of Pennsylvania, Philadelphia, Pa.

are the γ and δ bands [3], [4]. These four bands are largely characteristic of the iron porphyrin group itself, but are modified by the parts of the protein that absorb at 275 mμ. Because of small differences in the porphyrin and protein parts there are different and characteristic absorption bands in the visible and near ultraviolet regions for various iron proteins which are thus characterized by the positions of these bands. A particularly useful feature of these spectra is a sharpening that occurs at liquid nitrogen temperatures (see Fig. 2) [8], [9]. Under these conditions, the α and β bands show considerable fine structure and much more precise delineation is possible.

It should be noted that the absorption bands of the reduced (ferrous) forms are much sharper than those of the oxidized forms, and that a subtraction of the two spectra gives absorption bands that can still be closely identified with those of the reduced form (Fig. 3).

The most important group of these iron proteins is that associated with the electron transfer processes of the cell. For this reason, they have been named cell pigments or cytochromes [1]. These macromolecules do not exist in solution in the cell but are intimately bound in a solid structure within the cell, called a mitochondrion. Five different cytochromes are required to act in sequence for electron transfer, and it is probable that their appropriate geometric sequence is built into the structure of the mitochondria in order to facilitate the basic electron transfer reaction when the ferrous or reduced form of one transfers an electron to the ferric or oxidized form of the next:

$$(Fe^{++})_b + (Fe^{+++})_c \rightarrow (Fe^{+++})_b + (Fe^{++})_c. \qquad (4)$$

A sequence of cytochromes that has been found to occur in mammalian tissues such as heart, liver, and kidney is illustrated schematically in Fig. 4(a) [3], [4], [10]. Here four different cytochromes (a_3, a, c, b) are represented as bound to the cell structure with sufficient possibilities for vibration or rotation about their points of attachment to permit electron transfer by iron-iron

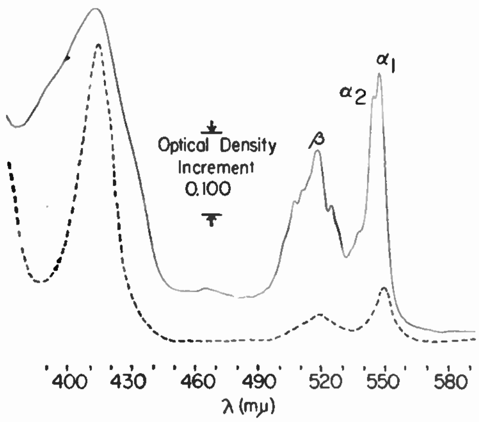


Fig. 2—The effect of temperature upon the absorption spectrum of purified and reduced cytochrome c. The dashed curve was obtained at 300°K and the solid curve was obtained at ~77°K. At the low temperature the α and β bands are considerably intensified and sharpened and show fine structure (RE-2). (Reprinted with permission of the *Journal of Biological Chemistry* [9].)

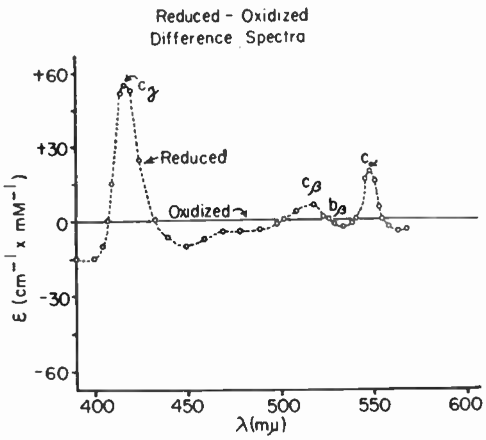


Fig. 3—A room-temperature difference spectrum of cytochrome c, obtained by subtracting the oxidized and reduced spectra of Fig. 1. This type of spectrum is obtained from living material by the methods described on page 1829. (MD-38).

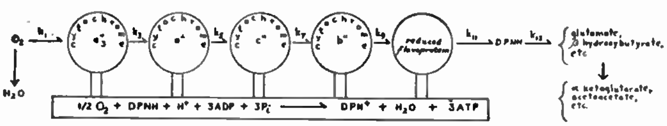


Fig. 4(a)—A schematic representation of the electron transfer system illustrating the chemical reactions involved in the reduction of oxygen and the oxidation of substrates (glutamate, etc.) by the transfer of electrons through a series of carriers (cytochromes a_3, a, c, b contain iron; flavoprotein contains flavinadenine dinucleotide; and DPNH represents diphosphopyridine nucleotide). The velocity constants for the reaction are identified by $k_1 - k_{13}$. The probable attachment of these iron proteins to a structure of a portion of the cell (mitochondrion) is also indicated. The over-all equations for the reaction involve the conversion of three molecules of ADP and P_i (adenosine diphosphate and phosphate) to three molecules of ATP (adenosine triphosphate) with the expenditure of a molecule of DPNH and a half-molecule of oxygen. The ATP supplies energy for many cell functions. Attachment of the molecules to a portion of cell in such a way so that energy transfer by collision is still possible is also indicated (MD-39).

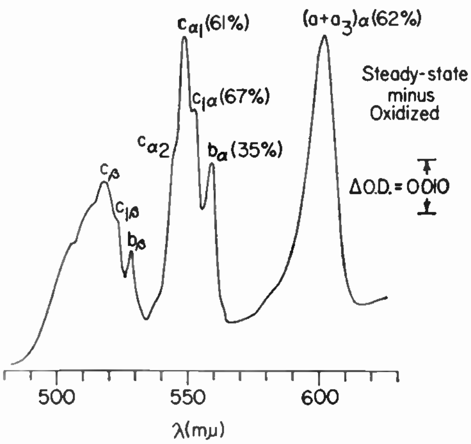


Fig. 4(b)—Spectrum representing the difference between the reduced and oxidized forms of cytochrome of particles isolated from beef heart, measured at the temperature of liquid nitrogen. The cytochromes are designated with the appropriate letter in the figure [cf. Fig. 4(a)]; cytochrome c_1 is between c and b in the sequence (Expt. 966b).

collisions. Thus the substrate, glutamate, is oxidized to α ketoglutarate and oxygen is reduced to water.

The accompanying formation of ATP occurs in the ratio of about three molecules of ATP per atom of oxygen [10a].

The identification of cytochromes of this chain by low temperature spectroscopy is indicated in Fig. 4(b) which represents the difference of absorbancy between the steady-state reduced and oxidized cytochrome components of beef heart mitochondria measured at temperature of liquid nitrogen (see below). The sharp peaks appropriate to the components of Fig. 4(a) are shown. In addition, cytochrome c_1, which acts between cytochromes *b* and *c*, is also found.

In addition to the possibility of collisions suggested by Fig. 4(a), two other possibilities may be considered: 1) that the iron proteins are immobilized in the structure of the cell (Fig. 5); 2) that conduction bands exist which permit electron transfer according to (1) and (2); and 3) that resonance energy transfer mechanisms are operative here. Detailed studies of the conductivity of purified iron-proteins suggest that inadequate conductivities exist to explain the observed rate of electron transfer in the living material [12], [13]. Thus, the bulk of the experimental evidence suggests that the collision hypothesis still has to be considered seriously for electron transfer in the biological systems [14]. Of particular relevance are the recent experiments which suggest that ferric and ferrous forms of the iron proteins can remain stable adjacent to one another for as long as five days at liquid nitrogen temperature [14]. In addition, changes of viscosity appear to influence the speed of electron transfer, although the possibility of a side reaction is not excluded in this case [14].

In the past several years light-induced ferrous-ferric changes in organisms whose life depends upon sunlight have been observed [15], [16]. This reaction is particularly easy to observe in the photosynthetic bacteria such as *Chromatium* and *Rhodospirillum rubrum*. In these reactions, the magnesium-porphyrin, bacterial chlorophyll, absorbs red light and then transfers energy to the iron atom of a cytochrome in a highly efficient reaction (2 quanta per electron [17]). It is of considerable interest to know whether a collision process such as that described above for the inter-cytochrome reactions is operative in the chlorophyll-cytochrome reaction. Sensitive spectrophotometric techniques described below afford a ready measurement of the disappearance of the γ or Soret band of a reduced cytochrome component of the living bacteria caused by its oxidation upon infrared illumination. A double-beam spectrophotometric recording is indicated in Fig. 6. In the left-hand record, obtained at 28°, illumination of the bacterium *Chromatium* with infrared light causes a downward sweep of the trace at a rate of 0.13 μM Fe/sec. On the cessation of illumination, the trace rises more slowly due to dark reactions which reduce the oxidized cytochrome. If the sample is now frozen to a temperature of −22°, illumination causes an oxidation of cytochrome at a rate approximately equal to that obtained at room temperature (0.11 μM Fe per second). Cessation of illumination under these conditions causes an extremely slow reduced reaction. The light-induced reaction can be observed at temperatures down to −190°C. It is apparent that two types of reactions are exhibited here: 1) a nonthermal oxidation reaction presumably due to a direct chlorophyll-cytochrome energy transfer, and 2) a much slower thermal reduction reaction which probably involves collision of this particular cytochrome with other members of its transfer chain.

Calvin [18] has observed electron spin resonance (esr) signals in the very similar bacterium *R. rubrum*. The signals are caused by infrared illumination at +25° and at −15° (Fig. 7). Similar esr signals have been ob-

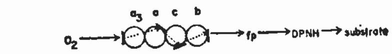


Fig. 5—A representation of a hypothetical process through the cytochrome chain in which all electrons must pass through the iron atoms of the porphyrin group in order to account for the observed oxidation and reduction of the iron (MD-41). (Reprinted with permission of Interscience Publishers, Inc., New York, N. Y. [13].)

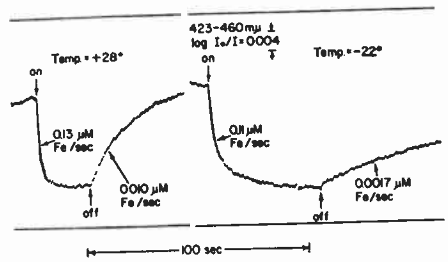


Fig. 6—A double-beam spectrophotometric recording of the effect of temperature on the light-induced oxidation of cytochrome 423 of the purple sulfer bacterium *Chromatium*. On the left the temperature is +28°C, and on the right the temperature is −22°C. The rates of oxidation of the cytochrome are computed on the basis of a molecular extinction coefficient of 100 $cm^{-1} \times mM^{-1}$. The cuvette of the double-beam spectrophotometer is contained in a Dewar flask and infrared illumination was reflected onto the sample from above (Expt. 3).

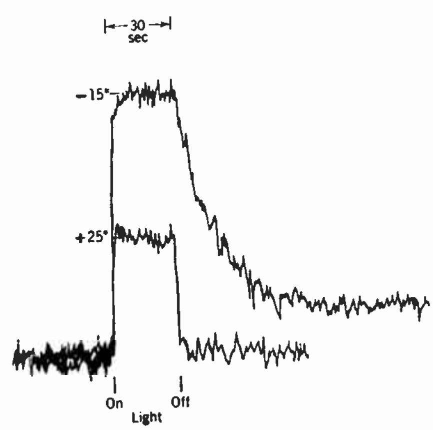


Fig. 7—Light-induced electron spin resonance signals from *Rhodospirillum rubrum* at +25° and −15° C. To improve the clarity of the figures, two traces given in the original reference are deleted and only those at temperatures similar to the spectrophotometric data of Fig. 6 are included (0-117). (Courtesy Dr. M. Calvin [18]).

tained from the type of bacteria used in the experiment of Fig. 6, *Chromatium* [18a]. The spectrophotometric data indicate that an electron (or proton) transfer from cytochrome to chlorophyll can occur in a process that is very nearly temperature-independent. Thus, the esr signal could be identified with the valency change of the iron or with the electron donated by cytochrome to chlorophyll. The first suggestion is unlikely, since iron proteins similar to cytochrome *c* give no measurable esr signal of $g=2$. The second is a more acceptable hypothesis and one that suggests the desirability of a spectrophotometric study of bacterial chlorophyll under these conditions. These experiments underline the importance of combined studies by different physical methods; the esr method suggests an odd electron, but does not identify the chemical species with which it may be associated. The spectrophotometric technique identifies the chemical species involved in the oxidation reaction, but does not indicate the odd element.

Structure of the Electron Transfer System

As indicated in Fig. 4(a), molecules of the electron transfer system are arranged to make a functional "assembly." Many of these molecules appear to be gathered together to form a definite structure within the cell called the mitochondrion. Some types of mitochondria contain $\sim 10^4$ of these molecular assemblies [19]. The accompanying electron micrograph (Fig. 8) shows three mitochondria in frog muscle. The cytochromes of the mitochondria utilize the oxygen brought in by the blood supply and, by the electron transfer process, ATP is formed which "recharges" the muscle fibrils for energy expended in a contraction. The intimate association of the energy source and energy sink is indicated by the contact of the mitochondria and the muscle fibrils.

A very close relationship exists between the expenditure of energy and the activation of oxidation-reduction reactions in the mitochondria, as indicated by the spectrophotometric tracings of cytochrome action in Fig. 9 described below (see Fig. 11).

Identification of Cytochromes with the Mitochondrial Structure

Direct identification of the cytochromes with the characteristic structure, or cristae, of the mitochondria by light absorption methods presents a considerable challenge to electronic techniques. The size of the mitochondrion, as indicated by scale on the diagram (Fig. 8), is between one and two micra, a dimension at which optical measurements are difficult to make because of the small amount of light passing through the specimen. In some biological materials a larger aggregate of mitochondria (known as the Nebenkern) is found, particularly in the grasshopper sperm cell. Apparently the immature cell gathers all its electron transferring capacities in one place in preparation for forming the sperm tail. Fig. 9 is an electron micrograph of a Nebenkern

Fig. 8—Electron micrograph of frog muscle showing three mitochondria (top center, lower left, and lower right) in intimate contact with the myofibrils to which they supply energy expended in muscle contraction. This micrograph affords an example of the intimate arrangement of energy sources and sinks in biological systems (0–95). (Courtesy Dr. R. Birks and Dr. A. Huxley, Oxford University, Eng.)

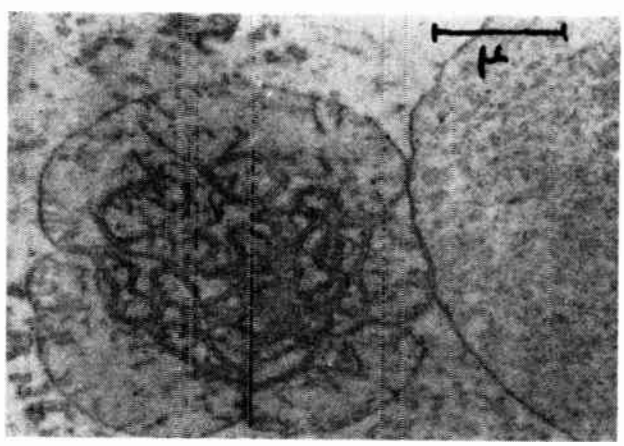


Fig. 9—A mitochondrial aggregate (Nebenkern) of the grasshopper sperm cell which is of sufficient size for accurate microspectrophotometric observation the Nebenkern is to the left of the center of the figure; the nucleus is to the right (IG-3). (Courtesy of Dr. I. R. Gibbons, Harvard University, Cambridge, Mass.)

of a fixed cell which shows a structure typical of mitochondria; but, in this case, the mitochondrion is about 3 μ in diameter, a more favorable size for a suitable signal-to-noise ratio in an electronic system. Fig. 10 illustrates the spectra that appear when a scanning microspectrophotometer [21] is set to record absorption bands in the Nebenkern portion of the living cell (Trace A). The band of cytochrome *c* occurs at 420 mμ and that of cytochrome *b* (with some contribution from cytochrome a_3) at 434 mμ. If the scan is taken from a part of the cell containing none of the characteristic structures of the mitochondrion, then no distinctive peaks can be observed (Trace B).

Dynamic Responses of the Electron Transfer System

As mentioned above, the purpose of the electron transfer system is to provide energy for the function of the cell. In muscle (cf. Fig. 11), contraction can cause a direct activation of electron transfer. A spectrophotometric record [22] (Fig. 9) shows that as the muscle is stimulated electrically and the tension rises (lower trace) the activation of electron transfer (upper trace) begins with a lag of less than 200 msec after the initiation of the contraction. This record is a composite of three traces: the electrical stimulus (top trace), the muscle tension as measured by a strain gauge (bottom trace), and the optical effects, recorded by the double-beam spectrophotometer (middle trace). The reduced form of the cytochromes absorbs more intensely, and thus the ferrous-ferric shift indicated by (2) can be measured as the increasing absorption of light at a specific wavelength (430 mμ). The experimental method will be discussed in more detail below.

Another interesting response of the cytochromes of the living cell is afforded by a suspension of ascites tumor cells which gives spectroscopic evidence of a metabolic control process which may be of considerable importance in regulating the metabolism of this material [23]. Fig. 12 shows that the activation of metabolism in these cells by the sudden addition of glucose causes an increase of respiratory activity (recorded polarographically by means of the platinum microelectrode) and a simultaneous increase of electron transfer resulting in a ferrous-ferric shift of the cytochrome *b* component (recorded here as an upward deflection of the spectrophotometric trace). The response in this case is transient and the activated respiration terminates in about a minute, as does the spectroscopic response. Thus the metabolism of the cell has been inhibited by the addition of glucose. This metabolic control response has certain characteristics similar to that originally described by Pasteur [24], [25], and that more recently studied by Crabtree [26], Racker [27], and Chance and Hess [23]. The computer formulation of such a metabolic control is described below.

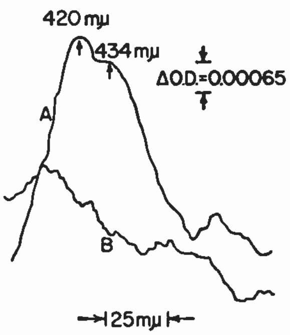


Fig. 10—A demonstration of the localization of cytochromes in mitochondria by means of microspectrophotometry. Trace A is a spectrum obtained with the microspectrophotometer focused on the mitochondrial body or Nebenkern (*cf.* Fig. 9). Trace B is obtained with the microspectrophotometer focused on the cytoplasm adjacent to the Nebenkern. In trace A the absorbancy peak at 420 mμ (cytochrome *c*) and a shoulder at 434 mμ (mainly due to cytochrome *b* but with some contribution from cytochrome a_3) can be identified—(RP-1). (Courtesy of Dr. R. Perry.) (Reprinted with the permission of the American Institute of Physics [21].)

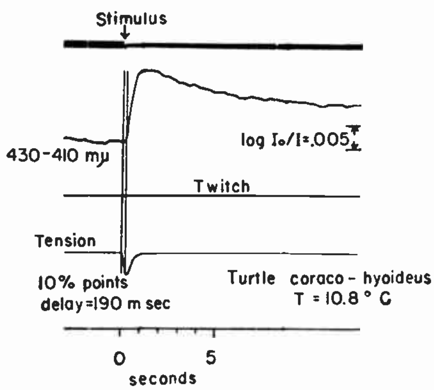


Fig. 11—The response of cytochrome *b* to muscular contraction. A correlation between the development of tension (lower trace) in turtle muscle and the development of a change in oxidation of cytochrome *b* (middle trace) as measured with the double-beam spectrophotometer. This record illustrates a significant delay in the cytochrome response (J-1). (Courtesy of Dr. F. Jobsis, Johnson Foundation University of Pennsylvania, Philadelphia, Pa.)

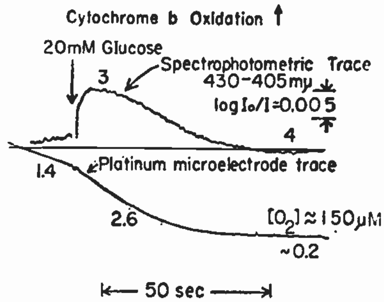


Fig. 12—The response of cytochrome to acceleration of metabolic activity. The metabolic activity is represented (lower trace) in terms of a continuous polarographic recording of oxygen utilization by a suspension of ascites tumor cells. The addition of glucose accelerates the rate of oxygen utilization from 1.4 to 2.6 μM $O_2 \times sec^{-1}$, but at the end of about a minute the metabolism is spontaneously inhibited to a rate of 0.2 μM$\times sec^{-1}$. The response of the cytochromes to this activation and inhibition phenomenon is recorded by the double-beam spectrophotometer (top trace) where the upper deflection corresponds to an oxidation of cytochrome *b* (the same component as was measured in Fig. 11). This oxidation subsides beyond the initial level as the metabolism is inhibited (498b). (Reprinted with the permission of the New York Academy of Science [23a].

Physical and Electronic Techniques

The foregoing experimental results suggest the great importance of simultaneous measurements of two or more parameters of the biological system so that some constant property of the living system may be continuously monitored at the same time that one of its less known physical attributes is being investigated [28]. Some of the techniques suitable for such measurements are described below.

Measurements of Oxygen Utilization

A general property of mammalian cells and tissues is their oxygen utilization, and the platinum electrode technique of Kolthoff and Lingane [29], as improved by Davies and Brink [30] and more recently by Harris and Lindsey [31], affords a versatile method for measuring the oxygen utilization of small amounts of cells and tissues. A method found to be more convenient for simultaneous studies of cell suspensions and isolated mitochondria is the vibrating electrode technique illustrated by Fig. 13. An "open" glass-coated electrode [30] is attached to the reed of a Brown Instrument "converter." The performance of this device in measuring the acceleration and deceleration of respiration in a suspension of ascites cells is indicated by the extreme right-hand portion of Fig. 12. The addition of glucose accelerates respiration and the change in rate is obtained without an appreciable stirring artifact. Thus the vibration rate is fast enough so that stirring causes no change of diffusion current. Another requirement which will become clear in the discussion of the associated spectrophotometer is that the vibration of the electrode be synchronous with the flickering of a light beam through the solution. The electrode is initially calibrated in an air-saturated medium and soon thereafter in an oxygen-free solution. Thus, long-time stability of the electrode sensitivity is not required. The method is, however, unsatisfactory for suspensions having very slow respiration rates because oxygen is slowly stirred into the solution. The electrode is also sensitive to substances which may be added to the reaction mixture, such as cyanide, and covered electrodes have been used by Davies [33] and by Clark [34], [35].

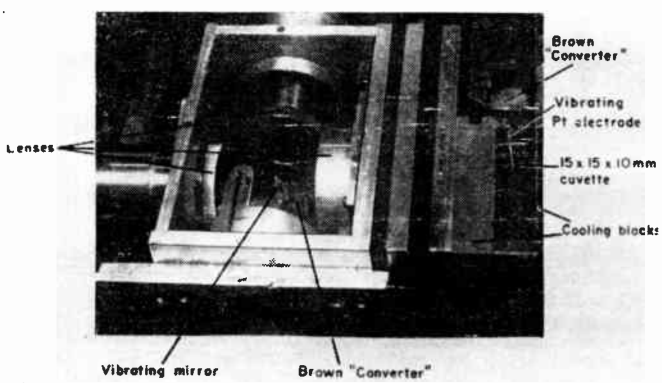


Fig. 13—An illustration of the adaptation of the vibrating platinum microelectrode to simultaneous measurements of oxygen utilization and cytochrome absorption, as illustrated by Fig. 12. The electrode, inserted into the optical cuvette at the extreme right, is vibrated by a Brown converter at 60 cps and the whole assembly is enclosed between cooling blocks to maintain the constancy of temperature. The function of the vibrating mirror of the double beam spectrometer is indicated by Fig. 16 (FA-30). (Reprinted by permission of Academic Press, Inc., New York, N. Y., [32].)

In the input circuit to the vibrating platinum electrode (Fig. 14), the vibrating reed which moves the electrode also actuates the contacts which modulate the electrode current and allow it to be passed through the input transformer. The reed may be polarized at the proper potential (−0.6 volts). In addition, a compensating voltage may be introduced through the 10 megohm resistor and potentiometer. The ac waveform from the secondary winding of the transformer may be amplified by any conventional ac amplifier and demodulated by an appropriate synchronous demodulator, for example, a switch identical to that used for vibrating electrode. The output can operate a recorder.

Measurements of Cytochromes

A single-beam spectrophotometer (such as the Beckman DU) has proved unsuitable for measuring the very small absorbancy changes of the cytochromes in suspensions of cells or mitochondria because of the large nonspecific light-scattering changes that accompany the specific absorbancy changes. Fig. 15 illustrates the problem of optical measurements in a plant tissue in which the removal of oxygen causes a nonspecific shift corresponding to the separation of two traces, while at the same time, absorption bands appear (520–580 mμ) [36]. The instrumentation problem is to measure these absorption bands and reject the nonspecific light-scattering changes. It is apparent that measurements with a single-beam spectrophotometer set at 560 mμ would give a poor representation of the actual spectral shift. We have, therefore, taken advantage of the relatively small change of light-scattering occurring at two adjacent wavelengths and have devised a double-beam spectrophotometer which records the difference of absorbancy ($\Delta D_1 - \Delta D_2$) at λ_1 and λ_2 [37], [38]. The optical arrangement of light-source and monochromators illustrated by Fig. 16 is especially suitable for observations in the range 250–650 mμ. Tungsten or hydrogen light sources illuminate the two monochromators (Bausch & Lomb, 250 mm focus) via the half reflecting mirror. The slit is focused upon the vibrating mirror, which alternately selects the light from one of the two monochromators and flashes it upon the sample (see also Fig. 13). The grating is imaged upon the sample and accurate superposition of the images is obtained. The scattered and transmitted light from the sample is received by a closely placed end-on photomultiplier which gathers as large a solid angle of light as is feasible for the particular experimental conditions. When the light intensities from the two monochromators are appropriate, there is no ac signal from the photomultiplier, but when a small change in the difference of transmission along the two optical paths occurs, there is an ac error signal which is amplified, demodulated, and recorded ("differential output" [Fig. 17]). To calibrate the system electrically, the

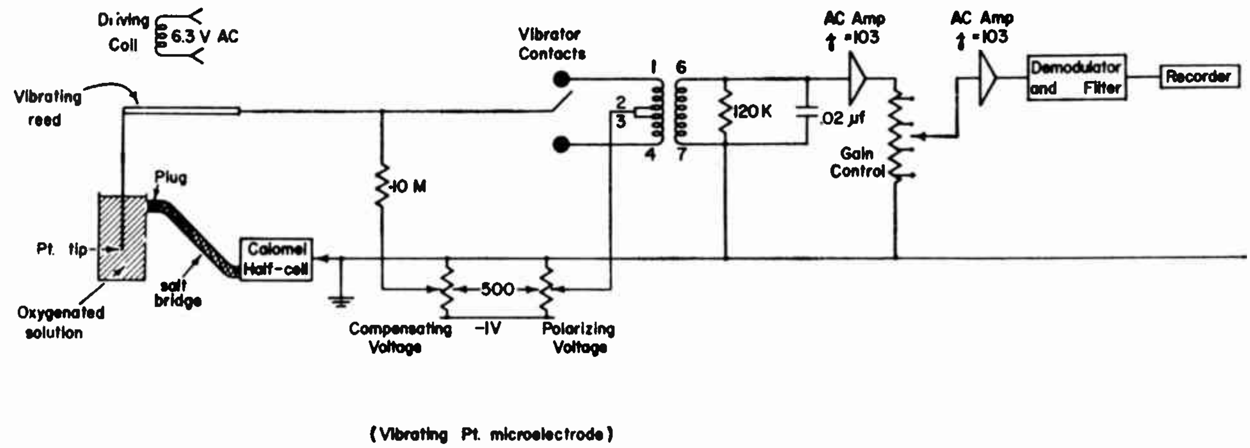


Fig. 14—Input circuit for vibrating electrode in which the reed of a Brown Instrument Co. converter used for vibrating the electrode is also used to chop the electrode current and, thereby, allow transformer coupling of the electrode signal and subsequent ac amplification and demodulation. The electrode input circuit can be isolated from ground potential due to the transformer coupling (MC-11).

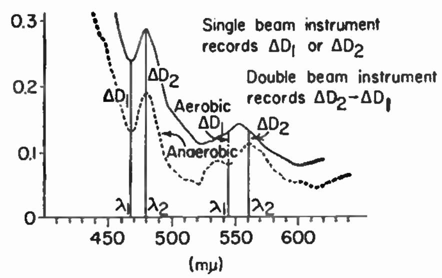


Fig. 15—An illustration of the experimental difficulties involved in absorbancy measurements with a single-beam technique with certain types of biological materials (bundles of wheat root). The two spectra taken from the work of Lundegardh show that the transition from aerobiosis to anaerobiosis (oxygen to nitrogen) causes a large displacement of the absorbancy records due to nonspecific light scattering changes. In addition, there are certain specific changes in the region of 550 mμ. Experimental difficulties of this type suggest the use of a double-beam technique rather than a single-beam technique (0-118).

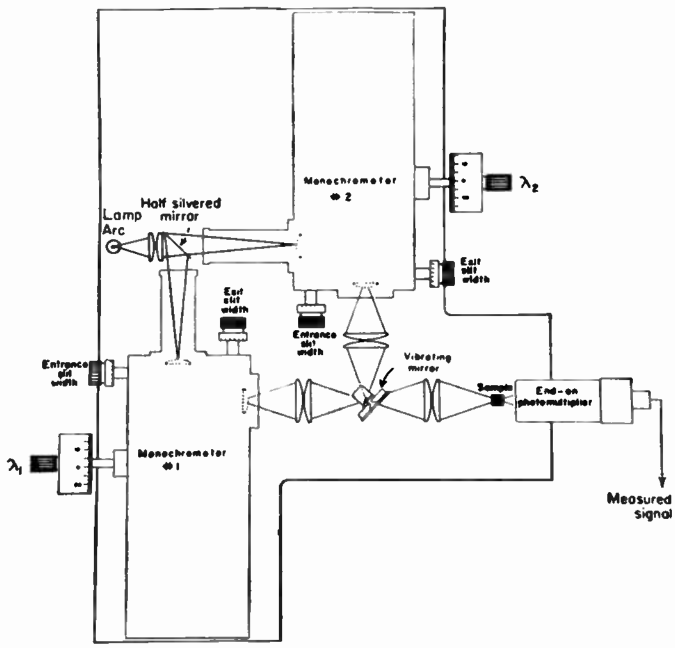


Fig. 16—The optical system of a double-beam spectrophotometric technique. The monochromators illustrated are Bausch & Lomb 250 mm focus. The vibrating mirror is founded on the arm of a Brown Instrument Co. converter (MD-34).

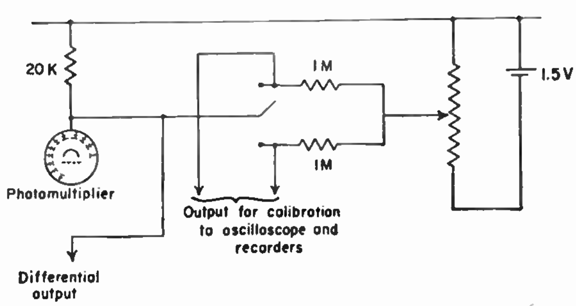


Fig. 17—Electrical input circuit for the double-beam spectrophotometer which allows "differential output" in which the amplitude of the ac signal indicates any change of absorbancy at the two wavelengths. In addition, the contacts on the vibrating mirror allow "time-sharing" measurement of absorbancy changes at the two wavelengths as well as a calibration of the photocurrent used for the measurement of the differential output (MD-52).

amplitude of the two light pulses is referred by intermittent connections to a standard dc voltage with switch contacts associated with the vibrating mirror (see Fig. 17). When the ac component of the waveform obtained at either of the two outputs denoted "output for calibration" is zero, then the amplitude of the light pulse has been accurately measured. It should be noted that single-beam operation at the two wave lengths can be obtained by simply recording the amplitudes of the ac error signal at the points labeled "output for calibration." In addition to measurement of the amplitude of the light flashes, the voltage sensitivity of the ac amplifying circuit for the difference measurement is calibrated by the response to an incremental change of the calibrating voltage. In this way the system can be calibrated electrically in terms of percentage transmission. Since small changes of transmission are measured, the optical density or absorbancy changes are linearly recorded. A very large nonspecific change of light transmission will cause an alteration of sensitivity, but this can readily be corrected by a control circuit for the dynode voltage operated by the average value of the photocurrent.

A photograph of a portion of the double-beam spectrophotometer using Bausch & Lomb monochromators is given in Fig. 13 (see also Fig. 18). The apparatus is provided with appropriate cuvettes of various optical paths and with control for the temperature of the solutions.

Measurement of Rapid Reactions

The usual method of initiating spectroscopic changes by rapidly stirring a reaction-initiating chemical into the cuvette with a stirring rod gives mixing times of ~0.5 second. To extend the time range beyond this limit, rapid flow devices operating on the principal of Hartridge and Roughton [39]–[41] have been developed specifically for timing rapid reactions in suspensions of living cells [42]. A particular form of this apparatus developed for use with the double-beam spectrophotommeter is indicated in Fig. 18. Here the regenerative flow apparatus is attached to the double-beam spectrophotometer using Bausch & Lomb monochromators. The flow velocity obtainable with this particular form of the apparatus allows a time after mixing of ~10 msec. An additional advantage of the regenerative flow apparatus is the increased signal-to-noise ratio that can be obtained, since the flow velocity can be maintained constant for an interval of several seconds, whereas the time after mixing is about 10 msec. In effect, a bandwidth reduction of over 100-fold is obtained, with a consequent 10-fold increase of signal-to-noise ratio.

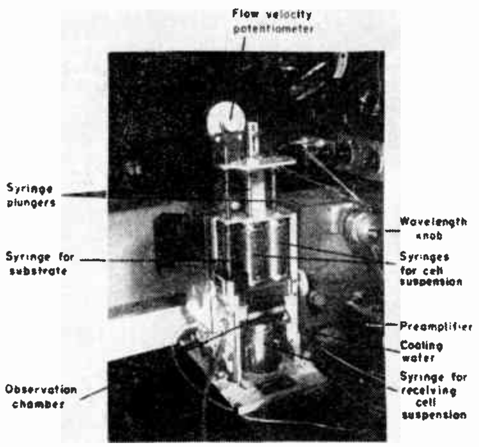


Fig. 18—A rapid flow apparatus attached to the double-beam spectrophotometer suitable for increasing the time range to about 10 msec. The syringes for driving the reactions are indicated as is the observation chamber. The photomultiplier has been removed for this photograph (FA-28). (Reprinted by permission of Academic Press, Inc., New York, N. Y. [32].)

Measurements in Muscle

Sensitive optical measurements of rapid and specific absorbancy changes in living tissues (for example, those associated with contracting muscle) are also highly desirable. In collaboration with Dr. C. M. Connelly a muscle holder (Fig. 19) has been developed for the study of such reactions following the contraction of a frog's sartorius muscle [38], [43]. The perfused muscle is held between two perforated lucite plates, past which flow solutions containing oxygen and various other reagents. The ends of the muscle are in contact with two silver electrodes which provide electrical stimulation. The muscle is illuminated by the double-beam spectrophotometer and the wavelengths used are sufficiently close together so that light-scattering changes which occur when the muscle is stimulated do not obscure the specific effects. Fig. 20 illustrates the absorbancy changes that occur in response to muscle contraction. It is seen that the artifact occurring at the cessation of stimulation is negligible compared with the specific absorbancy change. A comparison of the spectroscopic effect with the tension record is given in Fig. 9. More recently, Ramírez [39], working in this laboratory, has applied a similar technique to the study of rhythmic contraction in cardiac muscle (Fig. 21) of *Bufo marinus*. Here double-beam recordings (lower trace) in response to the initiation of cardiac contraction show very little of the nonspecific effects associated with each contraction. Simultaneous measurements with the single-beam technique afforded by the circuit connection illustrated in Fig. 17 show large light-scattering response to each contraction on a scale of 1/7 that of the lower trace. This apparatus has also been used for studies of a variety of living materials, such as brain and liver tissues.

Split-Beam Spectrophotometer

The double-beam apparatus is specifically adapted for recording, as a function of time, a change of absorbancy in a single sample at a pair of predetermined wavelengths. However, it is often desirable to record, as a function of wavelength, the difference of absorbancy between two samples that are identical in all respects except in the state of oxidation of the electron transfer system. To fill this need, we have developed a wavelength scanning spectrophotometer that has the advantages of rapid recording and high sensitivity in spite of the use of turbid solutions [38], [45], [32]. This instrument is called a split-beam spectrophotometer because

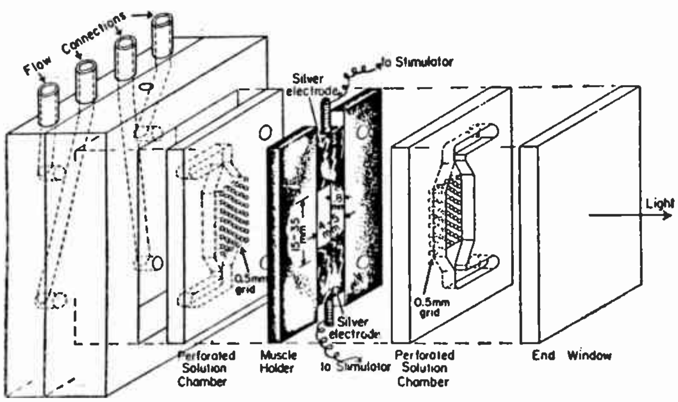


Fig. 19—A muscle holder suitable for spectroscopic observations of changes in the oxidation-reduction level following stimulation of excised muscle (0–54). (Courtesy Dr. C. M. Connelly.) (Reprinted by permission of *Science* [38].)

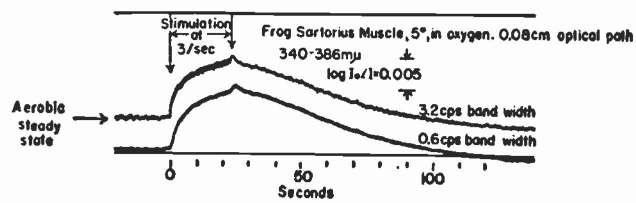


Fig. 20—A double-beam spectrophotometric record of the time course of absorbancy changes during and after the stimulation of frog sartorius muscle held in the chamber illustrated by Fig. 19. The absorbancy changes are recorded at two band widths (3.2 and 0.6 cps). The smallness of the optical artifact recorded during stimulation is evidence of the effectiveness of the double-beam technique for this type of measurement (Expt. 303). (Reprinted by mission of *Science* [38].)

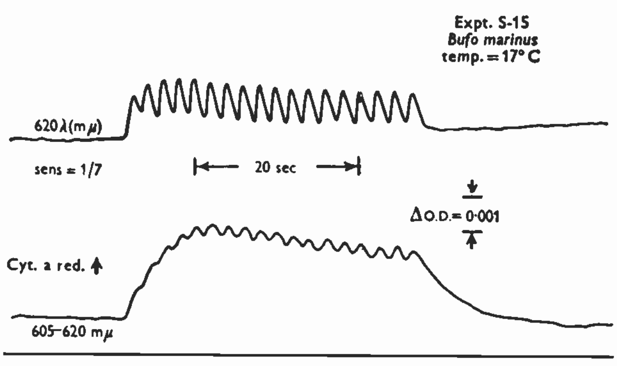


Fig. 21—Single and double-beam recording of absorbancy changes during and after contraction of cardiac muscle of *Bufo marinus*. The top trace is a single-end recording obtained by the "calibration" connection of Fig. 17 and the bottom trace is a differential recording. The latter shows as an upward deflection the reduction of cytochrome *a*. The rejection of the light-scattering artifacts by this technique is indicated by a comparison of the deflection of the two traces in response to the contraction, taking into account that sensitivity for the top trace is reduced 7-fold over that of the bottom trace (J.R.-3). (Courtesy Dr. J. Ramírez.) (Reprinted by permission of the *Journal of Physiology* [44].)

a single light beam passes through two solutions. The light may be chopped, either by a vibrating mirror or by an oscillating mirror. The optical path is illustrated by Fig. 22 where a light from the monochromator is reflected by a front surface mirror onto the oscillating mirror which is cam-operated by a synchronous motor. The light beam then sweeps past the two cuvettes to give a sequence: reference material, dark, measured material, dark. Suitable switching circuits operated by the cam compare the photocurrents corresponding to the reference flashes with a fixed voltage, as illustrated by Fig. 17. Thus the magnitude of the light pulse through the reference material can be held at a standard value by a feedback circuit controlling the dynode voltage of the photomultiplier. In this way, the compensation of the multiplier gain for changes of light source emissivity and sample transmission with wavelength can be compensated. The transmission of the sample is measured by the difference between the reference and measure pulses. A conversion of this quantity to logarithms by means of a segmented diode characteristic has been described. The sensitivity of this record is sufficient that detection of absorption bands in frozen samples of biological material is readily obtained; an example of the recording of the difference between an oxidized and reduced sample of a beef heart preparation already has been indicated in Fig. 4(b), which illustrates the excellent discrimination of the cytochromes that can be obtained with this technique.

A Highly Sensitive Microspectrophotometer

High sensitivities in optical measurement can be achieved on a macro scale by measurements of a small change of absorbancy in a large light intensity, as is done in the double-beam spectrophotometric technique above, where sensitivities of 10^{-10} mole of respiratory enzyme are obtained. However, microspectrophotometry affords an even more effective use of optical

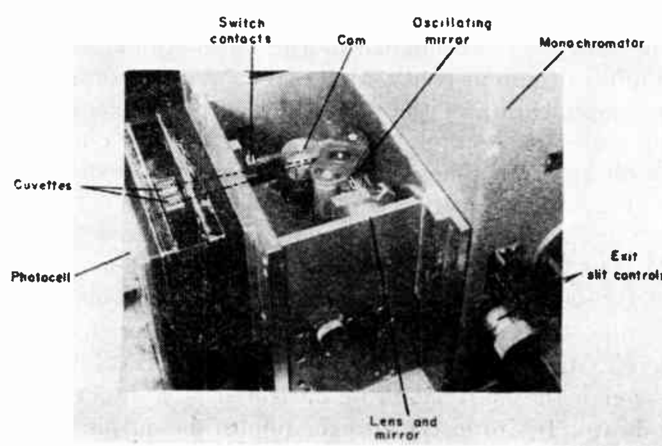


Fig. 22—Optical path of split-beam recorder showing monochromator, oscillating mirror, cam-operated switch contacts, and the two cuvettes. The photomultiplier is shown on the extreme left (FA-32). (Reprinted by permission of Academic Press, Inc., New York, N. Y. [32].)

geometry for detecting small amounts of material. In fact, this method has been effectively used by Caspersson [46] to evaluate the amounts of protein and nucleic acids which are present in high concentration in the cell. More recently, Thorell [47] has used this method in the measurement of hemoglobin formation. However, the application of this method to the measurement of the substances in the single living cell which are present in "enzyme" concentrations requires a high signal-to-noise ratio in measurements where a marginal photocurrent is obtained through the microscope optics. It has been possible recently to achieve the required spectrophotometric sensitivity to record automatically and directly the spectra of respiratory enzymes in groups of mitochondria in a single living cell [20], [21], [48]. A microfluorometer of comparable sensitivity has been developed for localizing DPNH (see Fig. 4(a)) in the single cell [49], [50], [51].

This microspectrophotometer has a 1.5 μ aperture and is capable of measuring, in the spectral region 340–600 mμ, absorbancy changes of approximately 1×10^{-4} in optical density at a time constant of ~2–3 seconds. The sensitivity of the apparatus is sufficient that respiratory enzymes, in a portion of the living cell, in the amount as small as 1×10^{-18} mole, can be detected at a signal-to-noise ratio of 50 to 1. Thus the noise level corresponds to 2×10^{-20} mole, or roughly 10^4 molecules. This detectibility considerably exceeds that of contemporary techniques such as electron spin resonance.

In principle, the method represents an application of the dynode feedback circuit of Picard [51a] and a refinement of the circuits for the split-beam spectrophotometer developed by Yang [45] and Åkerman [52]. Significant advantages of this type of wavelength scanning spectrophotometer are the automatic electronic compensation of changes of 1) emissivity of the lamp and sensitivity of the photocell, and 2) transmission of the reference material over an extremely wide dynamic range. Two additional advantages are 3) the reference signal is intermittently referred to the phototube dark current so that a change of the latter value will not diminish the sensitivity, and 4) the circuit incorporates ac amplification of the phototube output waveform so that grid current of the amplifying system does not cause an error. The result is a circuit of an extremely high inherent sensitivity and stability.

The time sequence of light pulses is illustrated by Fig. 23. The input waveforms shown on the top line consist of a series of pulses corresponding to the darkened phototube, light transmission through the reference material, the darkened phototube, light transmission through the material to be measured. The function of Vibrator I is to insert into the phototube output a signal equal to the difference between the dark and the reference pulses. When the phototube is dark, the input waveform is connected to the amplifier in series with the reference voltage. When the phototube is illuminated, the input waveform is connected to the amplifier without the reference voltage. For the remainder of the duty cycle, the switch contacts are opened so that the transients caused by the transition from the light to dark do not saturate the amplifier and the control circuits. A diagram of the circuit used in these experiments is shown in Fig. 24. The output of the reference voltage insertion circuit is connected to a packaged amplifier, type II [53]. The output of the upper type A18 transformer is used to operate a synchronous demodulator for the "reference minus measure" signal. Switch IIb which removes the "measure" signal from the control circuit and operates the lower type A18 transformer to activate synchronous demodulator IIIA to give the difference between the "reference" and the "dark" signals. This signal is remodulated and demodulated by chopper IIIb to give further drift-free gain. The output is fed into dc amplifier, type 6AU6, and is resistance coupled from point A to the grid of a 2C53 triode which controls the dynode voltage across the photomultiplier, type 1P29. In this instrument, it is unnecessary to convert from the absorbancy difference to logarithms, since the apparatus is usually used on a full scale absorbancy change of less than one per cent.

The performance of the apparatus without biological material is indicated by Fig. 25 which records the noise level as a function of time at 420 mμ. The peak amplitude of the noise over the 30-second interval is less than 5×10^{-4} optical density units, the average being 1×10^{-4}. The rise time of the output (10 to 90 per cent) is seven seconds, and the primary photocurrent is about 2×10^{-12} amperes. Thus, the shot noise level corresponds to approximately 4×10^{-16} amperes. The scanning time of the instrument has been set at 50 mμ per minute as a compromise between errors due to cell movement associated with too slow a recording speed and errors in recording the location of the absorption peaks. With the 2.5-second time constants there is an approximate error of 5 mμ in location of the peaks, which is satisfactory for the application to measurements of cytochromes whose absorption peaks are approximately 10 mμ apart in the region 410–450 mμ.

An experimental result obtained with this apparatus has been presented above, and the results clearly show that it is possible to record clearly the spectra of the respiratory enzymes localized in a small portion of the living cell.

Noise in Light Sources for Precise Spectrophotometry

Investigations carried on some years ago show that convection noises in the gas-filled lamp, or irregularities in the temperature along the tungsten wire, can considerably diminish the signal-to-noise ratio obtainable for precise spectrophotometry [54]. A typical result is shown in Fig. 26 where the lower curve shows that operation of the lamp at higher voltages gives a lower signal-to-noise ratio than that at the lower voltages and the result differs considerably from that expected theoretically (upper curve). This fault varies considerably

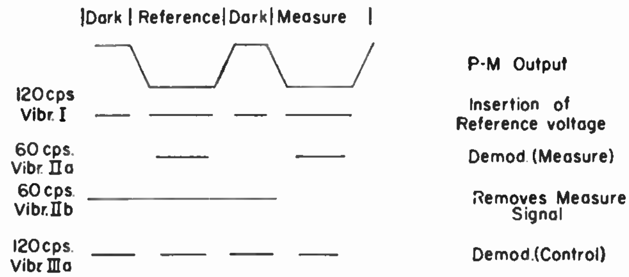


Fig. 23—Time sequence of switch circuits employed in the split-beam recording microspectrophotometer. The vibrating devices labeled I, II, and III are indicated on the circuit diagram of Fig. 24 (MD-81).

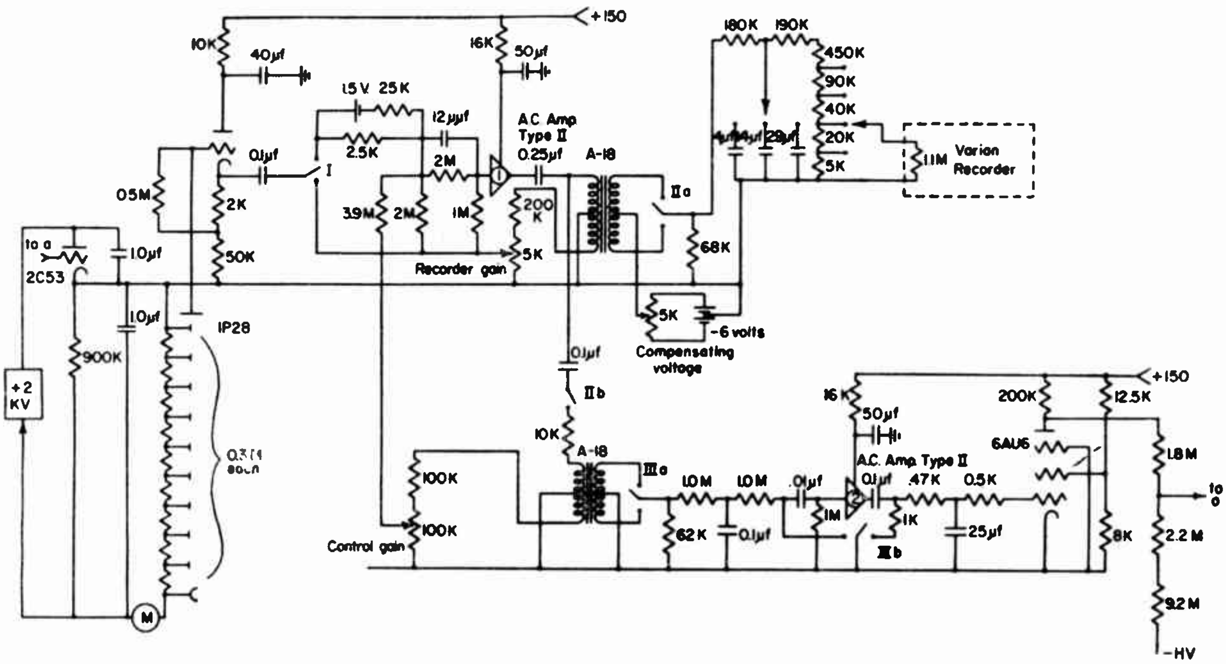


Fig. 24—Circuit connections for recording and control channels of the microspectrophotometer. The time sequence of switches I, II, and III is indicated in Fig. 23. The connection from the voltage divider attached to the plate type 6AU6 to the grid 2C53 is omitted (point A) (MEC-100a). (Reprinted by permission of the American Institute of Physics [21].)

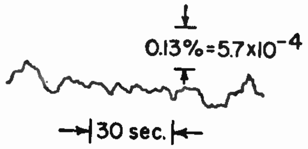


Fig. 25—A recording of the noise level of the microspectrophotometer operating at a primary photocurrent of approximately 10^{12} amp and a rise time (10 to 90 per cent) of seven seconds. The average value of the noise is approximately 1×10^{-4} in optical density (RS-2).

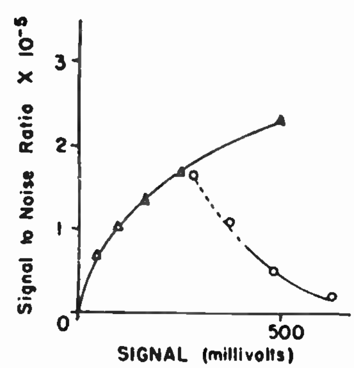


Fig. 26—Effect of lamp voltage on signal-to-noise ratio obtainable in sensitive spectrophotometry. The upper trace is a theoretical curve and the lower trace represents the experimentally observed data. This record was obtained with the 8.5 amp "sound-track excitor lamp," using a gas-filled coil. It is seen that the departure from the theoretical signal-to-noise ratio occurs at the higher lamp voltage (signal current) (Expt. 583d).

with the type of lamp; the so-called "sound track" exciter lamp has been found to be reasonably quiet while others, particularly some types designed for microscope illumination, have been found to be quite noisy. It is highly desirable to stabilize the light intensity from the lamp with negative feedback circuits [55], and a number of designs have been published.

Computer Representations of Biochemical Reactions

Model Systems

Model representations of the living processes have been useful in understanding their operation [56]–[59]. There are, however, several levels at which the model may be constructed. It may be an analog of the processes themselves; for example, in a chemical system, simpler compounds may be used which may simulate, perhaps on a slower time scale, the more complex biological systems. In other cases, an hydraulic analog may be used to represent some aspects of the chemical system [58]. Similarly, electrical networks may be used to simulate properties of the biological systems, either in terms of resistance networks where no dyanmic responses are obtained, or in terms of those containing both resistance and capacitance elements. Usually voltages are the analog of the chemical concentration, and, in the hydraulic system, liquid volume may represent chemical concentration. A more sophisticated analog is described by MacNichol [60], where a degree of simulation of the properties of the living system is rather good. However, in neither case is the model based upon the physical laws which are an inherent property of the biochemical-biological system.

Analog Computers

A more general approach to computer representation of the biological system is one in which the basic equations representing the living system afford the basis for the computer system. When the system is based upon mathematical equations derived from physical laws representing the system, it assumes considerably greater significance than in a model which may operate on an entirely different natural law. Thus we feel that this approach to computer representation is of fundamental significance and may possibly be applied to other systems.

In these studies, the law of mass action forms the basic equation [56]; chemical rate is proportional to the product of the concentrations of the reactants. For a multicomponent system, the net rate of change of a given substance is proportional to the sum of the rates of the reactions leading to the formation minus those leading to the decomposition of the substance. For example, referring to (1) and (2)

$$\frac{d\mathrm{Fe}^{++}}{dt} = k_1[\mathrm{Fe}^{+++}][e] - k_2[\mathrm{Fe}^{++}][-e]. \quad (3)$$

Eq. (3) may be rewritten for a pair of iron proteins of the respiratory chain. The letter and subscript represent the kind of cytochrome and the + signs its valency.

$$a^{++} + a_3^{+++} \xrightarrow{k_2} a_3^{++} + a''' \quad (4)$$

$$a_3^{++} + O_2 \xrightarrow{k_1} a_3^{+++} + O_2^{-}. \quad (5)$$

The differential equation is then

$$da_3^{++}/dt = k_1 a^{++} a_3^{+++} - k_1[O_2][a_3^{++}]. \quad (6)$$

There are similar equations for each of the components of the sequence of Fig. 4, which is simplified as follows:

$$\mathrm{DPN} \xrightarrow{k_4} c \xrightarrow{k_3} a \xrightarrow{k_2} a_3 \xrightarrow{k_1} O_2. \quad (7)$$

It is seen that such differential equations are nonlinear in the sense that they involve the product terms of two variable concentrations. Thus, a general solution cannot be obtained; the analog computer is essential for such studies, and such a computer has been used at the Johnson Foundation for about nine years.

The computer connections which are suitable for solving these equations are given in the schematic diagram (Fig. 27). It is seen that the terms on the right-hand side of the differential equations [*e.g.*, (6)] are obtained separately for each component and are integrated to give each of the variables.

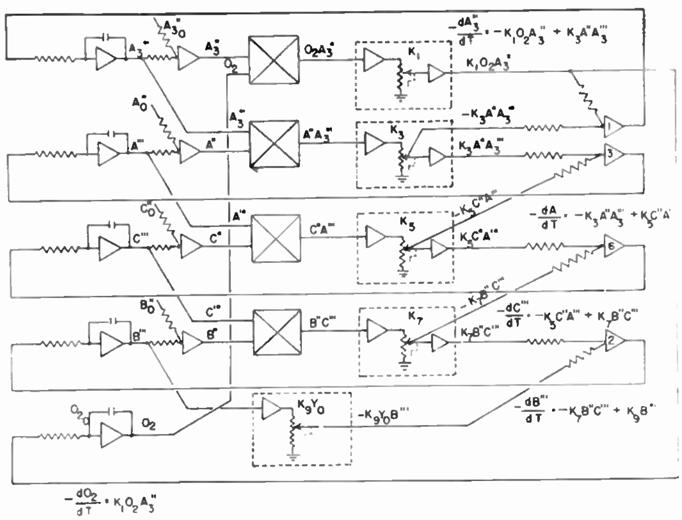


Fig. 27—A block diagram of analog computer interconnections required for a solution of nine differential equations representing a sequence similar to that of (7), but modified as follows:

$$Y \rightarrow B \rightarrow C \rightarrow A \rightarrow A_3 \rightarrow O_2.$$

The differential equations corresponding to the time course of the concentrations of the components are similar to those of (6) and the terms of these equations are summed in the adders 1, 2, 3, and 6 on the right side of the diagram. The integration of these sums on the left of the diagram gives the values of the variables which are then connected to the multipliers (square). The appropriate parameters are inserted mechanically ($k_{1,3,5,7,9}$) and the appropriate terms are summed to complete the loop (AC-56).

The results of the analog computer representation of the electron transfer system are given in Fig. 28. The speed with which the system responds to a step function of oxygen concentration is given on slow and fast time scales. The system starts with all components in the ferrous state (Fe^{++}) and a step function of 40 μM oxygen is applied. The solution represents, then, the concentrations of the oxidized forms which have rise times in the sequence,

$$a_3 < a < c < \text{DPN}$$

suggesting that the time sequence of the oxidation reaction represents the chemical sequence. Thus, an experimental determination of this time sequence by means of rapid measurements on living cells with the regenerative flow apparatus and the double-beam spectrophotometer permits a direct comparison with the computer data.

The upper portion of Fig. 28 indicates, on a slower time scale, the time course concentration of the oxidized forms of the electron transfer components. After an abrupt rise, a steady state is obtained during which the added substrate (x_0) is rapidly utilized. As its concentration decreases, a departure from the steady-state values is observed first for a_3, second for a, third for c, and fourth for DPN. This is a general property of such systems, namely, that the component nearer the substrate will show an earlier departure from its steady-state value. If we measure the time interval required for the intermediate to rise and fall from half-maximal to its maximal value, we observe from the analog computer solutions that there is an ordering property of the $t_{1/2\text{off}}$ values

$$a_3 < a < c < \text{DPN}.$$

In fact, Higgins [62] has been able to derive a theorem for such systems verifying the observation based on analog computer data. In general, it has been found fruitful to derive from mathematical solutions of the chemical equations empirical relationships between the variables which allow a more incisive study of the experimental system. A general formula relating a reaction velocity constant to the amplitude of the kinetic curve and the value of its $t_{1/2\text{off}}$ was discovered in early mechanical differential analyzer solutions of chemical equations for enzyme systems [56].

A Crossover Theorem

It is of considerable interest to determine whether the energy-conserving reactions leading to the formation of ATP by the chain of cytochromes diagrammed in Fig. 4(a) involve the interaction of particular pairs of electron transfer components and if so, which pairs. Chemical methods for studying this require a disruption of the structure of the mitochondria with possible dislocation of the reaction sequence. Thus it is of the greatest importance to devise a physical test to identify pairs of components involved in the energy-conserving reactions without any disturbance of the chemical balance of the system. It is known that the rate of electron transfer through the system is decreased if ADP or P_i [Fig. 4(a)] is removed. Furthermore, when this decrease occurs, some carriers become more reduced and others more oxidized. The place in the cytochrome chain at which the change from a reduction to an oxidation occurs is called a "crossover point." Under various conditions, the crossover point can be made to move along the

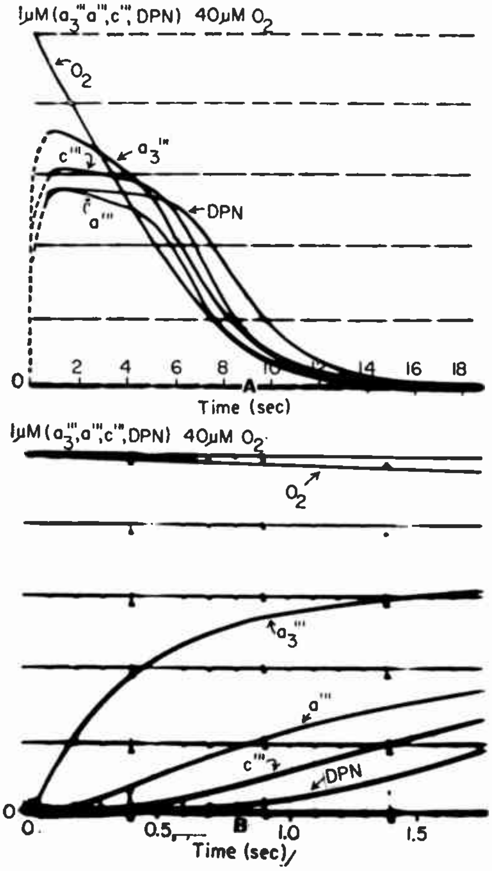


Fig. 28—Analog computer solution of transient and steady-state portions of the sequence of the reactions represented by (7). The program of the analog computer is similar to that illustrated by Fig. 27. The top trace shows computer solutions on a slow time scale, while the bottom trace shows in detail the initial portions of the transient response. (Solution obtained with Johnson Foundation electric analog computer) (AC-63). (Reprinted by permission of the *Journal of Biological Chemistry* [61].)

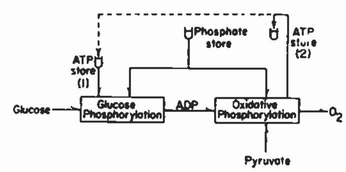


Fig. 29—A flow chart of a simplified metabolic control sequence for the utilization of glucose and oxygen by the cell. The addition of glucose, indicated experimentally by Fig. 12, depletes ATP in store I with the formation of ADP. ADP is then the control signal for the activation of oxygen utilization. Feedback from the ADP formed in oxidative phosphorylation from store II to store I (via dotted line) regulates the use of glucose (MD-72).

chain: a decrease of electron transfer rate causes it to move towards the substrate (glutamate, etc.) end, and vice versa. However, the movement is quantized and only three crossover points in a chain containing a total of eight such possibilities have so far been unambiguously identified [62a].

A consideration of the chain as a series electrical network immediately leads to the derivation of a fault location theorem [59], where instead of a single fault [62b], multiple faults are possible [62c]. Since, however, the resistive network fails to simulate the dynamic properties of the enzyme chain, we have derived the crossover theorem for the reaction mechanism indicated by (4)–(6) on the basis of physical arguments [61] of analog computer studies [59], and of inequalities [62d]. The theorem can be briefly stated for the chemical system in which a reduction is denoted by a minus sign and an oxidation by a plus sign.

For an effect causing a decreased rate of electron transfer,

1) an interaction site lies between a + to − change in the sequence from substrate to oxygen;
2) components between oxygen and the first site will always show + changes, and those between substrate and the last site will always show − changes;
3) a crossover point near the oxygen end of the chain can be shifted to the next site of interaction by a decrease of activity in the oxidase portion of the chain, and vice versa;
4) *a* + to − change (reversed crossover) does not identify an interaction site.

The theorem applies equally well to fault location in a linear resistive network. Its application to branched networks would be of some interest.

Digital Computers

Our experimental studies have not been restricted to the electron transfer system of the mitochondria but include studies of intact cells and tissues, as illustrated by Figs. 11 and 12, and 21, in which there is interaction between a function (such as muscular contraction), or a metabolic activity (such as the utilization of glucose), and the electron transfer system. The physical methods described above detect distinctive spectroscopic responses of the electron transfer system to those interactions. The digital computer provides an appropriate means for the representation of these more complex metabolic regulations. Such a representation is based upon the law of mass action for the chemical properties of the system and includes the compartments of the cell for particular chemical species. In addition, with small changes, processes governed by the rate of diffusion can be included [63].

Metabolic Control in Cancer Cells

The problem of metabolic control in cancer cells is of particular interest since current hypotheses on the cause of cancer [64] suggest that special control mechanisms lead to a greater formation of lactic acid in some types of tumor cells than is observed in the corresponding "normal" type. Whether or not this hypothesis is true [65], the cancer cell, particularly that of the ascites tumor type, affords an excellent example for a study of the way in which metabolic control can be exerted by the biochemical processes of cell metabolism. Questions of particular interest are:

1) What is the mechanism by which the cell diminishes its expenditure of food stuffs when its energy needs are satisfied?

2) How does the Pasteur reaction [24] operate to suppress glucose utilization in the presence of air where more efficient processes are operative and less glucose is needed.

3) What is the nature of the Crabtree effect [22] in which the inverse of Pasteur's reaction is observed?

In addition, several more recently identified metabolic regulations that occur in the first few minutes after adding glucose to ascites tumor cells demand an explanation?

A simple flow chart indicating the interactions that may occur in these cells is given in Fig. 29. For our purposes it is not necessary to go into the details of the actual chemical reactions but only to observe that the addition of glucose to the cell causes immediate phosphorylation of glucose by ATP with the production of ADP. The latter activates the electron transfer system (oxidative phosphorylation) which causes an acceleration of oxygen utilization. Electron transfer provides energy to rephosphorylate ADP to ATP (cf. Fig. 4); ATP in store II can ultimately become available as ATP in store I. In this case, the utilization of glucose can proceed through a simple feedback path [66] indicate by the diagram.[1] The interactions of Fig. 29 are far from complete and a more nearly complete representation is provided by the complicated flow chart of Fig. 30. Here we have added to the functions of glucose phosphorylation and oxidative phosphorylation an alternative method for glycolytic phosphorylations of ADP of Embden and Myerhof, which serves in the absence of oxygen to conserve energy in a very different manner than in the electron transfer system. Such an energy conservation system operating in the absence of air is necessary in muscles, particularly those involving intense energy expenditures of brief duration. The system is, interestingly enough, highly reactive in

[1] It should be noted that the term feedback as applied to biochemical systems, is becoming more and more widespread and a comment on its propriety is desirable. It is obvious that feedback in the chemical system is not directly comparable to that obtained with negative feedback of electronic circuitry. None of the obvious features of gain stabilization or more favorable dynamic response and impedance transformation, are obtained. In fact, the essential feature of subtracting a portion of output from input does not find its analog in the chemical system; there a steady state based on the difference of supply and demand (*i.e.*, of rates) is established (a more complete discussion of this point is presented elsewhere [66]).

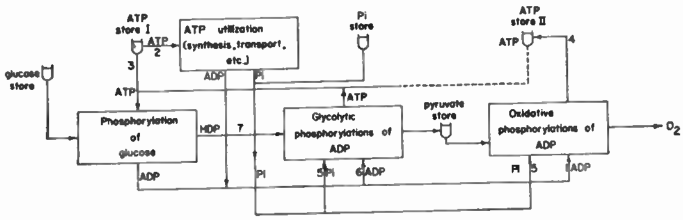


Fig. 30—Detailed flow chart for chemical interactions involved in metabolic control. Here, additional functions are interposed between glucose phosphorylation (glycolytic phosphorylations) and the use of ATP for cell function is independently specified (MD-65). (Reprinted by permission of *Science* [23].)

tumor cells. A second function of the flow chart is an enlargement of energy expenditure of the cell; for example, for synthesis of new material, for the transport of essential substances into the cell, for contractions in muscle cells, for electrical phenomena in nerve cells, etc. Basically, however, the flow chart is the same; that is, the metabolism of glucose requires energy in the form of ATP and thus causes the consequent production of ADP which, in this case, can be restored not only by oxidative phosphorylations (Path 5) but also by glycolytic phosphorylations (Path 6).

A compartmentation of the cell is indicated in the output of these two types of phosphorylation, that from oxidative phosphorylation (Path 4) going into ATP store II which is not immediately available to replenish ATP store I, and that from glycolytic phosphorylation being immediately available for this purpose. Another type of compartmentation is required, namely, that a chemical constituent, diphosphopyridine nucleotide, which occurs in oxidative and glycolytic phosphorylations, is not equilibrated between the two compartments.

Chemical Equations

Although it is beyond the scope of this discussion to analyze in detail the chemical equations representing the reactions that occur in the various functional compartments of the diagrams of Figs. 29 and 30, it is awkward to present the results of the digital computer application to this metabolic control problem without indicating the quantities involved. Table I, therefore, briefly presents the 20 chemical equations used so far in this representation and Table II gives the names of the chemicals used in the computer solution and the graphical procedure. Five equations are associated with the function of glucose phosphorylation, seven with glycolytic phosphorylations of ADP, five with oxidative phosphorylations of ADP, and three with ADP utilization and transfer. Three types of equations occur. In the first [(5) and (20)], the velocity is proportional to the concentration of a single substance. In the second [*e.g.*, (1) to (4)], the velocity is proportional to the product of the concentrations of two substances. In the third [*e.g.*, (14)], the velocity is proportional to the product of the concentrations of three substances. Whereas in the living system enzymes catalyse all these reactions, we have taken the liberty of omitting the enzyme in some of the reactions to provide a simpler solution [*e.g.*, (6) to (17)]. In (1) to (4), on the other hand, the enzyme is explicity represented (ENZ, ETZ). It is desirable, of course, to elaborate the program to include the appropriate enzyme in each step. (*Note:* The expenditure of a substance appearing on the left of the equation is indicated on the right of the equation by an accompanying minus sign.)

The velocity constants are written above the arrows. The initial concentration of each substance is written above and to the left of its symbol; the figure below this represents the full-scale value of the particular concentration to facilitate reading the concentrations from the graphs. The velocity constants are chosen so that the computer solutions occur at rates similar to those observed in the living cell. Thus, the numbers along the bottom of the graph (cf. Fig. 31) can be taken to indicate seconds.

Method of Solution

The differential equations representing the chemical processes (Table I) are solved by the usual numerical integration techniques. However, to achieve maximum speed in obtaining the solutions with Univac I, several additional features have been incorporated. Floating decimal arithmetic is avoided by representing all concentrations and rate constants in the range 0–0.999. This is achieved by representing the concentration of each chemical as the ratio of the actual concentration to its maximal concentration. Normalization is carried out automatically by the computer. The ratio of the maximal flux (product of concentration and reaction velocity constant) of the fastest reaction to the flux of the slowest reaction must not exceed 10^9 if an accuracy of three figures is to be maintained in view of the 12-digit limit of Univac I.

The numerical integration technique [67] can be represented by a Taylor extension in the form

$$y(t + \Delta t) = y(t) + y'(t)\Delta t + \epsilon, \tag{8}$$

where y represents the concentration of one of the chemicals, t the time, Δt the increment of time, and ϵ the error.

The error is given by

$$\epsilon = \frac{1}{2!} y''(t)\Delta t^2 + \frac{1}{3!} y'''(t)\Delta t_3 + \underset{\text{terms}}{\overset{\text{higher}}{\text{order}}} \frac{1}{2} y''(t)\Delta t^2 \tag{9}$$

where primes indicate derivatives. The first derivative obtained from the differential equation involves the calculation of $y(t+\Delta t)$. The second derivatives are obtained from the two first derivatives at times Δt apart and allow an approximate calculation of the

TABLE I
CHEMICAL EQUATIONS FOR METABOLIC CONTROL SEQUENCE (DC-46)

Phosphorylation of glucose	1)	${}^{0\to10^{-3}}_{10^{-3}}\mathrm{GLU} + {}^{10^{-6}}_{1.02\times10^{-6}}\mathrm{ENZ} \xrightarrow{1\times10^{10}} \mathrm{ENG} - \mathrm{ENZ} - \mathrm{GLU}$
	2)	${}^{0}_{10^{-5}}\mathrm{ENG} + {}^{5\times10^{-4}}_{1.5\times10^{-3}}\mathrm{ITP} \xrightarrow{1\times10^{10}} \mathrm{ADP} + \mathrm{ENZ} + \mathrm{GLP} - \mathrm{ENG} - \mathrm{ITP}$
	3)	${}^{0}_{10^{-3}}\mathrm{GLP} + {}^{10^{-5}}_{10^{-5}}\mathrm{ETZ} \xrightarrow{2\times10^{10}} \mathrm{ETG} - \mathrm{GLP} - \mathrm{ETZ}$
	4)	${}^{0}_{10^{-4}}\mathrm{ETG} + {}^{5\times10^{-4}}_{1.5\times10^{-3}}\mathrm{ITP} \xrightarrow{1\times10^{9}} \mathrm{GPP} + \mathrm{ETZ} + \mathrm{ADP} - \mathrm{ITP} - \mathrm{ENG}$
	5)	${}^{0}_{10^{-3}}\mathrm{GPP} \xrightarrow{10^{6}} \mathrm{GAP} + \mathrm{DHA} - \mathrm{GPP}$
Glycolytic phosphorylations of ADP	6)	${}^{0}_{10^{-3}}\mathrm{DHA} + {}^{10^{-4}}_{2\times10^{-4}}\mathrm{DPH} \xrightarrow{10^{9}} \mathrm{DPN} + \mathrm{AGP} - \mathrm{DHA}$
	7)	${}^{0}_{10^{-4}}\mathrm{GAP} + {}^{10^{-4}}_{2\times10^{-4}}\mathrm{DPN} \xrightarrow{10^{9}} \mathrm{DPH} + \mathrm{BGA} - \mathrm{GAP} - \mathrm{DPN}$
	8)	${}^{0}_{10^{-4}}\mathrm{BGA} + {}^{5\times10^{-3}}_{5\times10^{-3}}\mathrm{PI\Delta} \xrightarrow{10^{8}} \mathrm{DGA} - \mathrm{BGA} - \mathrm{PI\Delta}$
	9)	${}^{0}_{10^{-4}}\mathrm{DGA} + {}^{10^{-4}}_{10^{-3}}\mathrm{ADP} \xrightarrow{5\times10^{9}} \mathrm{ITP} + \mathrm{PGA} - \mathrm{DGA} - \mathrm{ADP}$
	10)	${}^{0}_{10^{-4}}\mathrm{PGA} + {}^{10^{-4}}_{10^{-3}}\mathrm{ADP} \xrightarrow{5\times10^{9}} \mathrm{ITP} + \mathrm{PYR} - \mathrm{PGA} - \mathrm{ADP}$
	11)	${}^{1.5\times10^{-3}}_{2\times10^{-3}}\mathrm{PYR} + {}^{10^{-4}}_{2\times10^{-4}}\mathrm{DPH} \xrightarrow{5\times10^{8}} \mathrm{LAC} + \mathrm{DPN} - \mathrm{PYR} - \mathrm{DPH}$
	12)	${}^{10^{-2}}_{1.2\times10^{-2}}\mathrm{LAC} + {}^{10^{-4}}_{2\times10^{-4}}\mathrm{DPN} \overset{1\times10^{7}}{\dashrightarrow} \mathrm{PYR} + \mathrm{DPH} - \mathrm{LAC} - \mathrm{DPN}$
Oxidative phosphorylations of ADP	13)	${}^{1.5\times10^{-3}}_{2\times10^{-3}}\mathrm{PYR} + {}^{10^{-4}}_{2.04\times10^{-4}}\mathrm{DIN} \xrightarrow{1.5\times10^{8}} \mathrm{DIH} - \mathrm{DIN} - \mathrm{PYR}$
	14)	${}^{10^{-4}}_{2.04\times10^{-4}}\mathrm{DIH} + {}^{5\times10^{-5}}_{1.5\times10^{-4}}\mathrm{X{\cdot}I} + {}^{1.5\times10^{-3}}_{1.5\times10^{-3}}\mathrm{OXY} \xrightarrow{2.4\times10^{12}} \mathrm{DIN} + \mathrm{XSI} - \mathrm{DIH} - \mathrm{X{\cdot}I} - \mathrm{OXY}$
	15)	${}^{5\times10^{-5}}_{1.5\times10^{-4}}\mathrm{XSI} + {}^{5\times10^{-3}}_{5\times10^{-3}}\mathrm{PI\Delta} \xrightarrow{4\times10^{8}} \mathrm{XSP} - \mathrm{XSI} - \mathrm{PI\Delta}$
	16)	${}^{5\times10^{-5}}_{1.5\times10^{-4}}\mathrm{XSP} + {}^{10^{-4}}_{10^{-3}}\mathrm{ADP} \xrightarrow{5\times10^{9}} \mathrm{2TP} + \mathrm{X{\cdot}I} - \mathrm{XSP} - \mathrm{ADP}$
	17)	${}^{5\times10^{-5}}_{1.5\times10^{-4}}\mathrm{XSI} + {}^{0\to2.5\times10^{-3}}_{2.5\times10^{-3}}\mathrm{DBP} \xrightarrow{4\times10^{8}} \mathrm{X{\cdot}I} - \mathrm{XSI}$
ATP utilization and transfer	18)	${}^{5\times10^{-4}}_{1.5\times10^{-3}}\mathrm{2TP} + {}^{0\to2.5\times10^{-3}}_{2.5\times10^{-3}}\mathrm{DBP} \xrightarrow{4\times10^{6}} \mathrm{1TP} - \mathrm{2TP}$
	19)	${}^{10^{-3}}_{3\times10^{-3}}\mathrm{1TP} + {}^{2\times10^{-6}}_{3\times10^{-6}}\mathrm{PUE} \xrightarrow{3\times10^{9}} \mathrm{PPP} - \mathrm{1TP} - \mathrm{PUE}$
	20)	${}^{10^{-6}}_{3\times10^{-6}}\mathrm{PPP} \xrightarrow{2\times10^{6}} \mathrm{PUE} + \mathrm{ADP} + \mathrm{PI\Delta} - \mathrm{PPP}$

TABLE II

CHEMICAL AND COMPUTER TERMINOLOGY FOR COMPONENTS OF METABOLIC CONTROL SEQUENCES

Chemical name	Symbol	
	Equations	Graphs
Glucose	GLU	*G*
Hexokinase	ENZ	*Z*
Hexokinase-glucose intermediate	ENG	*E*
Adenosine triphosphate in store I	1TP	*C*
Adenosine diphosphate	ADP	#
Glucose-6-phosphate	GLP	*L*
Phosphofructokinase	ETZ	—
Phosphofructokinase intermediate	ETG	*T*
Hexosediphosphate	GPP	*P*
Glyceraldehyde-3-phosphate	GAP	*A*
Extramitochondrial diphosphopyridine nucleotide	DPN	*N*
Extramitochondrial reduced diphosphopyridine nucleotide	DPH	—
Acyl enzyme intermediate of glyceraldehyde-3-phosphate dehydrogenase	BGA	*B*
Inorganic phosphate	PI	$
1,3-diphosphoglycerate	DGA	*D*
3-phosphoglycerate	PGA	*Q*
Pyruvate	PYR	*R*
Lactate	LAC	—
Intramitochondrial diphosphopyridine nucleotide	DIN	—
Intramitochondrial reduced diphosphopyridine nucleotide	DIH	*H*
Low-energy intermediate in oxidative phosphorylation	X·I	*X*
High-energy intermediate in oxidative phosphorylation	XSI	—
Oxygen	OXY	*O*
Phosphorylated intermediate in oxidative phosphorylation	XSP	*
Adenosine triphosphate in store II	2TP	*V*
Dibromophenol	DBP	—
Enzyme concerned in ATP utilization	PUE	—
Enzyme intermediate concerned in ATP utilization	PPP	%

error according to (9). Since y'' varies in the course of the solution, a fixed value of Δt either produces too large an error or makes the solution too slow for a given error. Thus Δt has been varied to obtain a more rapid solution with a given error. The computer automatically adjusts the size of Δt so that the estimated error is never larger than some specified error, ϵ, and never smaller than 0.1 ϵ. This automatic error control produces a solution in an optimum time at the desired accuracy. The solutions presented here are based upon a one per cent error, as can be seen by the fluctuations in the traces (Fig. 31).

To present the output data in analog form and thereby afford a ready comparison with the spectrophotometric data (*e.g.*, Fig. 12), a graph routine has been devised [63] in which the total interval of computations is divided into 200 equal parts. The computer then finds and plots the concentrations corresponding to the nearest integral in the actual solutions. Each chemical is now represented by a single letter [oxygen = O, etc. (Table II)]. The full scale of the graph is the distance between rows of dots, representing the maximum value for a particular chemical. These values appear in Table I. Chemical concentrations that overlap on the graph are printed at the top margin with the letter with which they overlap printed last, as, for example, in Fig. 31, where a number of chemicals overlap. One has the letter sequence #*G*, indicating that # has the same value as *G*. A time scale is printed at the bottom of the graph which, as mentioned above, can be read for the purpose of comparison with the living cell in terms of seconds. The graphs can be printed with up to 10 variables on a sheet.

Table II identifies the symbols used in printing the material as well as the three letter code used in equation writing and the proper chemical names.

The computer representation of the metabolic control [23], [68], phenomena of the ascites tumor cell is illustrated by Fig. 31. The solution is in three phases:

1) 0–32 time units represents the adjustment of the system to a steady-state value;

2) 32–114 time units represents the response of the system to glucose addition. At 32 time units, glucose is added (note that *G* goes from zero to its maximal value) and is rapidly utilized, as indicated by the downward sweep of the *G* trace (Fig. 31). Oxygen utilization, which was rather slow prior to the addition of glucose, is rapidly accelerated for some interval due to 1TP expenditure (trace *C*) and ADP formation (#). Thus, control of oxygen utilization by glucose addition is demonstrated by the computer solution.

A response mentioned above, related to the Crabtree effect, is that glucose addition will inhibit oxygen utilization. The computer solution shows that at approximately 100 time units the utilization of oxygen has fallen almost to zero. Apparently the control chemical for oxygen utilization, ADP, has fallen to zero because ATP store I is exhausted (trace *C*). Thus a second metabolic regulation can be demonstrated.

3) 114–end time units represents the effect of "uncoupling agents." It is observed experimentally that these reagents will reactivate the inhibited metabolism of the tumor cell. This is simulated at time units 114 by the addition of dibromophenol (DBP) which, in our representation, allows ATP in store II (2TP) (*V*) to become mixed with the ATP in store I (1TP) (*C*). Thus, the effect of DBP is to cause a rapid diminution of *V* and a rapid increase of *C*. Simultaneously there is a rapid increase of oxygen utilization (*O*) and a reactivation of glucose utilization (*G*). Thus, a third well-established biochemical response is exhibited by the computer solution.

This digital computer program provides a most powerful tool for the detailed investigation of a variety of physical and chemical systems. Sufficient flexibility is provided so that it is possible to solve and graph the solutions of complex sequences of differential equations which represent biological systems. Furthermore, representation on the basis of other laws than that of mass

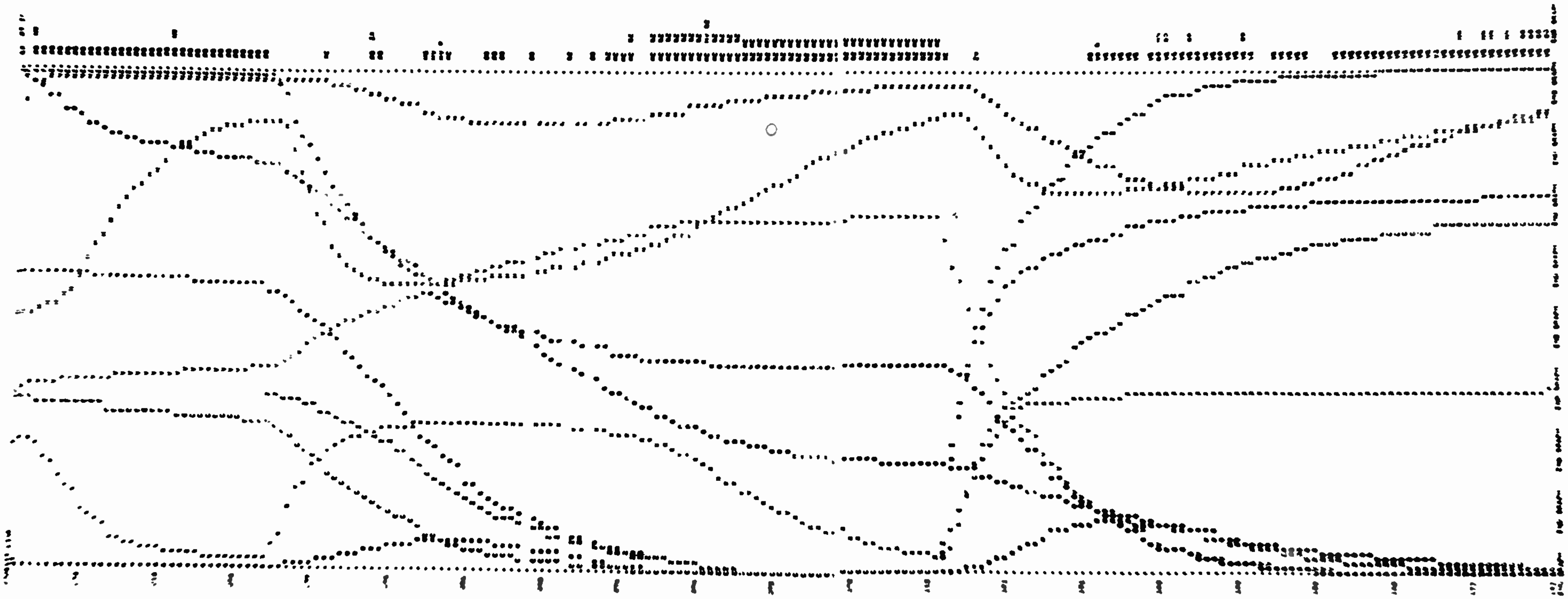


Fig. 31—High-speed printer record of the kinetics of ten intermediates in metabolic control systems of Figs. 29 and 30. The identification of the chemicals is given in Table II. The identification of the time scale at the bottom of the graph, the normalization of the concentrations for the variables and the key to the crossing over of the traces at the top of the graph is explained in the text (DC-4). (Data obtained with Univac I, with the aid of the University of Pennsylvania, Computer Center.)

action is surely possible, although the scope and applicability of this approach remains to be fully explored.

Bibliography

[1] D. Keilin, "On cytochrome, a respiratory pigment common to animals, yeast, and higher plants," *Proc. Roy. Soc. (London) B*, vol. 40, pp. 312–339; August, 1925.
[2] O. Warburg, "Heavy Metal Prosthetic Groups and Enzyme Action," Oxford University Press, London, Eng., pp. 1–230; 1949.
[3] B. Chance, "Spectra and reaction kinetics of respiratory pigments of homogenized and intact cells," *Nature*, vol. 169, pp. 215–230; February, 1952.
[4] D. Keilin and E. C. Slater, "Cytochrome," *Brit. Med. Bull.*, vol. 9, pp. 89–97; February, 1953.
[5] M. Dixon and E. C. Webb, "Enzymes," Academic Press, Inc., New York, N. Y., pp. 1–782; 1958.
[6] F. Lipmann, "Metabolic process patterns," in "Currents in Biochemical Research," D. E. Green, Ed., Interscience Publishers, Inc., New York, N. Y., pp. 137–148; 1946.
[7] A. Ehrenberg and H. Theorell, "Stereochemical structure of cytochrome c," *Nature*, vol. 176, pp. 158–159; July, 1955.
[8] D. Keilin and E. F. Hartree, "Effect of low temperature on the absorption spectra of hemoproteins; with observations on the absorption spectrum of oxygen," *Nature*, vol. 164, pp. 254–259; August, 1949.
[9] R. Estabrook, "The low temperature spectra of hemoproteins I. Apparatus and its application to a study of cytochrome c," *J. Biol. Chem.*, vol. 223, pp. 781–794; December, 1956.
[10] B. Chance and G. R. Williams, "A method for the localization of sites for oxidative phosphorylation," *Nature*, vol. 176, pp. 250–254; April, 1955.
[10a] H. Lardy, "Energetic coupling and the regulation of metabolic rates," in "Third International Congress of Biochemistry, Brussels, 1955," C. Liebecq, Ed., Academic Press, New York, N. Y., pp. 287–294; 1956.
[11] M. H. Cardow and D. D. Eley, "The semiconductivity of organic substances, pt. 3, haemoglobin and some amino acids," *Faraday Society Discussions*, in press; 1959.
[12] C. P. S. Taylor, "Energy transfer with special reference to biological systems," *Faraday Society Discussions*, in press; 1959.
[13] B. Chance and G. R. Williams, "The respiratory chain and oxidative phosphorylation," in "Advances in Enzymology," F. F. Nord, Ed., Interscience Publishers, Inc., New York, N. Y., vol. 17, pp. 65–134; 1956.
[14] B. Chance and E. L. Spencer, Jr., "Stabilization of 'steady states' of cytochromes at liquid nitrogen temperatures," *Faraday Society Discussions*, in press; 1959.
[15] L. N. M. Duysens, "Reversible photo-oxidation of a cytochrome pigment in photosynthesizing *Rhodospirillum rubrum*," *Nature*, vol. 173, pp. 692–693; April, 1954.
[16] B. Chance and L. Smith, "Respiratory pigments of *Rhodospirillum rubrum*," *Nature*, vol. 175, pp. 803–809; May, 1955.
[17] J. M. Olson, "Cytochrome reactions in *Chromatium*," *Biochim. et Biophys. Acta.*, vol. 28, pp. 227–228; May, 1958.
[18] M. Calvin, "Free radicals in photosynthetic systems," *Revs. Mod. Phys.*, vol. 31, pp. 157–161; January, 1959.
[18a] B. Chance, L. Piette, and L. Bicking, unpublished data.
[19] R. W. Estabrook and A. Holowinski, unpublished results.
[20] R. Perry, B. Thorell, L. Åkerman, and B. Chance, "Localization and assay of respiratory enzymes in single living cells. Absorbancy measurements on the Nebenkern," *Nature*, in press.
[21] B. Chance, R. Perry, L. Åkerman, and B. Thorell, "A highly sensitive recording microspectrophotometer," *Rev. Sci. Instr.*, vol. 30, pp. 735–741; August, 1959.
[22] F. Jobsis and B. Chance, "Time relations between muscular contraction and respiration of cytochrome chain," *Federation Proc.*, vol. 16, p. 68; March, 1957.
[23] B. Chance and B. Hess, "Spectroscopic evidence of metabolic control," *Science*, vol. 129, pp. 700–708; March, 1959.
[23a] B. Chance and B. Hess, "On the control of metabolism in ascites tumor cell suspensions," *Ann. N. Y. Acad. Science*, vol. 63, pp. 1008–1017; March, 1956.
[24] T. K. Walker, "Pasteur's work on fermentation and its significance for present-day studies in biochemistry," *Nature*, vol. 181, pp. 940–942; April, 1958.
[25] F. Lynen, "Phosphatkreislauf und Pasteur-Effekt," *Proc. Internatl. Symp. on Enzyme Chemistry*, Tokyo-Kyoto, Japan, 1957, Maurzen Co., Ltd., Tokyo, Japan, pp. 25–34; 1958.
[26] H. Crabtree, "Observations on the carbohydrate metabolism of tumors," *Biochem. J.*, vol. 23, pp. 536–545; April, 1929.
[27] E. Racker, "Carbohydrate metabolism in ascites tumor cells," *Ann. N. Y. Acad. Sci.*, vol. 63, pp. 1017–1027; March, 1956.
[28] B. Chance, "Enzymes in action in living cells: the steady state of reduced pyridine nucleotides," in "The Harvey Lectures, Series 49, 1953–54," Academic Press, Inc., New York, N. Y., pp. 145–175; 1955.
[29] J. M. Kolthoff and J. J. Lingane, "Polarography," Interscience Publishers, Inc., New York, N. Y., vol. 1, pp. 297–349; 1941.
[30] P. W. Davies and F. Brink, Jr., "Microelectrodes for measuring local oxygen tension in animal tissues," *Rev. Sci. Instr.*, vol. 13, pp. 524–533; December, 1942.
[31] E. D. Harris and A. J. Lindsey, "Vibrating electrodes in polarography," *Nature*, vol. 162, p. 413; September, 1948.
[32] B. Chance, "Techniques for assay of respiratory enzymes," in "Methods in Enzymology," S. P. Colowick and N. O. Kaplan, Eds., Academic Press, Inc., New York, N. Y., vol. 4, pp. 273–329; 1957.
[33] P. W. Davies, personal communication.
[34] L. C. Clark, Jr., "Monitor and control of blood and tissue oxygen tension," *Trans. Amer. Soc. Art. Int. Org.*, vol. 2, p. 41; 1956.
[35] B. Chance, "Cellular oxygen requirements," *Federation Proc.*, vol. 16, pp. 671–680; September, 1957.
[36] H. Lundegårdh, "Properties of the cytochrome system of wheat roots," *Arkiv Kemi*, vol. 5, pp. 97–146; February, 1953.
[37] B. Chance, "Rapid and sensitive spectrophotometry III. A double-beam apparatus," *Rev. Sci. Instr.*, vol. 22, pp. 634–638; August, 1951.
[38] B. Chance, "Spectrophotometry of intracellular respiratory pigments," *Science*, vol. 120, pp. 767–775; November, 1954.
[39] H. Hartridge and F. J. W. Roughton, "A method for measuring the velocity of very rapid chemical reactions," *Proc. Roy. Soc. (London) A*, vol. 104, pp. 376–394; 1923.
[40] F. J. W. Roughton, "Rapid reactions," in "Technique of Organic Chemistry: Investigations of Rates and Mechanisms of Reactions," S. L. Friess and A. Weissberger, Eds., Interscience Publishers, Inc., New York, N. Y., vol. 8, pp. 669–690, 710–738; 1953.
[41] B. Chance, see pp. 690–710 of [40].
[42] B. Chance, "Regeneration and recirculation of reactants in the rapid flow apparatus I. Design criteria," *Faraday Society Discussions*, vol. 17, pp. 120–123; April, 1955.
[43] B. Chance and C. M. Connelly, "A method for the estimation of the increase of concentration of adenosine diphosphate in muscle sarcosomes following a contraction," *Nature*, vol. 179, pp. 1235–1236; June, 1957.
[44] J. Ramírez, "Oxidation-reduction changes of cytochromes following stimulation of amphibian cardiac muscle," *J. Physiol.*, vol. 147, pp. 14–32; June, 1959.
[45] C. C. Yang and V. Legallais, "A rapid and sensitive recording spectrophotometer for the visible and ultraviolet region I. Description and performance," *Rev. Sci. Instr.*, vol. 25, pp. 801–807; August, 1954.
[46] T. Caspersson, "Methods for the estimation of the absorption spectra of cell structures," *J. Roy. Microscop. Soc.*, vol. 60, p. 25; March, 1940.
[47] B. Thorell, "Cellular formation of intermediates during haemoglobin synthesis," in "Ciba Foundation Symposium on Porphyrin Biosynthesis and Metabolism," G. E. W. Wolstenholme and E. C. P. Millar, Eds., J. and A. Churchill, Ltd., London, Eng., pp. 174–182; 1955.
[48] R. Perry, B. Thorell, L. Åkerman, and B. Chance, "Microspectrophotometric measurements of the cytochromes in a single *in vivo* mitochondrion," *Biochim. Biophys. Acta* (in press).
[49] B. Thorell and B. Chance, "Localization and assay of respiratory enzymes in single living cells. Absorbancy measurements on liver and kidney cells," *Nature*, in press.
[50] B. Chance and B. Thorell, "Localization and assay of respiratory enzymes in single living cells. Fluorescence measurements of mitochondrial pyridine nucleotide in aerobiosis and anaerobiosis," *Nature*, in press.
[51] B. Chance and V. Legallais, "A differential microfluorimeter for the localization of reduced pyridine nucleotide in living cells," *Rev. Sci. Instr.*, vol. 30, pp. 732–735; August, 1959.
[51a] R. G. Picard, "Applications for Normalizing Amplifier, Type EMS-4A," RCA Engrg. Memo. (AS-6113); 1949.
[52] L. Åkerman, unpublished data.
[53] B. Chance, J. N. Thurston, and P. L. Richman, "Some designs and applications for packaged amplifiers using subminiature tubes," *Rev. Sci. Instr.*, vol. 18, pp. 610–616; September, 1947.
[54] B. Chance, "Rapid and sensitive spectrophotometry I. The accelerated and stopped-flow methods for the measurement of the reaction kinetics and spectra of unstable compounds in the visible region of the spectrum," *Rev. Sci. Instr.*, vol. 22, pp. 619–627; August, 1951.
[55] B. Chance, "A light-regulator," *Electronics*, vol. 13, pp. 24–25; February, 1940.
[56] B. Chance, J. G. Brainerd, F. A. Cajori, and G. A. Millikan, "The kinetics of the enzyme-substrate compound of peroxidase and their relation to the Michaelis theory," *Science*, vol. 92, p. 455; November, 1940.

[57] B. Chance, "The kinetics of the enzyme-substrate compound of peroxidase," *J. Biol. Chem.*, vol. 151, pp. 553–577; December, 1943.
[58] A. C. Burton, "The basis of the principle of the master reaction in biology," *J. Cellular Comp. Physiol.*, vol. 9, pp. 1–14; December, 1936.
[59] B. Chance, W. F. Holmes, J. Higgins, and C. M. Connelly, "Localization of interaction sites, in multicomponent systems: theorems derived from analogues," *Nature*, vol. 182, pp. 1190–1193; November, 1958.
[60] E. F. MacNichol, "An analog computer to simulate systems of coupled bimolecular reactions," this issue, p. 1816.
[61] B. Chance, G. R. Williams, W. F. Holmes, and J. Higgins, "Respiratory enzymes in oxidative phosphorylation V. A mechanism for oxidative phosphorylation," *J. Biol. Chem.*, vol. 217, pp. 439–451; November, 1955.
[62] J. Higgins, "A Theoretical Study of the Kinetic Properties of Sequential Enzyme Reactions," Ph.D. dissertation, University of Pennsylvania, Philadelphia, Pa.; 1959.
[62a] B. Chance and G. R. Williams, "Respiratory enzymes in oxidative phosphorylation III. The steady state," *J. Biol. Chem.*, vol. 217, pp. 409–427; November, 1955.
[62b] F. A. Laws, "Electrical Measurements," McGraw-Hill Book Co., Ltd., New York, N. Y., p. 672; 1917.
[62c] F. A. Holton, "Spectrophotometric evidence bearing upon the phenomenon of oxidative phosphorylation," *Chem. Weekblad*, vol. 54, p. 368; July, 1958.
[62d] W. F. Holmes, "Locating sites of interactions between external chemicals and a sequence of chemical reactions," *Trans. Faraday Soc.*, vol. 55, pp. 1122–1126; July, 1959.
[63] D. Garfkinkel, J. Higgins, and J. Rutledge, in preparation.
[64] O. Warburg, "On the origin of cancer cells," *Science*, vol. 123, pp. 309–314; February; 1956.
[65] J. P. Greenstein, "Some biochemical characteristics of morpho; logically separable cancers," *Cancer Res.*, vol. 16, pp. 641–653- August, 1956.
[66] B. Chance, in Discussion, "Ciba Foundation Symposium on the Regulation of Cell Metabolism," G. E. W. Wolstenholme and C. M. O'Connors, Eds., J. and A. Churchill, Ltd., London, Eng., pp. 219–229; 1959.
[67] J. Higgins, in preparation.
[68] B. Chance, "Quantitative aspects of the control of oxygen utilization," pp. 91–121 of [66].

CORRECTION

R. G. Allen and J. E. Mezei of IBM Research Center, Yorktown Heights, N. Y., have notified E. Goto, author of "The Parametron, a Digital Computing Element which Utilizes Parametric Oscillation," which appeared on pages 1304–1316 of the August, 1959 issue of PROCEEDINGS, of the following error in his paper.

The five input parity-check circuit (Fig. 13, page 1310) which should give a "1" output when an odd number of inputs are "1" does not appear to be correct, possibly because of an error in drafting.

Fig. 13 represents the following logical function:

$$f(x, y, z, u, v) = [[x\bar{y}z\bar{u}v][\bar{x}y\bar{z}uv][x\bar{y}z\bar{u}\bar{v}][\bar{x}y\bar{z}uv][\bar{v}]]$$

where the square brackets represent the majority function.

This five majority function can be reduced to the simpler function:

$$f(x, y, z, u, v) = [[x\bar{y}z\bar{u}v][\bar{x}y\bar{z}uv][\bar{v}]]$$

which does not yield the desired result in four cases (Table I) of the thirty-two possible combinations of five binary variables. A correct logical function for a five input parity check is

$$f(x, y, z, u, v) = [[x\bar{y}z\bar{u}\bar{v}][\bar{x}y\bar{z}uv][\bar{x}yz\bar{u}v][x\bar{y}\bar{z}uv][\bar{v}]]$$

which would be represented by the logical circuit shown in Fig. 1.

TABLE I

x	y	z	u	v	Goto Circuit Result	Desired Result
1	0	1	0	0	1	0
0	1	0	1	0	1	0
1	0	1	0	1	0	1
0	1	0	1	1	0	1

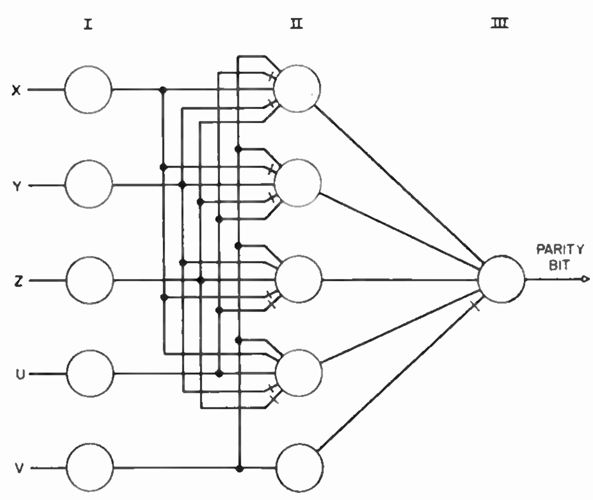


Fig. 1

Alternating Current Spectroscopy of Biological Substances*

H. P. SCHWAN†, FELLOW, IRE

Summary—The electrical properties of live matter are analyzed. The article first summarizes general principles which pertain to the frequency dependence of the electrical properties of any type of matter. It then states the particular mechanism which, at various parts of the total frequency spectrum, are predominantly responsible for observed data. They include time-dependent interface polarization, accumulation of charges due to inhomogeneous structure and orientation of polar molecules. The electrical properties of water and electrolytes, of protein suspensions, of subcellular and cellular structures are outlined in terms of previously mentioned mechanism. This, in turn, permits synthesis of the experimentally observed dielectric parameters of tissues. The treatment encompasses the total range of frequencies, from 1 cps to 100,000 mc. The article concludes with a chapter which outlines the application of the data and pertinent impedance techniques to a variety of basic and applied problems in biology and medicine.

Introduction

Spectroscopy, utilizing the visible, infrared and ultraviolet part of the electromagnetic spectrum, has been used extensively for the investigation of biological material. Its analytical value is based on the specific interaction of the light quanta with atoms and molecules. Other physical forces which are employed extensively for biological research are the constant gravitational pull (ultracentrifugation) and electrostatic forces (electrophoresis). This article will attempt to summarize the principles of alternating current spectroscopy of biological substances; *i.e.*, the mechanisms which are responsible for certain typical frequency relationships of the dielectric parameters of the biological substances under consideration.

From the earliest time of the knowledge of electricity, its interaction with biological material has intrigued mankind. The discovery of electricity itself was due to biophysical experimentation, as conducted by Galvani. The recognition that conduction of nerve impulses is based essentially on the propagation of electrical stimuli along "cable conductors" has only strengthened this interest. In view of the strong relationships between electricity and biology it appears surprising that during the last three decades a quantitative understanding of the factors which determine the electrical properties of biological material was obtained first. Earlier work, extending over many decades, was purely descriptive. During the 1930's, sound theoretical foundations first were applied to a more detailed understanding of the interesting dielectric properties of biological materials. This rather recent development is by no means concluded, but it is sufficiently advanced to warrant a review to familiarize the engineer and physicist, interested in electrical phenomena, with the present state of the art.

* Original manuscript received by the IRE, September 10, 1959. The work summarized in this article has been supported by research grants H 1253 from the National Institutes of Health, and by contract NR 102-289 between the Office of Naval Research, Dept. of the Navy and the University of Pennsylvania.

† Electromedical Division, The Moore School of Elec. Engrg., University of Pennsylvania, Philadelphia 4, Pa.

In the following article the "linear" and "passive" properties of biological substances will be summarized. The term "linear" applies to the validity of Ohm's law. It is found to be fulfilled for small signal strength, for example in muscular tissue for current densities below 1 ma/cm^2, this limit applying to low frequency currents and being less critical as the frequency increases. The term "passive" indicates that potentials applied to the biological material do not evoke specific biological response characteristics. This is true, for example, for excitable cells if the potential drop over the cell membrane which is evoked by the applied current stimulus is much smaller than the "resting potential" of the membrane; *i.e.*, smaller than about 70 mv.

Approach and Terminology

The electrical properties of matter are primarily of interest to the physicist and engineer. Pertinent tools and terminology will be applied to biological material. Biological material which encompasses such substances as tissues, body fluids, protein, and cellular suspensions. The magnetic properties of body tissues and proteins are practically identical with those of vacuum. The linear electrical properties of matter are completely characterized by its capacitance C and conductance G. In the case of a uniform field the following equations hold.

$$G = \kappa \frac{A}{d}; \qquad C = \epsilon\epsilon_r \frac{A}{d}, \tag{1}$$

where A is the electrode area and d the electrode distance. A uniform field is readily achieved with parallel electrode configurations if $A \gg d$. ϵ is the dielectric constant relative to free space, ϵ_r a numerical factor (8.84×10^{-14} if C is measured in farad and length dimensions in cm), and κ is the electrical conductivity (expressed in units of mho/cm if G is expressed in mho and A and d in cm units). The dielectric constant product $\epsilon\epsilon_r$ is the factor of proportionality between charge and potential across a unit volume of matter, and the conductivity is the one between current and potential.

The complex term $Y=G+j\omega C$ is the admittance and its inverse the impedance of the sample Z. The quantity

$$\epsilon^{+} = \epsilon - j\epsilon'' = \epsilon - j\frac{\kappa}{\omega\epsilon_r} \tag{2}$$

is the complex dielectric constant. The fact that we restrict our presentation to linear properties is identical with the statement that ϵ and κ are independent of potential. Of course, this does not imply independence of frequency or time. However, the independence from potential simplifies the presentation in either frequency or time domain. Suppose a step function potential is applied across a sample of dielectric material at the time $t=0$. Furthermore, consider a simple expotential function of the type $1-e^{-t/T}$, a good approximation of the often observed time dependence of the charge accumulated due to this step function potential. The equivalent frequency dependence of the complex dielectric constant is expressed by

$$\epsilon^{+} = \epsilon_{\infty} + \frac{\epsilon_0 - \epsilon_{\infty}}{1 + j\omega T} \tag{3}$$

which separates into

$$\epsilon = \epsilon_{\infty} + \frac{\epsilon_0 - \epsilon_{\infty}}{1 + (\omega T)^2}; \qquad \kappa = \kappa_{\infty}\frac{(\omega T)^2}{1 + (\omega T)^2}, \tag{4}$$

where the indexes 0 and ∞ refer to zero and infinite frequency and

$$\frac{(\epsilon_0 - \epsilon_{\infty})\epsilon_r}{T} = \kappa_{\infty}. \tag{5}$$

Thus, the existence of any mechanism which introduces a time constant T is seen to give rise to the frequency dependence as displayed in Fig. 1. From a formal point of view it is possible to express any time dependence of change invoked by a step function potential by replacing the simple equation $e^{-t/T}$ with a series of expotential functions, thus describing the frequency dependence of any linear system as a sum of curves of the type given in (3).

In a similar manner the transient behavior of the current induced by a step function potential represents either a simple expotential time dependence or can be thought of as composed of several expotential curves. The corresponding alternating current steady-state solution is derived for the case of one time constant T in the form

$$\Delta = \kappa_{\infty} + \frac{\kappa_0 - \kappa_{\infty}}{1 + j\omega T} \tag{6}$$

where Δ is the complex conductivity. A combination of the two processes is achieved by

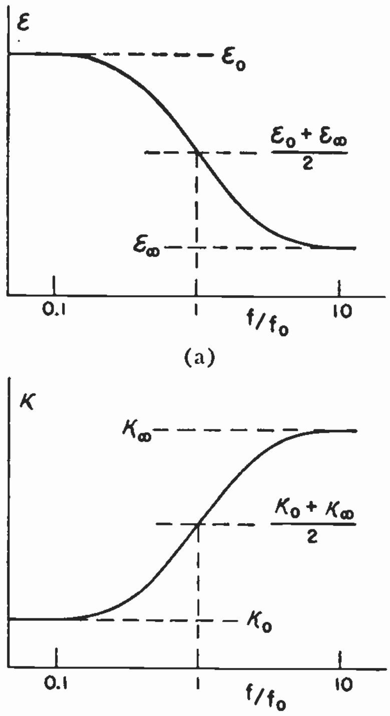


Fig. 1—Frequency dependence of dielectric constant and conductivity. The curves correspond to an exponential response of polarization of the type $1-e^{-t/T}$ to a step function potential.

$$\epsilon = \epsilon_{\infty} + \frac{\epsilon_0 + \epsilon_{\infty}}{1 + (\omega T)^2}; \; \kappa = \kappa_0 + (\kappa_{\infty} - \kappa_0)\frac{(\omega T)^2}{1 + (\omega T)^2} \tag{7}$$

$$\kappa_{\infty} - \kappa_0 = \frac{(\epsilon_0 - \epsilon_{\infty})\epsilon_r}{T}. \tag{8}$$

This cannot be any more readily combined into a simple expression of either complex dielectric constant or conductivity as given before in (3) and (6).

The formalistic considerations offered above do not state anything about the mechanism which gives rise to experimentally observed frequency dependences. They merely point out the type of frequency dependence which must exist if transient phenomena are of a non-instantaneous character. However, they suggest the following approach with regard to the analysis of dielectric data obtained at various frequencies.

1) Measurement of the frequency dependence of the electrical properties of the substance under investigation.

2) Determination of the time constants or distribution functions of time constants from an analysis of the experimental data. This task is complicated by the fact that we do not know in advance if time constants ought to be assumed in terms of charging currents (polarization) or conduction currents; *i.e.*, if we should attempt to present experimental data in terms of sums of expressions of the forms of (3) or (6).

3) Investigation of the possible physical mechanism which give rise to such time constants as obtained from

the experimental material. Pertinent models often suggest experiments which involve variation of parameters either on a structural level or by means of temperature, and greatly aid in the successful conclusion of this task. With it the task of "explaining" the observed dielectric data is then largely accomplished, or it is reduced to other problems which are concerned with charge distributions in molecules or with the dielectric properties of various constituents of the investigated substance.

The above outlined approach proved to be extremely powerful in the analysis of the factors which determine dielectric properties of biological material. It is not primarily concerned with making new contributions to biology, but we will see that such contributions invariably result, and that the application of alternating current techniques to biology establishes a new research tool which can solve problems which cannot be attacked by other means.

Relaxation Mechanism and Equivalent Circuits

Several groups of mechanism are known to give rise to electrical relaxation, *i.e.*, the particular frequency- and time-dependence discussed above.

The Existence of Polar Molecules

For such molecules the "center of gravity" of all its positive charges is not identical in space with the location of the "center of gravity" of all its negative charges. Thus, they will behave under the influence of an electrical field like a dipole and accordingly contribute to the total polarization. The ability of the molecule to rotate its dipole in the field direction is more restricted the higher the frequency, and also decreases as thermal movement increases. The theory of this type of polarization was developed by Debye [1] under simplifying assumptions. If the molecular shape is that of a sphere and the distance to the next polar molecule is large compared to the molecular dimensions, (3) is fulfilled with only one time constant involved. The value of the time constant can be exactly predicted under such circumstances in its relation to viscosity and molecular radius. A somewhat more complex case exists in case of a molecular shape which can be approximated by an ellipsoid with three different axes. In this case, for large distances between polar molecules, the solution has been given by Perrin [2] and is expressed by the sum of three expressions of the type in (3). The above outlined theory explains in many cases quantitatively and in all cases at least qualitatively the observed frequency dependence of the dielectric properties of molecular solutions. It has, in a major way, contributed to the knowledge of molecular structure.

Inhomogeneous Structure

Inhomogeneous structure of the dielectric material of interest, *i.e.*, variation of the dielectric constant and conductivity throughout the dielectric substance, cause interfaces, separating regions of different dielectric properties, to be charged if a step function potential is applied to the total dielectric. As a consequence, the total charge accumulated by the dielectric matter displays a transient behavior of the type discussed before, and, consequently, must give rise to dielectric relaxation effects. The theory of this type of dielectric relaxation is usually associated with the names Maxwell [3] and Wagner [4]. Maxwell was the first to discuss the case of the conductance κ of a suspension of spheres of conductivity κ_i, occupying a volume fraction p in a solvent of conductivity κ_a. His result

$$\frac{\kappa - \kappa_a}{\kappa + 2\kappa_a} = p\,\frac{\kappa_i - \kappa_a}{\kappa_i + 2\kappa_a} \tag{9}$$

can be applied to any type of dielectric property of both phases i and a by replacing the conductivities κ by complex conductivities $\Delta = \kappa + j\omega\epsilon\epsilon_r$. The separation of the resultant complex equation into real components permits for small p values to prove the validity of (7) where

$$T = \epsilon_r\,\frac{\epsilon_i + 2\epsilon_a}{\kappa_i + 2\kappa_a} \tag{10}$$

and

$$\epsilon_0 = \epsilon_\infty + 9p\,\frac{(\epsilon_i\kappa_a - \epsilon_a\kappa_i)^2}{(\epsilon_i + 2\kappa_a)(\kappa_i + 2\kappa_a)^2} \tag{11}$$

as shown by Wagner [4]. Wagner also treated the case of cylinders and the case of a series arrangement of different dielectric slabs. A substantial addition to this field was provided by biophysicists, interested in the analysis of cellular material during the 1930's. Useful solutions for the case of spheres surrounded with a shell of different dielectric properties were given by Fricke [5]–[8] in an attempt to approximate the geometry of cells, which are always surrounded by a membrane structure. For simplicity it was assumed that the inside and outside of the sphere are purely conductive; *i.e.*, that their dielectric constants can be neglected and that the membrane is purely capacitive (nonconducting). Recently, Pauly and Schwan [10] gave the complete theory of the shelled sphere, showing that it involves two relaxation terms of the form of (7) and stating how the parameters κ_0, κ_∞, T, ϵ_0, ϵ_∞ are related to the dielectric properties and geometrical dimensions of the three phases involved. More recently, Fricke gave a general and useful theorem pertaining to the case of spheres, each surrounded by several shells concentrically arranged [6]; he also recently calculated the low-frequency dielectric constant of a suspension of ellipsoids surrounded by shells [7]. The case of ellipsoidal shape has been treated only for the two phase system; *i.e.*, not assuming a shell [8]. The case of very elongated cells is best treated by approximating them by shelled cylinders, as has been treated by Cole [11], both for the case of frequency-independent dielectric properties of the shell, and the case where the shell is characterized by a frequency-independent electrical phase angle, a situation of particular biological interest.

A case which is also of particular biological interest and still yields relatively simple formulas was treated by Schwan [12]. It assumes purely conductive properties of the solution surrounding a membrane covered sphere and inside the sphere. The shell itself is assumed to have a capacity C_M and conductivity G_M per unit surface area; *i.e.*, in comparison with the more general treatment of Pauly and Schwan the dielectric constants of the media inside and outside the shelled spheres are considered too small to contribute significantly, an assumption which is justified throughout the LF and RF range. The solution of this problem may be stated here since it is sufficiently simple to help us demonstrate later some of the unique features of the Maxwell-Wagner type of dispersion found in biological materials. The frequency behavior is characterized by (7) and (8), involving only one relaxation time. The parameters of the equations are

$$T = RC_M \frac{\kappa_i + 2\kappa_a}{2\kappa_i\kappa_a + RG_M(\kappa_i + 2\kappa_a)} \tag{12}$$

$$\epsilon_0 = \frac{9}{4\epsilon_r} \frac{PRC_M}{\left[1 + RG_M\left(\frac{1}{\kappa_i} + \frac{1}{2\kappa_a}\right)\right]^2} \tag{13}$$

$$\kappa_\infty = \kappa_a\left[1 + 3p\frac{\kappa_i + \kappa_a}{\kappa_i + 2\kappa_a}\right] \tag{14}$$

$$\kappa_0 = \kappa_a\left[1 - \frac{3}{2}p\frac{1 + RG_M\left(\frac{1}{\kappa_i} - \frac{1}{\kappa_a}\right)}{1 + RG_M\left(\frac{1}{\kappa_i} + \frac{1}{2\kappa_a}\right)}\right]. \tag{15}$$

The particular properties of biological cells ($\epsilon_i, \epsilon_a \sim 60$; $\kappa_i, \kappa_a \sim 10^{-2}$ mho/cm; $C_M \sim 10^{-6}$ farad/cm^2; $G_M \sim 10^{-2}$ mho/cm^2; $R \sim 10\mu$) place the dispersion in the low RF range. The more general solution of Pauly and Schwan can be shown to be very closely identical with the sum of the solution expressed by (12) to (15) and the solution for solid spheres expressed by (10) and (11). The latter dispersion occurs for biological cells at ultrahigh frequencies and its magnitude is very weak in comparison with that of the dispersion in the low RF range.

Interface Polarization

Small particles, such as biological cells, colloidal particles, protein molecules, etc., are known to carry an electrical charge if suspended in an electrolyte solution. This charge can be either positive or negative and depends on the hydrogen ion concentration of the solution. It is compensated by the charge of a cloud of ions of opposite sign, which forms around the particle. The decrease in density of this ionic cloud with distance from the particle causes a corresponding change in potential, thus giving rise to a boundary potential. If suspension of charged particles is exposed to an alternating current field, part of the resulting alternating currents will pass through the boundary potential region of the particles and the existing dc boundary potential will be modulated by an ac component. The modulation potential at any point is proportional to the current density at the same point provided that we limit our discussion to small levels of alternating currents. The proportionality which exists between alternating current and potential permits the introduction of the concept of a surface conductance surrounding the particle. Indeed, if the impedance of a suspension of charged particles in an electrolyte is investigated by alternating current techniques, the total impedance can readily be analyzed in terms of the properties of the particle *per se* those of the surrounding electrolyte and a surface conductance. However, not much detailed knowledge is available about surface conductances. Data have been given particularly for glass particles (White, *et al.* [13] and Rutgers and de Smet [14], to mention only a few), which place the surface conductance in the range of 10^{-8}–10^{-9} ohm^{-1} per cm^2 surface. Work by Fricke and Curtis [15] has proven that the surface conductance is frequency dependent. From a formal point of view we may therefore characterize surface conductance behavior by a sum of expressions of the type in (3) with a sufficient number of time constants to allow for the broad frequency dependence involved. Since the validity of (3) necessitates the simultaneous applicability of (4), it comes as no surprise that a capacitive element also is involved. Indeed, the measurements of Fricke and Curtis revealed this capacitive component. It is, of course, also frequency dependent over the total range of observation. Hence it appears more reasonable to replace the concept of a surface conductance by that of a surface admittance, the latter one being frequency dependent. This frequency dependence appears of necessity, reflecting the applicability of sums of relaxation processes of the type in (3). If this concept is correct, it would appear that at sufficiently low frequencies, the surface admittance must eventually become frequency independent. This has indeed been proven by results obtained by Schwan and Maczuk [16]. In the latter case, measurements were carried out with sufficiently sensitive techniques and at sufficiently low frequencies to obtain stabilization of dielectric parameters as the frequency is lowered. A number of results, which have been obtained with bacteria, blood cells, polystyrene spheres of different size, fat particles, etc., suggest that the time constants involved in the phenomena of surface admittance vary in proportion to particle size. So far no detailed explanation of this fact has been formulated.

Finally we wish to point out that the properties of surface admittance are quite similar to those of electrode polarization impedances; *i.e.*, those related to the boundary potentials separating metal electrodes and electrolytes. In both cases the boundary potential is best characterized by an impedance or admittance which is strongly frequency dependent. This is, of course, not surprising in view of the fact that we deal in both cases with boundary potentials at the interface of a solid and an electrolyte and that the polarization impedance or surface admittance value is determined by the mode of modulation of a dc potential in an electrolytic phase.

All three of the above-mentioned mechanisms contribute to the frequency dependence of biological material. Surface admittance polarization appears predominantly at lower frequencies (at or near audio frequencies), the Maxwell-Wagner type of structural relaxation at higher frequencies (radio frequency range) and polar molecular rotation at radio and higher frequencies.

Equivalent Circuits and Presentations in Complex Planes

Very often the frequency range which is available to study a particular relaxation phenomena is too restricted to accurately obtain the "limit" values ϵ_0, ϵ_∞, κ_0, and κ_∞. In this case Argand circles are useful for purposes of extrapolation. Eq. (3) may be transformed into

$$\frac{\epsilon^* - \epsilon_0}{\epsilon_0 - \epsilon_\infty} = j\omega T, \tag{16}$$

yielding a circle in the complex dielectric plane as shown in Fig. 2(a) and a frequency dependence of u/v which is characterized by a slope of 45° in a logarithmic presentation. On the other hand,

$$\Delta = \kappa_\infty + \frac{\kappa_0 - \kappa_\infty}{1 + j\omega T}$$

yields a circle in the admittance plane [see Fig. 2(b)]. However, those in (7) do not yield a circle in either dielectric or admittance plane. One must first subtract κ_0 in order to obtain a circle in the complex dielectric plane or ϵ_∞ to obtain the circular presentation in the admittance plane. A circular presentation in the admittance plane always, of course, necessitates a circular presentation in the impedance plane where X is plotted against R; $Z = R \times jX$ being the total impedance.

It is easily shown that (7) may be represented by the circuit in Fig. 3(a), the case where κ_0 may be neglected (circle in dielectric plane) by the circuit in Fig. 3(b) and the case where $\epsilon_\infty = 0$ (circle in admittance plane) by the circuit in Fig. 3(c).

The time constants T involved in the relaxation equations are obviously identical with the time constants of the series combinations RC. In case of a distribution of time constants, a number of series RC arrangements are to be introduced in parallel, instead of the single ones depicted in Fig. 3(b) and 3(c). It is not possible to transform the circuit of the former to that of the latter. However, a generalized circuit as in Fig. 3(b), involving many time constants, *i.e.*, RC series arrangements in parallel, can be transformed into a circuit as in

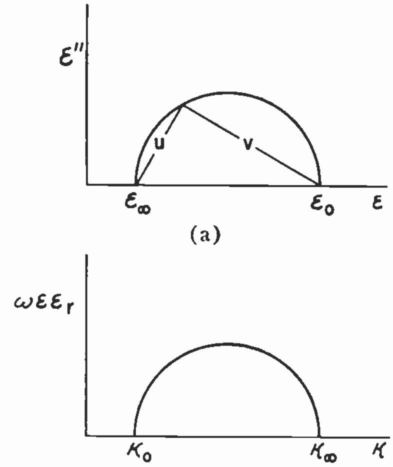


Fig. 2—Circular arcs in dielectric and admittance plane. The arcs are characteristic of a relaxation behavior which is characterized by one time constant.

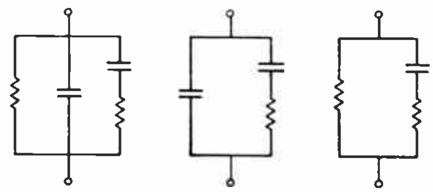


Fig. 3—Equivalent circuits of dielectrics whose behavior is characterized by one time constant. Circuits (b) and (c) correspond to circular arcs in the dielectric or admittance plane, respectively. Circuit (a) does not yield a circle in either dielectric or admittance plane.

Fig. 3(c) with many RC arrangements. Of course, the distribution of time constants (number of T values per T unit as function of T) will then be different in both circuits. This means that it is not possible to speak in terms of time constants without stating if the time constants are defined with regard to either dielectric or admittance plane.

Presentation of experimental observations are often found to result in circular plots with a depressed center (Cole and Cole [17]). The explanation for this behavior can be provided in two ways:

1) A circular behavior characterized by a circle with depressed center in the dielectric plane is obtained when a distribution function of time constants T

$$\frac{\sin \alpha\pi}{\cosh (1 - \alpha) \log T/T_0 - \cos \alpha\pi} \tag{17}$$

is assumed to represent the existence of a sum of dielectric relaxation expressions of the form in (3), (Cole and Cole [17]). The Cole-Cole factor α is defined by the equation of the circle

$$\epsilon^* = \epsilon_\infty + \frac{\epsilon_0 - \epsilon_\infty}{1 + (j\omega T)^{1-\alpha}}. \tag{18}$$

Its value is also easily obtained from the angle under which the circle intersects with the abscissa, the latter's value being $(1-\alpha)\pi/2$. The distribution function (17) is not statistically indicated; however, it is sufficiently close to statistically reasonable functions to make it practically impossible to differentiate experimentally between a true circle and the impedance locus based on a statistical function of Gaussian character (Schwan, [12]).

2) A replacement of the capacitor in the series RC arrangement of circuit in Fig. 3(b) by a polarization element with frequency independent phase angle can readily be shown to necessitate a circle with depressed center in the dielectric plane (Cole [18]). Pertinent work strongly indicates that many biological membranes and the corresponding C element in the equivalent circuit in Fig. 3(b) display this special frequency dependence. Hence it appears that in both mechanisms, variability of time constants due to corresponding variability of cell size and shape as well as the frequency dependence of biological membrane characteristics, provide adequate explanations for often-observed circular plots with depressed center in studies of biological material.

Biological Matter and Their Relaxation Properties

It appears convenient to analyze separately the dielectric properties of various tissue components, and then to synthesize these into the experimentally observed properties of tissues. A few words may be stated first to summarize the structure of tissue with high water content; such tissues are muscle and all body organs. Tissue is composed of cells, the latter ones being surrounded by cell membranes. While the cell membranes are structurally "solids" and of low conductivity, the material inside the cells and surrounding it is in the soluble state and of high electrical conductivity. In the intracellular solutions "swim" many subcellular elements, which in turn resemble cellular structure since they are equipped with membranes. Such components are the cell nuclei, mitochondria, and other more recently identified components. Furthermore the solution surrounding and inside these subcellular components contains many macromolecules of a size which is large compared to that of the molecules which compose the fluids itself. The most predominant classes of these macromolecules are the proteins and nucleic acids. Aside from their macromolecular concentration, the fluids inside and outside the tissue cells contain large amounts of salt. Thus, for our purposes, these fluids appear as suspensions of macromolecules in electrolytes. In the case of tissues, the cells are "interconnected," therefore providing the necessary mechanical stability for tissue. Blood cells, composed predominantly of red and various types of white cells, are freely movable. From a structural and electrical point of view blood appears quite similar to tissue. Tissues with comparatively low water content are bone and the fatty tissues. Their electrical properties, while less well investigated, support the fact that the basic electrical structure of such tissues is similar to those discussed before, the major difference being the comparatively small amount of electrolyte involved.

Water and Electrolyte Solutions

The water molecule is polar, due to the unsymmetric location of the two H atoms with respect to the O atom. Since it is smaller than other polar molecules, its characteristic frequency of polar rotation is considerably higher than that of other molecules. Its value varies with temperature and is near 20,000 mc. More detailed data have been given and summarized by Smyth [19]. The dielectric properties of water are characterized to an excellent degree by one single time constant, the Cole-Cole factor α being equal to zero accurate to 0.02 (Schwan [20]). More recent unpublished measurements carried out by Li in our laboratory refine this statement to $\alpha = 0 \pm 0.005$. The "high frequency" dielectric constant ϵ_∞ is quoted between 4 and 6; *i.e.*, higher than suspected from the optical index of refraction (suggesting less than 3) and ϵ_0 near 78.

The addition of salts does not affect the time constant of the polar relaxation of water noticeably. However the low frequency dielectric constant ϵ_0 is slightly affected; it decreases as the ionic strength increases about one dielectric unit for a tenth molar solution KCl. From the above data it is recognized that the dielectric constant of electrolytes is practically frequency independent up to 1000 mc. The same applies for the conductivity for the range of ionic strength values of biological interest. This simply reflects the fact that in this case the low-frequency conductance term κ_0 in the conductance relaxation equation is large compared with the Debye term

$$(\kappa_\infty - \kappa_0)\frac{(\omega T)^2}{1 + (\omega T)^2}$$

for frequencies below 1000 mc.

The rather simple dielectric behavior of water and electrolyte solutions enables us to discuss and analyze the behavior of suspensions of macromolecules in electrolytic solutions.

Macromolecular Suspensions

Proteins and other biological macromolecules carry a charge distribution which may be simulated by a dipole from an electrical point of view. As a consequence, these molecules are polar and display a strong frequency dependence of their electrical constants. Actually two Debye terms of the form in (3) describe the frequency dependence in most cases quite well (Oncley [21]). This is in agreement with Perrin's treatment of polar ellipsoids of revolutions [2] and reflective of the fact that the shape of many of the proteins investigated may be approximated by an ellipsoid of revolution. It has been possible, from the relationships between characteristic frequency, size and shape of polar molecule and observed magnitude of dispersion to establish permissible combinations of axial ratio of molecule and its effective size (Oncley [21]). The latter in turn is related to the actual volume of the molecule and the added volume of "bound" water which moves with the molecule under the influence of the electrical field. Thus, pertinent dielectric work has proved useful in determining hydration of proteins (bound water per weight unit protein matter), using axial ratios established by independent techniques. The relaxation effect due to the polarity of proteins is centered in the low mc range (Oncley [21]) and the dispersions magnitude $\epsilon_0-\epsilon_\infty$ up to 10 dielectric units per g protein in 100 cc.

As the frequency is lowered substantially below the characteristic frequency of the polar dispersion, the dielectric properties assume frequency independent character. Measurements with albumin by Takashima and Schwan (unpublished data) down to 100 cps illustrate that no new dielectric phenomena appear which would give rise to other frequency dependences.

At frequencies substantially above the characteristic frequency of the RF dispersion of proteins, the latter's bulk polar properties can not contribute any more to the molecules' dielectric behavior. This simply reflects the inability of the polar molecule to follow extremely rapidly varying fields with its rotation. The protein molecule can be regarded as an extremely poor conductor and its dielectric constant substantially lower than that of water; hence, it appears appropriate to consider it at high frequencies simply as a "dielectric hole" in the surrounding electrolyte. The application of a modification of Maxwell's (9) for ellipsoids is immediately indicated. This modification has been given by Fricke [22] in the form

$$\frac{\kappa - \kappa_a}{\kappa + x\kappa_a} = P\,\frac{\kappa_i - \kappa_a}{\kappa_i + x\kappa_a} \qquad (19)$$

where x is a factor which depends on the axial ratio of the ellipsoid and was tabulated by Fricke [22] (1925); the index i refers to the hydrated protein *in toto*. It is permissible to replace the conductivity terms in (19) by dielectric constants if conductive currents are low in comparison with capacitive current; *i.e.*, at sufficiently high frequencies. Hence (19) can be utilized to calculate the "effective" dielectric constant of the hydrated protein, considering here protein proper and its hydration shell as one unit. Fig. 4, obtained by Schwan and Li [23], shows that the effective dielectric constant of hemoglobin, the only protein investigated so far in detail above 100 mc, changes strongly with frequency. The frequency range of interest here is far too high to permit an explanation in terms of the bulk dipole moment of hemoglobin. Two independent explanations can be suggested.

1) The polar molecule possesses in turn polar groups of smaller size and of a sufficient degree of freedom to partially move with the field. The small size of these molecular components places their relaxation to higher frequencies than characteristic for the total molecules rotation. Indeed, molecules of smaller size than proteins display relaxation effects in the high megacycle range.

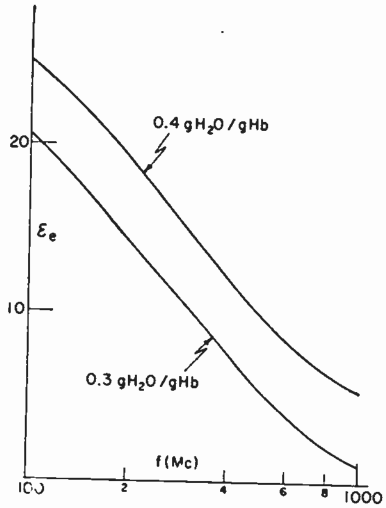


Fig. 4—Effective dielectric constant of hydrated hemoglobin as function of frequency. The two curves are based on different assumed hydration values as quoted. Data obtained by Li and Schwan.

2) The dielectric properties of bound water undergo a dispersion above 100 mc. The pertinent analysis places bound water with regard to its dielectric behavior between ice (relaxation frequency in the audio-frequency range) and free water (relaxation frequency near 20,000 mc) (Schwan [24]). Its low-frequency dielectric constant ϵ_i is found to be near 80, *i.e.*, comparable to that of ice and water, which lends weight to this interpretation. Its relaxation is characterized by a spectrum of time constants and "characteristic" frequencies $\frac{1}{2}\pi t$ extending from below 100 to above 1000 mc. This probably reflects the fact that bound water is composed of several layers of water which are coupled with decreasing tightness to the macromolecular surface. Further work is necessary to arrive at a definite decision between the two offered hypotheses. We have furthermore observed that the effective dielectric constant of hydrated hemoglobin increases rapidly with increasing hb concentration, while the frequency dependence for hemoglobin decreases. As the concentration increases beyond 10 per cent in volume a frequency-independent value above 100 mc is approached which is nearly identical with the value existent for albumin at a concentration of only 3 per cent (Grant [25]). While the exact interpretation of all these observations is as yet not available, there can be no doubt that they reflect molecular interactions. It is obvious that ultrahigh frequency studies of the dielectric properties of macromolecular suspensions have much to contribute to the knowledge about macromolecular behavior and bound water.

While the dielectric constant of hydrated protein matter is frequency dependent and high by comparison with the value anticipated from index of refraction data, it is not so high as to affect the over-all dielectric constant of the macromolecular suspension strongly. Hence, for a first-order approach, the dielectric constant of macromolecular suspensions can be understood in terms of perfect "dielectric holes" assuming an $\epsilon_i \ll \epsilon_a$ in (19). A more difficult situation pertains however with regard to the conductivity; data which are calculated by use of (19), assuming $\kappa_i = 0$, differ strongly from experimental observations. This has been demonstrated in detail for the conductivity of the interior of erythrocytes (Pauly and Schwan, unpublished, except in ONR-report form). Since the discrepancy is the more pronounced the higher the hemoglobin concentration, it must reflect a concentration dependent ion-binding effect of the protein component of the macromolecular suspension. Here again, electrical work, carried out at sufficiently high frequencies to enable determination of the internal conductance of cells (see below as to the appropriate approach of this problem), may well contribute to a more detailed understanding of the interaction of intercellular macromolecules with ions.

Biological Membranes

Biological membranes fulfill the following major tasks:

1) They surround the interior of biological cells and confine its content of electrolytes and macromolecules. Both are in most cases of a different composition and concentration than in the medium surrounding the cells.
2) They regulate the exchange of matter between the inside and the outside of cells in a manner which is specifically related to the task of any particular type of cell.

Hence, the structure of membranes involves two basic elements: a fairly "rigid" structural arrangement of molecules which serves to establish the barrier between the inside and the outside of the cell, and specific sides which regulate the exchange of matter from the outside to the interior and vice versa. While the former is now relatively well established, little is known about the latter. The exchange of matter may very well be thought of as taking place through small "holes." If this oversimplified model applies, the hole cross section occupies only an extremely small percentage of the total membrane surface and cannot be seen by conventional microscopic means. This statement includes electron microscopy at least at the present state of the art.

Electron microscope and X-ray diffraction have developed as major tools to study the structure of the molecular arrangement of membranes. It appears most likely that the latter is composed of a lipid layer which is covered on both sides with layers of proteins, while the lipid layer is probably of bimolecular character. The proteins seem to be oriented perpendicularly to the surface of the membrane. There can be no doubt that the over-all thickness of the membrane is about 100 to 150 Angstrom, while the innermost part of the membrane, *i.e.*, the bimolecular lipid film, is about 30 to 40 Å thick. From an analysis of dielectric constant measurements, predominantly based on the application of equations of the form (13), it has been possible to determine the capacity per unit area of a great number of cell membranes. Its value is found to be near 1 $\mu f/cm^2$. From the usual relationship between capacitance and dielectric constant,

$$C_M = \epsilon\epsilon_r \frac{A}{d}, \tag{20}$$

we must either conclude that the dielectric constant of the total membrane is 10–15 or that the capacitance originates only from the lipid layer, the latter having a dielectric constant of about 3. The second hypothesis must assume of course that the protein layers are sufficiently conductive to short circuit their own capacitance. Both possibilities must be considered. The first one is perhaps somewhat more likely in view of our considerations in the previous paragraph that hydrated proteins have at frequencies below 100 mc a high effective dielectric constant even though they need not be assumed to rotate with the field as a whole.

A refined analysis of the electrical properties of the bulk of the membrane matter is possible with large cells. This permits arrangements which include internally applied electrodes, so that measurements directly across the membrane are possible. External electrode arrangements of a simple field geometry are also possible with large cylindrically shaped cells such as the axon of the giant squid. Both types of investigations have been carried out (see, for example, Cole [26]). They show that the membrane capacitance often varies somewhat with frequency. At the same time, the conductivity of the membrane is observed to be frequency dependent in such a fashion that the electrical phase angle of the membrane appears practically frequency independent. This manifests itself in an impedance or admittance locus with a depressed center, as outlined before, the depression of the circle defining the phase angle of the membrane. A small frequency dependence of the capacitance and large frequency dependence of the conductance so that a nearly frequency-independent phase angle results, is characteristic for most low-loss dielectric materials. It represents the presence of a very wide distribution of relaxation times constants. It is not surprising to find that the bulk of the membrane matter is arranged in a manner which gives it similar electrical properties as observed for most low-loss dielectrics.

So far we have concerned ourselves with the properties of the bulk of the membrane matter. The presence of "holes," needed to accomplish the metabolic task of membranes, does not seem to contribute to the membrane capacitance. But it introduces an additional conductance term which appears, from an electrical point of view, shunted across the RC network which may be chosen to represent the bulk of the membrane matter

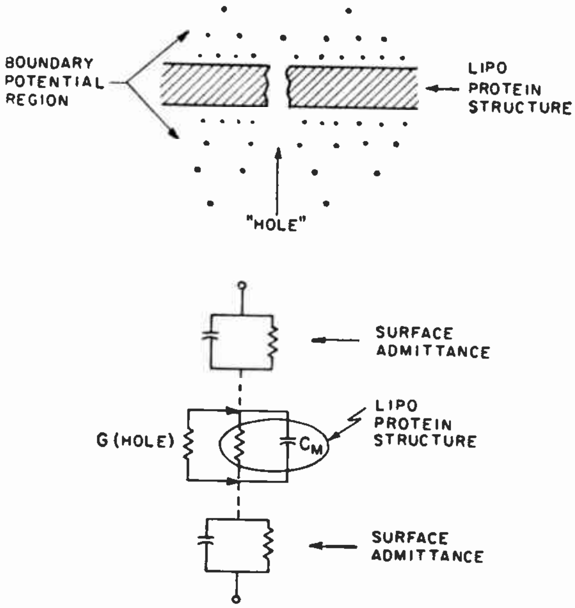


Fig. 5—Schematic of membrane structure and equivalent electrical circuit. Most of the elements of the circuit are frequency dependent.

(see Fig. 5). Such a frequency-independent resistive element, shunted across the membrane of constant phase angle does not disturb the appearance of an impedance locus with suppressed center. The validity of this statement is easily proven and omitted here due to limitations of space. Hence "holes" are virtually invisible in an investigation of the frequency dependence of the impedance of cells. Unfortunately the same applies if one attempts to determine the total membrane conductance, involving both the losses of the membrane matter as well as the "hole" conductance term, from an analysis of the conductance of a cellular suspension. This is due to the fact that in both (13) and (15) the RG_m terms in the denominator are too small to be detected. The reason for this becomes obvious if one considers likely values of the membrane conductance (order of magnitude 1 m mho/cm²). In essence, this argument represents the fact that the shunt provided by the external medium around the cells masks the conductance contribution arising from the membrane proper. Attempts are suggested by such considerations to replace the conductive medium inside and outside the cells by a highly resistive medium and thereby to emphasize the RG_m term. This was indeed attempted in the writer's laboratory with suspensions of erythrocytes. Even though it was possible to decrease both inside and outside conductivity κ_i and κ_a by several orders of magnitude, all attempts to detect the RG_m terms failed (Pauly, 1959; unpublished). The explanation for this result may be due to the fact that the G_m term adjusts proportionally with the terms κ_i and κ_a.

The only approach which succeeded in determinations of the membrane conductance G_m is restricted to large cells, employing techniques indicated before. Data obtained support the view that the total membrane conductance is about 1 to 10 mho per cm² membrane surface. However this approach has not been applied as yet over a sufficiently wide range of frequencies to permit a separation of bulk membrane matter losses and hole conductance *per se*.

All above statements pertain to frequencies in the radio and ultrahigh frequency region. At audio frequencies, we observe a strong change in the membrane capacitance and a corresponding change in its conductance in agreement with the relaxation equations of the type in (7). That this change is associated with the membrane surface can be readily seen from an inspection of (13) and (15) applicable for spherical cells. At frequencies which are substantially lower than those characteristic of the time constant or time constants which are associated with the structural "β dispersion" to be discussed later, they may be replaced by

$$\epsilon = \frac{9}{4\epsilon_r} pRC_M; \quad \kappa = \kappa_a\left(1 - \frac{3}{2} p\right) + \frac{9}{4} pRG_M. \quad (21)$$

Any experimentally observed change in ϵ and κ of a cellular suspension must therefore reflect a corresponding change in the membrane parameters C_m and G_m since the other constants involved can safely be assumed to be frequency independent at the low frequencies of interest. The observed change in the electrical parameters associated with the membrane is sometimes characterized by a single time constant (lysed erythrocytes, Schwan and Carstensen [27]) and sometimes associated with a broad spectrum of time constants (E.coli, Schwan [12]). It is also observed with solid particle suspensions, such as glass spheres, polystyrene spheres, fat particles, etc. Complete dispersion curves for these materials were obtained by Schwan and Mazcuk [16]. Other data, not completely covering the total range of dispersion but indicative of the same phenomena, have been obtained with such materials as glass and kaolin particles and with gelatin, soil, etc. Often amazingly high dielectric constants are obtained. The fact that solid and membrane covered particles display similar properties at low frequencies strongly indicates that the observed relaxation phenomena is not built into the membrane proper but rather originates at the surface of the particles. The common origin of all these low-frequency relaxation effects is furthermore strongly suggested by a relationship between particle size and average time constant of approximately linear character (Schwan [12]). According to this approximate proportionality, proteins should not be able to display this phenomena of increased dielectric constant at low frequencies. Indeed, the precise measurements on albumin suspensions by Takashima and Schwan mentioned before were unable to detect any change in ϵ and κ between 100 cps and 200 kcps.

The phenomenon discussed reflects the presence of surface polarization relaxation. In the case of particles surrounded with membranes, *i.e.*, biological cells, the surface polarization impedance appears as an integral part of the total membrane characteristic. No measurement, utilizing whatever electrode technique, is capable of separating between two impedance components one characteristic of the membrane *per se* and another characteristic of the surface polarization admittance. This is due to the fact that electrodes are too large to "push" through the ionic cloud responsible for the surface polarization term and measure the properties of the membrane *per se*. Our knowledge of the relative contribution of the surface admittance and membrane proper to the total membrane characteristic is therefore dependent on the study of solid particles carrying no membrane and consequently equipped with only a surface admittance.

If the total blame for the observed relaxation phenomenon at low frequencies is formally put on the membrane, its capacitance and conductance appear frequency dependent at low frequencies. For example, the membrane of lysed erythrocytes has then the properties illustrated in Fig. 6. Through a process of conversion from the frequency into the time domain, as indicated earlier, it follows that the electrical data of the membrane appear to be time dependent; *i.e.*, we observe that the conductivity, which is considered characteristic of the permeable function of the membrane, is time dependent if a step function potential is applied across the membrane. This phenomenon has indeed been observed and extensively studied for the membrane of the giant squid axon, (Cole [28], Hodgkin and Huxley [29]). From the material presented above it can be stated that this time dependence, found to be an essential prerequisite to biological processes of excitation (Cole [30]), may very well originate in part outside the membrane proper.

Cells and Subcellular Components

Not only cells, but many subcellular components are surrounded by a membrane. In the latter category, particles such as mitochondria and cell nuclei are of particular interest. No dielectric work has been conducted so far with cell nuclei. But a detailed investigation of the impedance of mitochondria has been conducted by Pauly, Packer and Schwan (1959, in print). It resulted in a membrane capacity of about 1 $\mu f/cm^2$; *i.e.*, is in agreement with the values published previously for many cell membranes. The work of Fricke, Schwan, Li and Bryson [31] was concerned with bacteria (E.coli) and also resulted in the same capacitance value. Cole [32] gave a survey of earlier results which pertains to a variety of biological cells such as erythrocytes of a variety of species, nerve cells, muscle cells, and even plant cells. The same value of about 1 $\mu f/cm^2$ membrane surface for the membranes electrical capacity seems always valid. The strongly indicated conclusion is that the particular lipo-protein combination, which establishes the bulk of the membrane structure and which is responsible for this capacitance value, is so uniquely suitable for the task assigned to it by nature, that it has been used by it universally. Differences in metabolic aspects of membrane performance, probably reflect only variability of that small part of the membrane, which we identified by the term "hole" before.

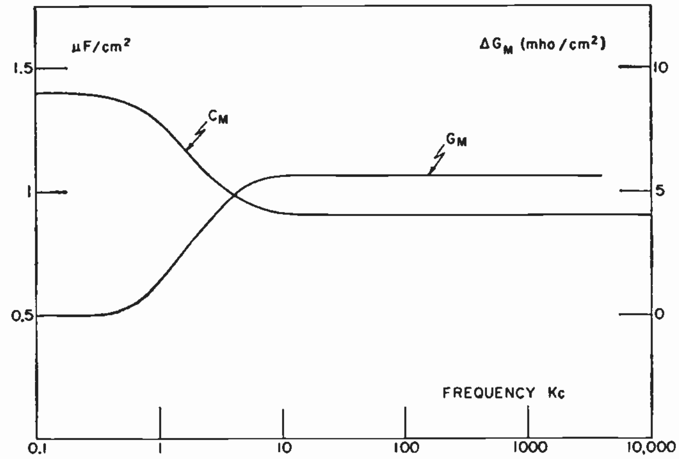
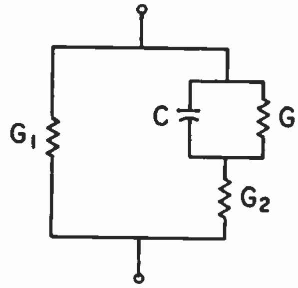


Fig. 6—Frequency dependence of the effective membrane capacitance and conductance of lysed erythrocytes. The ghost cells have been prepared by addition of distilled water. Data obtained by Schwan and Carstensen.

The frequency dependence of a suspension of cells is now easily understood; a dispersion of structural character can be predicted, which has been termed β dispersion (Schwan [12]). Its presence results from the existence of a time constant, which reflects the time needed to charge up the capacitance of the cell membranes through the "access impedances" (a term introduced by O. Schmitt) of cell interior and exterior. The equations which are applicable to this situation have been given for the case of a spherical shape of the cells: (7), (8), and (12)–(15). They hold in close approximation, provided that interior and exterior are predominantly resistive in character. For typical cell dimensions and interior and exterior impedance values (κ values of about 10^{-2} mho/cm) characteristic frequencies in the range between 100 kc and several mc result. This places the observed structural dispersion in the radio frequency range. Similar equations as given before have been obtained for ellipsoidal and cylindrical shape and applied to bacterial suspensions (Fricke, Schwan, Bryson and Li [31]).

It is of interest in this connection to discuss briefly an equivalent circuit for a cellular suspension. We have already pointed out that (7) and (8), the general relaxation equation, as applicable to a suspension of spherical cells whose inside and outside is predominantly resistive, can be expressed in terms of the equivalent circuit in Fig. 3(a). Introducing the values for the parameters of the relaxation equations from (12) to (15), it is readily possible to prove the validity of the equivalent circuit shown in Fig. 7. The physical significance of the circuit is obvious: part of the alternating current bypasses the



$$C = \frac{9}{4} pRC_M; \quad G = \frac{9}{4} pRG_M; \quad G_1 = \kappa_a \left(1 - \frac{3}{2} p\right);$$

$$G_2 = \frac{9}{2} p \frac{\kappa_i \kappa_a}{\kappa_i + 2\kappa_a}.$$

Fig. 7—Equivalent circuit for a suspension of spherical cells. Internal and external cell medium are assumed resistive.

cells, while another part passes through the membranes and permits the cell interior to participate in the conduction of electricity. The ratio of these two currents is, of course, strongly frequency dependent, due to the capacitive character of the membrane. It is also easily seen why the capacitor in the circuit of Fig. 7 is not only proportional to the membrane capacitance C_M but also to cellular volume fraction p and cell radius R. The radius is inversely proportional to the number of cells per unit distance in the field direction and characterizes how many capacitors C_M are arranged in series per unit distance. The volume fraction p is proportional to the number of cell membranes which can be charged.

It is apparent from all relationships, that it is possible from a determination of the frequency dependence of the impedance of a suspension of cells to extract the following information:

1) the capacity of the cell membrane;
2) the dielectric constant and conductivity of the cell interior; and
3) the dielectric constant and conductivity of the cell exterior.

The preparation of cellular suspensions permits in most cases the separation of the external medium from a packed cell sediment by standard techniques (centrifuge) and, furthermore, permits us to obtain suspensions of variable volume concentration. Hence, tests may be run as to the independence of the membrane capacitance and internal properties of the cell from the cell concentration, thereby substantially supporting the validity of the data obtained. Many applications of interest to the biological sciences are possible, one of which is chosen as an example. Erythrocytes, swollen and "contracted" mitochondria and E.coli have been investigated in our laboratory (Pauly, Schwan, Carstensen, and Packer, submitted for publication). The cells were prepared by repeated washing in solutions of different ionic strength, making it possible to equilibrate internal and external medium as much as possible. The analysis of pertinent impedance determinations permitted comparison of internal and external conductivity. Some of the results are demonstrated in Fig. 8 for mitochondria. Swollen mitochrondria and erythrocytes behave almost as perfect "osmometers"; *i.e.*, their internal ionic strength, as reflected by the proportionally related internal conductance, changes almost in proportion with the ionic strength of the external medium. In other cases such as E.coli, and particularly with contracted mitochondria, the internal conductance is at high ionic strength almost independent from the external medium, indicating that the release of ions is encouraged as the external mediums ion content is reduced. It appears that a residual amount of ions is fixed internally, yet free enough to move so that it can contribute to the conductance of the interior, only at physiological concentrations. The exchange of ions between cell interior and exterior is known to be controlled by the cell membranes. At the same time it appears from above quoted investigations that exchange of ions depends on organizational aspects of the cell interior, particularly the interaction between ions and macromolecules in the cells. This is an important research area, since it is intimately related to basic life processes.

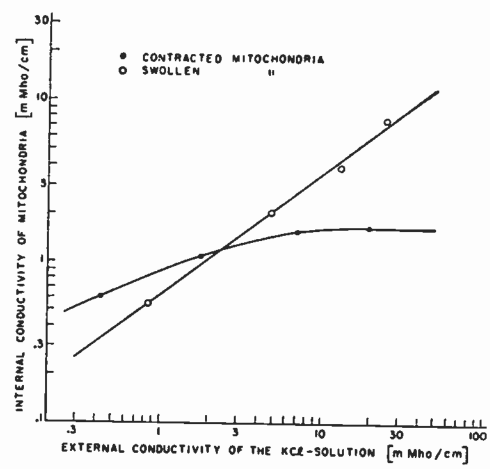


Fig. 8—Internal conductance of swollen and contracted mitochondria as a function of ionic strength of external medium. The latter quantity is presented in terms of electrical conductivity. Data obtained by Pauly, Packer, and Schwan.

Tissue Properties

Tissues may be considered as dense suspensions of biological cells. The synthesis of the dielectric properties of tissues by use of the aforementioned principles yields the type of over-all frequency response demonstrated in Fig. 9. The figure pertains to the dielectric constant of muscular tissue. The dielectric constant decreases with increasing frequency in three major steps, which are labeled as α, β and γ dispersion. At very high frequencies, the capacity of the cell membranes provides a very low and, therefore, "invisible" impedance in series with the other impedance elements provided by internal and external medium (see Fig. 7). Hence, the dielectric behavior is essentially that of a macromolecular suspension. It involves most markedly the

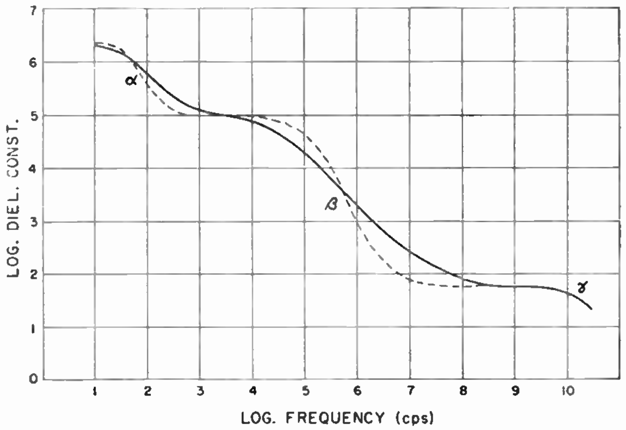


Fig. 9—Frequency dependence of the dielectric constant of muscular tissue. The dashed curve would characterize the frequency dependence if only three relaxation mechanisms would exist, each characterized by only one single time constant.

strong frequency dependence of the abundant water (about 70 per cent in muscular tissue). In the radio frequency range, the structural dispersion related to the presence of the cell membranes is predominant (β dispersion). It reflects the fact that the cell membrane capacitances contribute increasingly to the over-all impedance as the frequency becomes lower. At still lower frequencies, the access impedance elements become small in comparison with the membrane reactance and, finally, the dielectric properties of the membrane appear simply shunted across the frequency independent component, determined by the external medium; *i.e.*, the access impedance elements have now become "invisible" (see Fig. 7). Now the surface polarization effect eventually becomes pronounced and gives rise to the relaxation effect, labeled as α dispersion. Superimposed on these three major dispersion phenomena are by comparison minor additional relaxation effects:

1) The dispersion effects associated with the behavior of hydrated proteins and other macromolecular particles in the general frequency range between 100 and 1000 mc;
2) the protein dispersion in the low mc range associated with the bulk polar properties of proteins;
3) the structural dispersion of subcellular components, which are caused by their membranes and is also located in the low mc range.

In addition to this, variability of size and shape of the elements which contribute to the electrical properties are responsible for the existence of distributions of time constants rather than single terms. For example, an analysis conducted in this laboratory proves that the two-time constants associated with the structural dispersion of shelled spheres are to be replaced by six-time constants if we deal with ellipsoids which are surrounded by a confocal membrane (Shen and Schwan, in preparation for publication). This gives rise to a "flatter" frequency dependence. All these factors are responsible for the deviation of the experimentally observed data from the behavior, which is indicated by a dashed line in Fig. 9. The dashed line would be anticipated if only the three major relaxation phenomena (surface polarization, structural dispersion, polar water dispersion) were existent and if each were characterized by one single relaxation time constant.

It is possible to predict approximately with present-day knowledge and principles outlined above, the dielectric behavior of cell suspensions and tissue from the molecular, ionic, and cellular composition. These predictions are only semiquantitative in cases which involve very complex geometries. In other cases, particularly at high frequencies, they are quite accurate. A typical example pertains to the values of tissue dielectric constants near 300 mc; *i.e.*, at frequencies between the β and γ dispersion. The dielectric decrement for tissue solids, *i.e.*, the decrease in dielectric constant due to the addition of 1 g protein to 100 cc electrolyte is near 1 at 400 mc. Table I compares values calculated on the basis of this approach and compares these values with experimental data.

TABLE I

CALCULATED AND EXPERIMENTAL DIELECTRIC CONSTANTS AT 400 MC FOR SOME TISSUES AND BLOOD. THE VARIATION IN THE CALCULATED LIVER DATA REFLECTS CONSIDERABLE POSSIBLE VARIATION IN LIVER LIPID CONTENT

	Calculated	Experimental
Blood	57	57–59
Muscle	50–56	52–54
Liver	38–58	44–51
Skin	44	46–48

Dielectric properties reflect particularly at higher frequencies the amount of water in tissue. With higher water content or content of electrolytes both conductivity and dielectric constant increase from the values characteristic for the solid tissue components to those found for an electrolyte solution. Indeed the data for substances with low water content such as fatty tissues and bone are about one order of magnitude lower than those for the tissues with high electrolyte content such as muscle and most body organ tissues. The influence of the water content may be demonstrated for the case of fatty tissue in Fig. 10 and Fig. 11, obtained at a frequency of 900 mc. For an actual tabulation of dielectric constants and conductivity values of various body tissues we refer to Schwan [12].

APPLICATIONS AND PROBLEMS OF THE FUTURE

While the above outlined principles are of basic interest, several applications of biological impedance techniques to both basic and applied problems have developed or are indicated for further study. Of basic interest are applications of above outlined principles in the following areas.

1) Determination of the hydration of macromolecules. From the principles developed above it becomes evident that the total volume of the hydrated protein

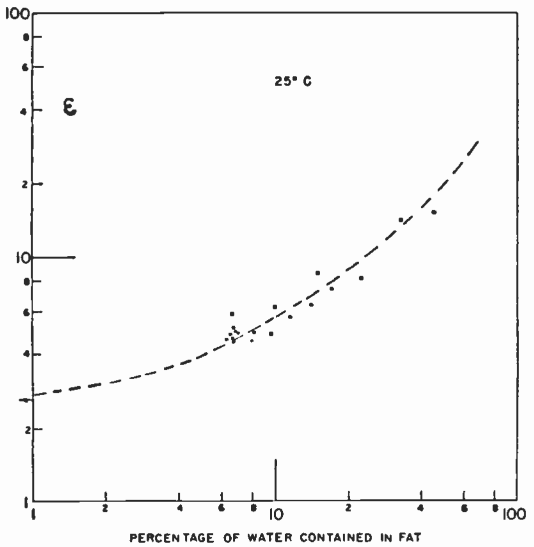


Fig. 10—Relationship between dielectric constant and water content of subcutaneous fat tissue. Each point represents one sample. The arrow marks a sample whose water was totally removed in the oven. Frequency 900 mc.

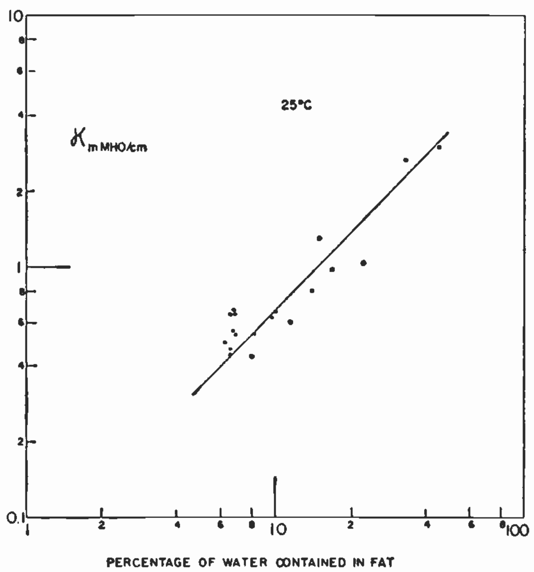


Fig. 11—Relationship between electrical conductivity and water content of subcutaneous fat tissue. Each point represents one sample. Frequency 900 mc.

molecule can be determined. This can be done by application of Fricke's equation from the effective dielectric constant of the protein solution, the dielectric constant of the solvent and the effective dielectric constant of the hydrated protein. From the electrically determined volume of the hydrated protein and the known protein volume, the hydration value is obtained by subtraction. If evaluations are based on measurements at very high frequencies, so that the dielectric constant of the hydrated protein has passed through its UHF dispersion, *i.e.*, above about 1000 mc, less uncertainty as to the hydrated proteins effective dielectric constant will exist. It is, in any case, very much smaller than that of the surrounding medium. Checks as to the independence of the result obtained at several very high frequencies, are indicated in view of the present status of this field. For further detailed discussions concerning this interesting possibility of determining hydration see Schwan [12] and, more recently, Grant [25].

2) Determination of hydration values from conductivity measurements. This technique, proposed more recently (Schwan [33]) is based on the assumption that the dispersion of the effective dielectric constant of hydrated proteins in the frequency range between 100 and 1000 mc and discussed above is due to a corresponding dispersion of bound water. The relationship between $\epsilon_0-\epsilon_\infty$ and $\kappa_\infty-\kappa_0$ formulated in (8), the fact that the total changes $\epsilon_0-\epsilon_\infty$ are nearly equal for bound and free water, and the difference in their characteristic frequency, suggest that $\kappa_\infty-\kappa_0$ is about 10 times smaller for bound than for free water. This and the fact that the conductance change for bound water occurs predominantly below 1000 mc, causes the frequency dependence of κ above 1000 mc to be due to free water. Hence a determination of this conductance increase and that to be due if only free water were existent, enables us to determine the volume taken by free water. This is accomplished by using a simplified version of Fricke's equation for conductance previously mentioned which is

$$p = \frac{\kappa_a - \kappa}{\kappa_a + \dfrac{1}{x}\kappa} \tag{22}$$

and by making the safe assumption that the hydrated protein conductance can be neglected. Subtraction of this volume from the total volume taken by protein *per se* and bound water yields the volume taken by bound water alone. This technique is less dependent than the dielectric technique mentioned above upon critical assumptions. The uncertain assumption as to the dielectric constant of the hydrated protein is now replaced by the much simpler and easily cross-checked assumption that the hydrated proteins conductance is small in comparison to that of the surrounding electrolyte at sufficiently high frequencies.

3) Studies of the exchange of ions through membranes are of importance with regard to its metabolic functioning. We have already pointed out the difficulties involved in any direct determination of the conductance of the membrane of biological cells. However, we can determine both internal and external conductance values (14) and can even monitor these values electronically as function of time. Hence we are able to follow internal ionic strength adjustments to imposed external changes of ionic strength. The transformation of pertinent records into permeability statements is straightforward.

4) It should be of interest to investigate how far the uniformity of membrane capacitance values, reflective of the same molecular membrane composition extends. Pertinent work with other bacteria, cell nuclei and even

virus should be of interest and may shed light on the question of on what level of organization live membrane structures are first used and how they are created.

There are other problem areas that are of more applied interest.

1) In electrocardiography, the analysis of body surface potentials which are due to the electrical activity of the heart, is advanced today to the point where the fine structure of the electrocardiogram in its relationship with the hearts activity is under theoretical and experimental study. Any attempt to elucidate from body surface potentials statements about the electrical heart generator without paying attention to small, but significant variability of the impedance of the tissues surrounding the human heart, is subject to criticism. Body surface potentials are affected by two factors: the electrical activity of the heart *per se*, and the characteristics of the media surrounding the heart. More details concerning tissue impedance characteristics are needed, particularly with regard to their unisotropic characteristics, to evaluate the tissue influence on the ECG.

2) In neurophysiology, conduction of nerve impulses by nerve fibers is accomplished by means of transmission line principles. The nerve cell is of cylindrical shape. Its cell interior and exterior provide the inner and outer conductor of a coaxial transmission line and the membrane establishes the dielectric, separating the two conductors. Nerve cell interior, exterior, and membrane determine the transmission constants; *i.e.*, characteristic impedance and propagation constant (O. Schmitt [34]). The membrane properties are frequency dependent, as discussed before, and furthermore become nonlinear for applied signals which compare with the signals propagated by the nerve. Hence, a complex nonlinear system of differential equations replaces the linear ones, normally used to characterize transmission line behavior. It has been shown that the time and amplitude dependence of membrane properties is a prerequisite for excitability (Cole [30]). Perhaps the most important advances in neurophysiology are due to the better knowledge of the electrical properties of the nerve membrane as achieved during the last decade. The synthesis of the propagated nerve signals from available membrane properties is at least in principle straightforward, to be attacked by appropriate and available principles. This statement is made with full realization that the nonlinear membrane characteristics cause this mathematical problem to be of sufficient size to require very extensive computations. On the other hand, more work must be undertaken to understand fully why nerve membranes are frequency, time, and amplitude dependent and why their electrical properties change as observed with environmental factors and signal strength. The pertinent detailed understanding will probably be provided when a better knowledge of the time dependent aspects of boundary potentials is achieved.

3) Diathermy, *i.e.*, the use of artificially induced local heat for therapeutic purposes, is based on the use of either high-frequency currents or the use of microwave radiation. While the former is applied with a frequency of 27 mc, the latter operates at 2450 mc. The use of ultrasonic radiation is not discussed here since it is unrelated and is mentioned at another place in this issue. The following problems arise in diathermy.

a) How deep does the artificially induced heat penetrate? This depends, of course, on the absorption coefficients of microwaves in tissue, or, in the case of the 27-mc current therapy (erroneously often called "ultra-short wave therapy"), the manner in which the currents spread and are conducted by the various tissues involved.

b) How can the amount of heat generated at any particular part of the body or in the whole area under treatment be judged from instrument readings?

The answer to both problems requires, as a prerequisite, a knowledge of the capacitive and resistive properties of body tissues at the frequencies of interest. With this knowledge available a good start has been made in recent years to solve above formulated problem areas [35]. More work ought to be encouraged if diathermy techniques is ever to become a scientifically controlled tool, to be applied to the patient in a precise manner.

4) More recently, considerable interest has been aroused by the possibility of injury due to exposure to strong sources of microwave radiation. The three major areas of potential harm are related with total body exposure, exposure of the eye resulting in cataracts, and exposure of the reproductive organs. A flux level of 10 mw/cm² absorbed energy has been suggested for tolerance purposes, (Schwan and Li [36]). However, there is no doubt that this figure establishes only an order of magnitude. It corresponds roughly to the case of light fever (about 1°C temperature rise) and irreversible and severe, perhaps lethal damage occurs first at approximately tenfold the tolerance level, *i.e.*, above 100 mw/cm². Sufficient dielectric work has been carried out, as outlined above, to establish all tissue properties of interest in the radar frequency range and to convert these data into absorption coefficients and depth of penetration constants by standard techniques. The knowledge of absorption coefficients has been obtained for such materials as muscle, fat, blood, eye-glass body fluid, eye-lens matter, bone, red and yellow bone marrow, liver, kidney, lung, white and gray brain matter, etc., (Schwan [37]) and is considered a prerequisite to any understanding of the effects of microwaves on mankind. We also know, from a discussion of simple geometries which neglect scattering, that subcutaneous fat can serve to match muscular tissue to air, so that under certain conditions almost all incident energy is absorbed (Schwan and Li [36]). We also know that for relatively slight variation in the respective thickness of skin, fat, etc., completely different amounts of radiation are made available for the body.

Nothing is as yet known about the scattering characteristics of mankind as related to the problem of what percentage of incident energy is absorbed by the exposed person. But little is known about the thermoregulation of the human body and its effectiveness to combat internal heat sources. There are indications that other important biological effects may be evoked by a mechanism of a nonthermal character (Schwan [38]). Pertinent basic research necessitates cooperation with physical chemists, biochemists, and neurophysiologists since protein denaturation, evoked nerve-response characteristics, and physical forces on microscopic particles are probably all involved. This is a very wide, new and interesting field of research with many exciting opportunities for those interested. It is obvious from the above discussions that a detailed knowledge of the dielectric properties of biological material is a prerequisite for all those who intend to do quality work in this area.

References

[1] P. Debye, "Polar Molecules," The Chemical Catalog Co., New York, N. Y.; 1929.
[2] F. Perrin, "Mouvement Brownien d'un ellipsoide dispersion diélectrique pour des molécules ellipsoidales," *J. Phys. Radium*, vol. 5, pp. 497–511; October, 1934.
[3] J. C. Maxwell, "A Treatise on Electricity and Magnetism," Oxford University Press, London, Eng., 3rd ed.; 1892.
[4] K. W. Wagner in "Isolierstoffe der Elektrotechnik," H. Schering, Ed., Springer-Verlag, Berlin, Ger., ch. 1; 1924.
[5] H. Fricke, "A mathematical treatment of the electric conductivity and capacity of disperse systems," *Phys. Rev.*, vol. 26, pp. 678–681; November, 1925.
[6] H. Fricke, "The complex conductivity of a suspension of stratified particles of spherical or cylindrical form," *J. Phys. Chem.*, vol. 59, pp. 168–170; February, 1955.
[7] H. Fricke, "The electric permittivity of a dilute suspension of membrane-covered ellipsoids," *J. Appl. Phys.*, vol. 24, pp. 644–646; May, 1953.
[8] H. Fricke, "The Maxwell-Wagner dispersion in a suspension of ellipsoids," *J. Phys. Chem.*, vol. 57, pp. 934–937; December, 1953.
[10] V. H. Pauly and H. P. Schwan, "Ueber die Impedanz einer Suspension von kugelformigen Teilchen mit einer Schale," *Z. Naturforsch.*, vol. 14b, pp. 125–131; 1959.
[11] E. Bozler and K. S. Cole, "Electrical impedance and phase angle of muscle in rigor," *J. Cellular and Comp. Physiol.*, vol. 6, pp. 229–241; June 20, 1935.
[12] H. P. Schwan, "Electrical properties of tissue and cell suspensions," in "Advances in Biological and Medical Physics," Academic Press Inc., New York, N. Y., pp. 119–289; 1957.
[13] H. L. White, E. Monaghan, and F. Urban, "Stream potentials and dc surface conductivities in small capillaries," *J. Phys. Chem.*, vol. 40, pp. 207–214; February, 1936.
[14] A. J. Rutgers and M. deSmet, "Electrosmosis, streaming potentials and surface conductance," *Trans. Faraday Soc.*, vol. 43, pp. 102–111; March, 1947.
[15a] H. Fricke and H. J. Curtis, "The determination of surface conductance from measurements on suspensions of spherical particles," *J. Phys. Chem.*, vol. 40, pp. 715–722; June, 1936.
[15b] H. Fricke and H. J. Curtis, "The dielectric properties of water-dielectric interphases," *J. Phys. Chem.*, vol. 41, pp. 729–745; May, 1937.
[16] H. P. Schwan and J. Maczuk, "Electrical relaxation phenomena of biological cells and colloidal particles at low frequencies," *Proc. First Natl. Biophysics Conf.*, Yale University Press, New Haven, Conn., pp. 348–355; 1959.
[17] K. S. Cole and R. H. Cole, "Dispersion and absorption in dielectrics," *J. Chem. Phys.*, vol. 9, pp. 341–351; April, 1941. See "1. Alternating Current Characteristics."
[18] K. S. Cole, "Alternating current conductance and direct current excitating of nerve," *Science*, vol. 79, pp. 164–165; February, 1934.
[19] C. P. Smyth, "Dielectric Behaviour and Structure," McGraw-Hill Book Co., Inc., New York, N. Y.; 1955.
[20] H. P. Schwan, "Eine genaue Methode zum Studium der Verteilung von elektrischen Relaxationszeiten, angewandt auf das System Wasser," *Z. Naturforsch.*, vol. 9a, pp. 35–37; 1954.
[21a] J. L. Oncley, "The investigation of proteins by dielectric measurements," *Chem. Rev.*, vol. 30, pp. 433–450; June, 1942.
[21b] J. L. Oncley in "Proteins, amino acids and peptides as ions and dipolar ions," E. J. Cohn and J. T. Edsall, Eds., Reinhold Publishing Co., New York, N. Y., ch. 22; 1943.
[22] H. Fricke, "A mathematical treatment of the electric conductivity and capacity of disperse systems," *Phys. Rev.*, vol. 26, pp. 678–681; November, 1925.
[23] See [12], pp. 192–197.
[24] H. P. Schwan and K. Li, "Electrical properties of free and bound water," in "Program and Abstracts," The Biophysical Society, M.I.T., Cambridge, Mass., Abstract G3; 1958.
[25] E. H. Grant, "Dielectric methods of estimating protein hydration," *Phys. Med. Biol.*, vol. 2, pp. 17–28; July, 1957.
[26] K. S. Cole, "Electro-Ionics of Nerve Action," Naval Medical Research Institute, Natl. Naval Medical Center, Bethesda, Md., Lecture and Review Series No. 54-6; December, 1954.
[27] H. P. Schwan and E. L. Carstensen, "Dielectric properties of the membrane of lysed erythrocytes," *Science*, vol. 125, pp. 985–986; May, 1957.
[28] K. S. Cole, "Ions, potentials, and the nerve impulse," in "Electrochemistry in Biology and Medicine," T. Shedlovsky, Ed., John Wiley and Sons, Inc., New York, N. Y., pp. 121–140; 1955.
[29] A. L. Hodgkin and A. F. Huxley, "Quantative description of membrane current and its application to conduction and excitation in nerve," *Proc. Cold Spring Harbor Symposia on Quantitative Biology*, vol. 17, pp. 43–52; June, 1952, and *J. Physiol.*, vol. 117, pp. 500–544; August, 1952.
[30] K. S. Cole, X. Antosiewicz, and P. Rabinowitz, "Automatic Computation of Nerve Excitation," Natl. Bur. Standards, Rept. 4238; July, 1955.
[31] H. Fricke, H. P. Schwan, V. Bryson, and K. Li, "A dielectric study of the low-conductance surface membrane in E. coli," *Nature*, vol. 177, pp. 134–135; January, 1956.
[32] K. S. Cole, "Impedance of single cells," *Tabulae Biologicae (Cellula)*, vol. 19, pp. 24–27; 1942.
[33] H. P. Schwan, "Microwave impedance determination and protein hydration," Advance Abstracts of Second Intermediate Conference on Medical Electronics, Paris, France; 1959.
[34] O. H. Schmitt, "Dynamic negative admittance components in statically stable membranes" in "Electrochemistry in Biology and Medicine," T. Shedlovsky, Ed., John Wiley and Sons, New York, N. Y., pp. 91–120; 1955.
[35] H. P. Schwan, "Biophysics of diathermy" in "Therapeutic Heat," S. Licht, Ed., Phys. Med. Library, E. Licht, New Haven, Conn., vol. 2, pp. 55–115; 1958.
[36] H. P. Schwan and K. Li, "Hazards due to total body irradiation by radar," Proc. IRE, vol. 44, pp. 1572–1581; November, 1956.
[37] H. P. Schwan, "Survey of Microwave Absorption Characteristics of Body Tissues," Rome Air Dev. Center, Air Res. and Dev. Command, Rome, N. Y. ASTIA Doc. No. AD 131-477; July, 1958.
[38] Schwan, H. P., "Molecular Response Characteristics to Ultra-high Frequency Fields," Rome Air Dev. Center, Air Res. and Dev. Command, Rome, N. Y. ASTIA Doc. No. AD 131-477; July, 1958.

General References

Some general references are suggested for purposes of introduction and greater detail.

C. J. F. Bottcher, "Theory of Electric Polarization," Elsevier Press, Houston, Tex.; 1952.
H. Froelich, "Theory of Dielectrics," Oxford University Press, London and New York; 1949.
P. Debye, "Polar Molecules," The Chemical Catalog Co., New York, N. Y.; 1929.
C. P. Smyth, "Dielectric Behaviour and Structure," McGraw-Hill Book Co., Inc., New York, N. Y.; 1955.

The four references above provide excellent and detailed introductions to the field of dielectric properties, particularly those of polar molecular origin.

H. P. Schwan, "Electrical properties of tissue and cell suspensions," in "Advances in Biological and Medical Physics," Academic Press Inc., New York, N. Y., vol. 5, pp. 119–289; 1957.
"Electrical properties measured with alternating current: body tissues," in "Handbook of Biological Data," W. S. Spector, Ed., Table 279, NAS-NRC, ASTIA Doc. AD 110501.

These provide detailed numerical impedance data of tissues at various frequencies.

H. P. Schwan, "Biophysics of diathermy" in "Therapeutic Heat," S. Licht, Ed., Phys. Med. Library, E. Licht, New Haven, Conn., vol. 2, pp. 55–115; 1958.

This article considers, in detail, the implications of impedance work for the area of diathermy as employed by the medical profession (physical medicine).

Comments on Microelectrodes*

R. C. GESTELAND†, B. HOWLAND‡, J. Y. LETTVIN†, AND W. H. PITTS||

Summary—**Metal-filled microelectrodes are best for high-frequency work; fluid-filled ones are best for low frequencies and dc. Both have advantages and drawbacks. This paper gives the results of experience with both sorts of probe. Practical hints and recipes are included because these seldom appear in detail.**

Introduction

Physiological amplifiers pose no exotic problems. They are designed to operate within the dc to 100-kc band and are usually required to have high input impedance, low current issuing into the input, and low drift. Many excellent models exist. What is of interest, however, is the nature of the coupling to the tissue. Existing microelectrodes are of two sorts, as pointed out by Svaetichin [1]: low-pass electrolyte-filled pipettes and high-pass metal-filled pipettes. There is no universal electrode. Under conditions where one is interested in propagated signals only as signals, we have found the metal probe to be most useful. When one is interested in membrane processes the fluid probe is best. A metal tip, while it will penetrate and record from a cell body, is useless for interpreting membrane potential and shape of the cell spike. The fluid tip will not see external triphasic spikes of vertebrate axons above the noise level. The complementary nature of the two probes is perhaps best given by stating that the fluid pipette is good for intracellular records and other slow changes of potential elsewhere, and the metal tip is good for extracellular records and rapid changes. In a device which handles the former, one is most concerned with low current through the probe (*i.e.*, low grid current), very high input impedance, and low drift, and will sacrifice for these features the upper frequency response and a low noise level. In a device which handles the latter, one is most concerned with low noise and good high-frequency response and will sacrifice input impedance and drift, and avoid electrode current by capacitative coupling.

We shall show some measurements on metal electrodes in inorganic electrolytes and body fluids, and shall also present a new recipe for a metal microelectrode.

* Original manuscript received by the IRE, September 3, 1959. This work was supported in part by the U. S. Army (Signal Corps), the U. S. Air Force (Office of Sci. Res., Air Res. and Dev. Command), and the U. S. Navy (Office of Naval Res.), and in part by Bell Telephone Labs., Inc.

† Res. Lab. of Electronics and Dept. of Biology, Mass. Inst. Tech., Cambridge, Mass.

‡ Lincoln Lab., Mass. Inst. Tech., Lexington, Mass.; formerly with Res. Lab. of Electronics, Cambridge, Mass.

|| Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Mass.

Considerations on the Use of Metal Microelectrodes

It is usual in physical chemistry to distinguish two extreme kinds of electrodes. The first is the reversible type, in which ions from the solution are actually charged and discharged, so that a steady current is possible and the dc potential of the electrode has a well-defined value depending on the current and the composition of the solution. The second type is the ideal polarized electrode, in which no transformation of ions takes place, no steady current can pass, and such as does represents the charging and discharging of a double layer made up of the electrode and ions very close to its surface, forming a structure which acts like a capacitance whose value is dependent on the voltage across it. This electrode has no well-defined dc potential; the latter may vary wildly under apparently identical circumstances, and is enormously influenced by traces of impurities. Actual electrodes are always combinative of both types, and their impedance as a function of frequency shows the extent to which one or the other mechanism dominates their behavior, as may be seen in the figures.

These measurements were taken with a special bridge, designed by Howland and later modified considerably [2]. Both bridge circuits were direct reading in series-equivalent resistance and capacitance, and included provision for canceling the effect of the external capacitance to ground of the microelectrode and the wire connecting it. Only a very small ac signal, generally less than 50 mv, was applied to the electrode; more would have produced harmonic distortion and/or irreversible change in the surface of the electrode.

A bright Pt electrode in the spinal cord or in spinal fluid, both rich in adsorbable compounds which presumably greatly retard oxidations and reductions at the interface, behaves like a capacity even at low frequencies; its impedance is then proportional to ω^{-1}, over a very wide range (Fig. 1) (actually $\omega^{-0.95}$ in the figure).

Another electrode behaves in a more complicated way. Thus it will often happen that the electrochemical reaction is itself quite rapid, and the current is limited by the diffusion of the reacting ions and the products between the surface of the electrode and the bulk of the solution. In this case the impedance for high frequencies will decline as $\omega^{-1/2}$, as appears in the case of bright Pt in physiological saline. If the electrochemical reaction is itself slow, the impedance will be lower for very small ω and higher for very large ω than in cases where diffusion is predominant (Fig. 2, R_{Pt}, $X_{C_{Pt}}$).

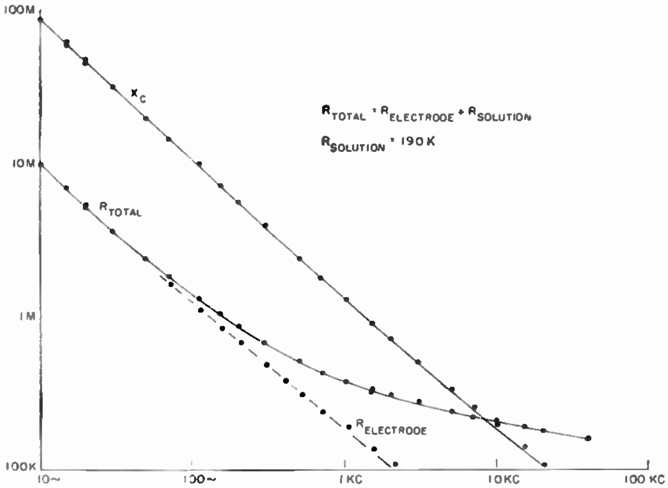


Fig. 1—Variation of impedance with frequency, series resistance-capacitance measure. Bright platinum (851) 10μ (diameter) electrode in spinal cord. Electrode is 90° cone.

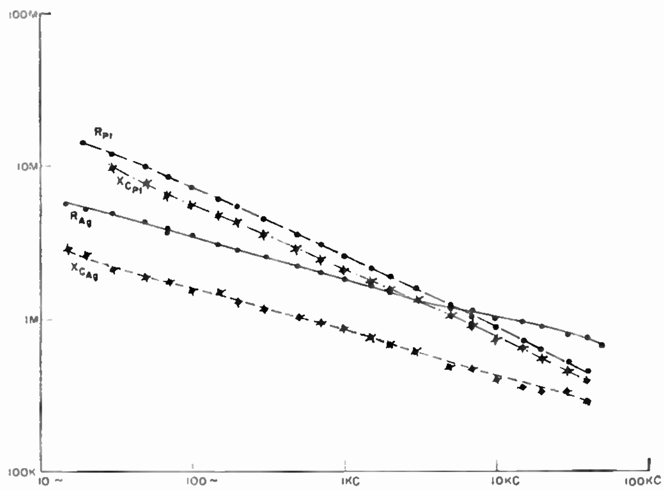


Fig. 2—Variation of impedance with frequency, series resistance-capacitance measure. Bright platinum cone and Ag-AgCl cone adjusted to same surface area, about 50 μ^2. Both tips are in physiological saline.

In practical cases, it may happen that inhomogeneities in the electrode material will spread out the band of frequencies for which the impedance declines only slowly, since the rate of the reaction is much dependent on the dc potential of the electrode, and this will vary over the surface. In cases where an insoluble reaction product covers the electrode, the multiplicity of diffusion paths of different length may have a similar effect. Such a case is the Ag-AgCl electrode in Fig. 2, whose impedance is approximately proportional to $\omega^{-1/4}$ over a wide range and independent of the presence of gelatin. The dc potential of the electrode can make a considerable difference in the impedance curve, by changing the rate or even the nature of the reaction carrying current, and therefore the balance between it and diffusion as responsible for the impedance.

The importance of these curves in selecting an electrode for a particular study rests upon two sorts of consideration, distortion and noise.

Distortion of a Metal Electrode

The response of a metal electrode to a source of current suddenly appearing and remaining constant in the medium can be guessed roughly from the well-known theorem on Fourier transforms, which gives the result that if

$$Z(\omega) \to \frac{A}{\omega\gamma}, \qquad \gamma < 1$$

as $\omega \to 0$ or $\omega \to \infty$, then the integrated current

$$\int_0^t I(t)dt \to \frac{1}{A\Gamma(2-\gamma)}\, t^{1-\gamma},$$

as $t \to \infty$ or $t \to 0$, respectively. A and γ are empirical constants. If the current is unidirectional,

$$I(t) \to \frac{1}{A\Gamma(1-\gamma)}\, t^{-\gamma}$$

for $t \to \infty$ and $t \to 0$. Thus the diffusion electrode would report a signal proportional to $t^{-1/2}$ for both short and long times, the Ag-AgCl one proportional to $t^{-1/4}$ for short times, and also for long, if the behavior of $Z(\omega)$ continued to be the same for low frequencies. In fact, of course, the relevant Z is

$$\frac{1}{1 + \dfrac{Z_E}{Z_L}},$$

when Z_E is the electrode impedance and

$$Z_L = \left(\frac{1}{R_L} + j\omega C_L\right)^{-1}$$

is that of the external path to ground, so that the response of the Ag-AgCl electrode, $Z_E \sim A\omega^{1/4}$ to the square wave would behave initially like

$$(AC_L)^{-1}\Gamma(\tfrac{1}{4})^{-1}t^{-3/4}.$$

Actually $Z(\omega)$ must be bounded for $\omega = 0$ since a steady dc current is possible in this electrode, but the recorded signal should still decline slowly to its asymptote as $t^{-3/4}$.

Noise of a Metal Electrode

The rms noise voltage generated by a metal microelectrode in a given narrow frequency band is most easily specified in terms of the equivalent noise resistance at room temperature, R_N the relation between these quantities being:

$$E_{\text{rms noise}} = \sqrt{4kTR_N\Delta F}.$$

Here k is Boltzmann's constant and T is the absolute temperature. Now, Nyquist's Thermal Noise Theorem states that R_N should be equal to the real part of the electrode impedance at the same frequency. The proof of this theorem assumes that the system is in thermal equilibrium, which would in this case require that the current through the electrode be zero, and that there be no chemical reaction taking place at the surface of the electrode. (These conditions are difficult to satisfy in practice.)

The equivalent noise resistance of a metal microelectrode as a function of frequency was measured by means of a resistance substitution method, using interpolation, for a number of frequency bands defined by a Spencer-Kennedy variable band-pass electronic filter. The noise output of the electrode was suitably amplified, filtered, and measured by a thermistor bridge-type rms voltmeter.

Results of these measurements as a function of frequency for a bright platinum microelectrode in normal saline solution are shown in Fig. 3. Good agreement between the equivalent noise resistance and the real part of the electrode impedance is evident. Similar agreement was also obtained for a Ag-AgCl microelectrode also in normal saline solution. Little if any noise in addition to the expected thermal noise was generated by the microelectrode.

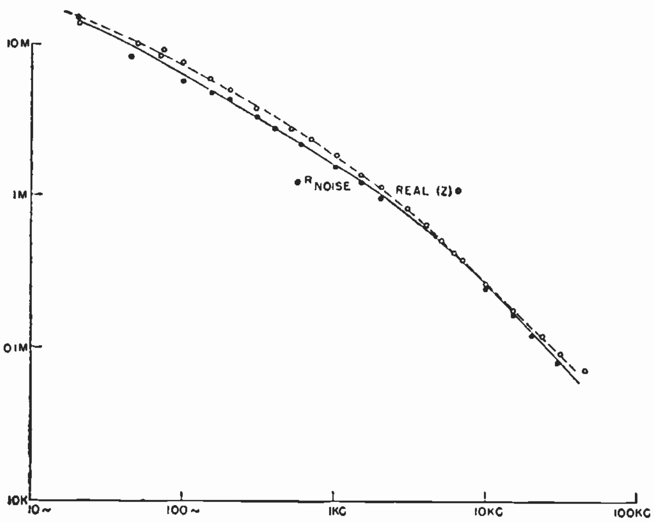


Fig. 3—Variation of impedance with frequency, series resistance-capacitance measure. Bright platinum (851) conical tip, 10μ in diameter, 5μ high, in physiological saline. Noise measure by substitution method (see text).

DC Measurement with Metal Tips

We ought to say something about the use of microelectrodes for measuring dc or classical thermodynamic potentials in living tissue. The conditions here are so difficult that not one of the classical reversible electrodes has ever been shown to perform its office with even modest reliability. The Ag-AgCl fails to measure Cl^-, apparently for a number of reasons. First, the solid AgCl diffuses away from the tip of a microelectrode with great speed and can hardly be continuously formed with an imposed current. This will be the bother generally for very small electrodes requiring saturation with a sparingly soluble salt, since the time of diffusion varies with the square of the linear dimensions of similar structures. Second, silver forms extremely stable complexes with organic molecules having attached amino and sulfhydril groups which occur in plenty where the electrode locally damages tissue. The stability of almost all of these is greater than the insolubility of AgCl. This difficulty is more important than the other, for it applies to almost all heavy metals one might hope to use: Hg, As, Sb, Bu, Pb, Cd, Cu, etc. Finally, the reduction-oxidation potential of axoplasm is low enough to reduce methylene blue, which places it below hydrogen. At this potential AgCl is reduced to metallic Ag and Hg_2Cl_2 to Hg, the insolubility not being adequate to the nobility of the metals. (AgI might possibly be capable of existence, and AgS is capable, under these circumstances; but sulfide ions could hardly exist at the pH of axoplasm. In fact ingested silver is deposited as metal in tissue, as in argyria.) This difficulty also applies to other heavy metals and to oxide electrodes (Sb and Bi) used to measure pH.

One may note that if the Ag-AgCl electrode were actually reversible to Cl^-, since this ion can move freely across the membrane and is in fact distributed in equilibrium with the membrane potential, such an electrode should not record the membrane potential at all. This is also true with a true pH electrode. In one instance it has been reported that an Ag-AgCl microelectrode measures membrane potential. Such a result seems to confirm that an Ag-AgCl microelectrode becomes a noble metal electrode within a cell.

The inert, or noble metal electrode, should be free of these difficulties and measure the reduction-oxidation potential of the solution. The difficulty here is that the oxidations and reductions only take place through enzyme-substrate complexes of enormous molecular weight, activation energies being far too high to permit any equilibrium to be attained otherwise; the consequence is an impedance far too high to make the small electrodes usable at dc or very low frequencies even with the best control of grid current. If there were a rapidly reversible system present to buffer the potential, such as $Fe^{++} \rightleftarrows Fe^{+++}$, this could be avoided, but no inorganic system seems usable, and organic ones seem to be all of the quinonoid type which is adsorbed to proteins with great avidity.

Recipe for an Easily Made Metallic Microelectrode

Our study of probes has been cursory at best. Nevertheless we feel that metal microelectrodes are seldom good at low frequency and are excellent at high. So we have tried to devise a probe that would be easy to make and would have extremely low impedance in the upper physiological band, 1–50 kc. The reason for wanting such a device is that clearly we would have a more favorable signal-to-noise ratio if the admittance could somehow be increased. Thus we might record more easily the external current of nerve fibers which, in unblocked axons, is the second derivative of the traveling spike—the familiar triphasic transient, a very much smaller signal than the membrane spike.

Svaetichin [1] and Dowben and Rose [3] have plated electrodes with Pt black to increase the surface-volume ratio at the tip. This seemed the best idea and we tried the method. Very soon we found an odd property attaches to Pt black in tissue. Having an enormous surface, it catalyzes local reactions, and, quite literally, burns onto itself a shell of very adherent stuff. Thus, although such an electrode may show a moderately low impedance at high frequencies, it is kept from intimate contact with the tissue around it by this shell. The gain from working at low impedance is offset by the lack of intimacy. Such an electrode records very well when close to or touching cells, or if it has blocked a nerve fiber, but then the signals are so large that any electrode will do.

We set about redesigning Dowben and Rose's electrode. First we discovered that gold flashing is not necessary. Second we found that if we add gelatin to the chloroplatinic acid bath from which we plate the Pt, the ball is not only made adherent to the tip but is, in a sense, prepoisoned and does not burn a shell onto itself. Finally, as is often recommended, we used a little lead acetate in the bath to make the Pt deposition uniformly fine.

In constructing the metal-filled pipette we used the following procedure. Woods metal was doped with enough indium to give a slight shrinkage on cooling. The resulting alloy melted well below the boiling point of water, and the molten metal could be sucked up in ordinary polyethylene tubing (spaghetti) which slipped off easily after the metal was cooled. Thus, we could make a wire of the alloy. We pulled a glass micropipette of the usual sort, and, under a microscope, broke off the tip until the O.D. was about 2–3 μ. Then we put some of the alloy wire in the shank and pushed it as far down as it would go with a piece of copper wire just thin enough to fit into the shank. Placing the pipette on a hot plate just warm enough to melt the alloy we pressed constantly against the copper wire. The alloy softened and became fluid except where it touched the copper wire which had a much higher heat conductivity. During the period when the alloy was hard at the copper wire and fluid at the upper end of the pipette taper, the alloy acted both as its own plunger and fluid, and under the constant pressure applied to the copper wire flowed down the pipette to the tip. Then the alloy around the copper wire melted and the copper became imbedded in it. When the pipette was removed from the heat the alloy was found to have retracted from the tip or to have formed a vacuole along the taper. When the tapered part of the pipette was then held close to some source of heat (say a quarter inch away from a 60-watt bulb) for about 5 seconds, the tip was found to be filled or the vacuole gone.

These electrode blanks can be made at the rate of one in three minutes. The Pt ball is plated on just before use. The bath we used was 1 per cent chloroplatinic acid and 0.01 per cent lead acetate. 100 cc of this solution was added to 2 cc of viscous but ungelled gelatin solution at room temperature. The current for plating was obtained from a 1.5-volt cell through a variable 1-megohm resistor having a least resistance of 20 k. Contrary to Svaetichin's statement, bubbles need not form. The plating should be done under direct vision through a microscope.

The impedance curve for one of these gelatinized balls about 3 μ in diameter in normal saline is shown in Fig. 4.

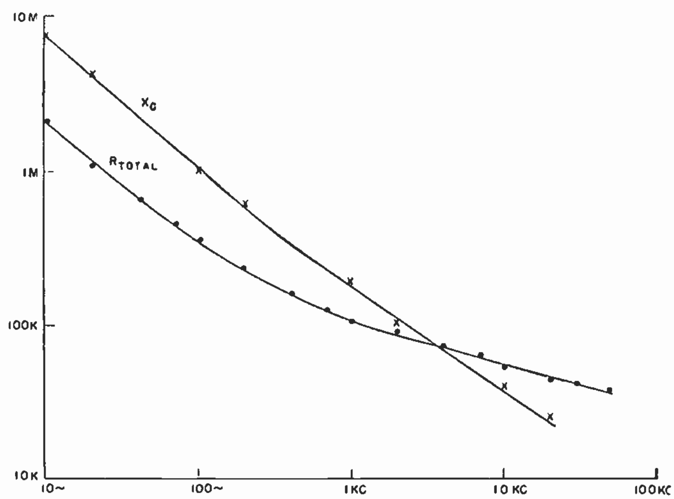


Fig. 4—Variation of impedance with frequency, series resistance-capacitance measure. Platinum black ball, about 3μ in diameter, in physiological saline.

When it is compared with the curve from a bright Pt 90° cone 10 μ in diameter in a protein solution (Fig. 1), it is noted that at 10 kc the real component is less than 100 KΩ. The high admittance in the 1–100-kc band has allowed us to detect single unmyelinated fibers in the optic nerve well above the noise level. The impedance varies almost as $(\omega)^{-1}$ just as it does for bright Pt in protein media but the curve has been translated downward. We found such a probe is biased in either direction well over one volt by a polarizing current of 10^{-10} amp.

We use simply a cathode follower of high transconductance, couple the electrode to the grid by 0.03 μf, and shunt the grid to ground by 10 megohms. Thus, we achieve very good external records from single nerve fibers, for at the high frequencies of the triphasic spike the admittance of the electrode is very high. We know that an electrode is operative if the noise at the input is much reduced when the electrode makes contact.

Recipes and Other Miscellany for Fluid-Filled Pipettes

The complementary electrode, or low-pass type, is the electrolyte bridge—the familiar KCl micropipette. This sort of probe is so well known as to need no exegesis.

Glass micropipettes are made in a standard fashion whether one intends to fill them with metal or saline. They are seldom drawn by hand, as originally described by Ling and Gerard [7], except when one needs odd bends or tapers. Alexander and Nastuk [8] developed an electrode puller which has become available commercially in several models. It is simply a solenoid and heating coil so arranged that when a glass tube is held at one end, extends through the coil, and is gripped at the other end by the solenoid, the softening of the tube under the coil is detected by the solenoid pulling gently (*i.e.*, having an appropriate series resistance) until the tube has lengthened enough to close a switch shorting out the series resistance and applying the maximal pull of the solenoid. The softened glass tube extends in one of Plateau's [9] volumes of minimal surface, the unduloid (modified of course by the high viscosity of glass), and the appearance of the tip of a glass micropipette is that of a series of constrictions or steps of decreasing period superimposed on a continuous narrowing. The glass breaks at the smallest constriction while the glass is still plastic, and so it is not easy to draw consistently anything but the finest tip diameters (less than 1 μ). The period and amplitude of the undulations leading to the tip are functions of the heating and pulling program. Control of the period by suitable adjustment allows one to set the penultimate constrictions before the tip to be whatever diameter one chooses over a wide range. The ultimate cycle or cycles of glass can be broken off under the microscope, the break generally occurring cleanly and circularly at the constriction one chooses.

In making fluid-filled pipettes it is usually thought proper to leave the last glass cycles on, for the finer the tip the less damage it does in penetrating nerve or muscle membrane. Filling these pipettes with KCl solution used to be considered difficult because of the trapping of bubbles in the tips. Several recipes for filling have been published and they all work. Our method is not very much better than the others, but it has the advantage that electrodes are ready for use a few minutes after manufacture. Several micropipettes are mounted, tip down, on a glass spindle and are held to it by a rubber band. KCl solution (slightly less than 3 molar to allow for evaporation) is heated in a side-arm flask to about 80°C and then the spindle is inserted into the hot solution, pipette tips down, until the pipettes are totally immersed, whereupon the top of the flask is stoppered. We make the spindle and stopper as one piece. Suction is then applied to the solution until it boils moderately vigorously. Since the boiling bubbles come off most easily at irregular surfaces, they appear at the tops of the pipettes and at the rubber band. If the flask of hot solution has been removed from the heat and made to sit on a surface at room temperature for a few minutes before suction is applied, then indeed it is almost impossible to get boiling bubbles from below the tips of the electrodes, *i.e.*, from the bottom of the flask. Thus the danger of tip breakage, such as occurs with direct boiling over a flame, is minimized. After about 20 seconds of moderate vacuum boiling, suction is cut off and air readmitted, then suction is applied again, and in this way we go through three cycles of low-pressure boiling. The notion behind this procedure is this: Most of the air in the pipette will be exhausted in the first boiling and be replaced with steam. On readmission of normal pressure most of the electrode will fill, leaving an entrapped air bubble. This bubble will be again diminished in the next cycle because it dissolves in the hot solution bounding it under normal pressure and is evacuated by the turbulence of reboiling. Three boiling cycles appear to fill nineteen out of twenty electrodes. The spindle can then be transferred to a cooler solution, or, more reasonably, the whole flask cooled under running tap water for a few minutes.

The impedance of micropipettes with electrolyte core is that of a moderately large resistance shunted by a small capacitance. Since the resistance is concentrated in the tip and the medium around the tip, the distributed capacitances can be disregarded and the shunt capacitance to the grounded medium around the electrode, and to ground from the connecting lead and the grid can be lumped together. If the electrode couples to the tissue with 10 megohms resistance, a shunt of 10 $\mu\mu$f will degenerate the recorded signals quite seriously for there are 50-kc components in an intracellularly recorded cell spike.

Since the physiologist is very concerned with the exact shape of a transient, the signal is usually regenerated from an amplifier having in-phase gain greater than unity through a trimmer capacitor back to the grid of the input tube. This positive feedback is called a "negative capacitance" and it has been exhaustively handled by several authors. We particularly recommend Amatniek's paper [4]. There are many such commercial devices already available, most of them good. The principle behind them is to use an electrometer tube at the input to insure low grid current and high input impedance initially, and to amplify the output of the tube by a factor of 2–3, keeping the output in phase with the original input to as high a frequency as possible. The output of the amplifier is fed back to the grid of the electrometer through a variable trimmer. This raises the effective input impedance at high frequency.

Such a method of recording is inherently noisy, for the signal-to-noise ratio at the input is not changed by the feedback. Thus the more the signal is degenerated without the negative capacitance, the more noisy will be the restored signal. With this compensation the noise always increases with the frequency [5]. The strategy of using such an amplifier is to arrange for absolutely minimum shunt capacitance at the input initially. That is, the lead from electrode to input ought to be as short as possible and any grounded material should be kept reasonably far from the lead and grid. Our version of this device, shown in Fig. 5, has no greater virtues than other circuits—in fact the parts are more expensive—but a student can throw one together out of hand and the transistors, being silicon, are pretty stable. If one wants extremely high dc stability it is best to chopper-stabilize as Moore [6] has done.

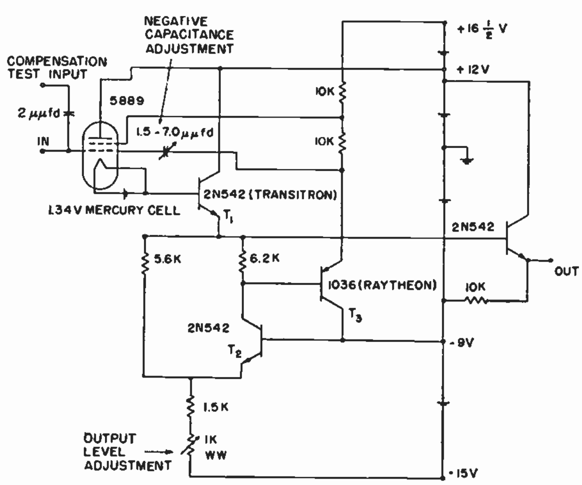


Fig. 5—Circuit for headstage.

Use of Square Pulses of Current Through Fluid-Filled Pipettes

Of course one should not use a compensation scheme without being able to determine if it is properly adjusted. A simple method of monitoring an electrode is to use a ramp of voltage (triangular or trapezoidal waves) through a small capacitance, say 1–2 $\mu\mu$f to the grid, as designed by Howland [5]. The recorded derivative, of course, is a square pulse whose height measures the resistance of the electrode and whose squareness at the corners measures the compensation. One can adjust the ramp to occur, say, at the end of every sweep of an oscilloscope and change the sign of the ramp with each sweep. Without such a monitor on the resistance and reactance one can get into all kinds of difficulty, as for example in the reports of "synaptic noise" at a membrane surface which generally is nothing more than the thermal noise of an electrode sitting up against a very high impedance such as a patch of membrane.

The ramp function can be used not only to test patency and compensation for a fluid-filled pipette, but also to stimulate through the same electrode that is used to record. If the square pulses of current are not very large or long the resistance of the fluid-filled tip will remain reasonably constant, rectification will enter only as a second-order effect, and the pulses can then be balanced out in the amplifier by simple subtraction of a square pulse of voltage. Suppose one is using a micropipette in a dorsal or ventral root of a cat's spinal cord. It penetrates large fibers easily and reliably and the records stay constant for between 30 minutes and 1 hour. We have evidence that these penetrations only occur at nodes at Ranvier, but, to avoid controversy, let us remark only that penetration of myelinated fibers occurs. In a large mammalian axon the current necessary to set off a single node, whether it is directly under the electrode or decoupled by the internodal resistance, is rarely larger than 2×10^{-9} amp delivered for 0.1 msec. The artefact, 20 mv at most through a 10-megohm electrode resistance, is easily balanced out, after compensation, by subtraction of an appropriate square pulse, leaving only a small make-and-break transient due to the uncompensatable distributed capacitance of the electrode tip and delay in the circuit of the headstage. Thus one can measure electrotonus, local response (damped nonlinear behavior) and many other properties of axons. It is an especially good and reliable demonstration for students.

One can also use an ac signal, say 5 kc, through the small capacitor to the grid, and balance it out in the amplifier, *i.e.*, by subtracting it after appropriate phase shift and amplitude setting. Such a virtual bridge measures the changes which occur primarily in the immediate tip environment. Thus one can show, with this arrangement, that mammalian axoplasm changes its resistivity late on the falling phase of the action potential, just as Strickholm found it does for myoplasm [10].

Finally, the use of square pulses of current to measure threshold successfully distinguishes fibers from cells, at least in the spinal cord. Elementary computations show that if one assumes threshold current density to be the same for all nerve membrane then there should be about a 50:1 difference between the current necessary to excite the smallest cell and that to excite the largest node, given a local source of current well within the neuron. Since in fact it seems that the cell membrane has a higher threshold than axonal membrane, the discrepancy should be even more pronounced. We have found this difference useful, although in practice not as extreme as in the calculation. No cell, as determined by shape of spike, etc., seemed to be excited with currents below 5×10^{-8} amp, and no axons required more than 3×10^{-9} amp at about 0.15 msec duration. Most units one penetrates in the dorsal half of the cord seem to be axons by the criterion, according to Wall [11]. This would explain why many investigators cannot see "synaptic potentials" in interneurons. Incidentally, the higher currents needed to stimulate cells cannot be balanced out at all well because the impedance at the tip fluctuates drastically. These impedance changes return to normal only very slowly.[1]

Conclusions

Comparing the two sorts of electrode, metal and fluid, it is easy to see that the signal-to-noise ratio for high frequencies (such as appear in the externally recorded triphasic spike of single fibers) is extremely poor for the fluid tips even with compensation, and extremely good for the spongy metal tip. It is also easy to see that at low frequencies the fluid tip is far superior to the metal; however, anyone who has attempted to pass even small ac current through a micropipette filled with saturated KCl solution and inserted into tissue also knows that these electrodes rectify at low frequencies as would be expected.

Therefore we advise that nervous physiologists do not try to make one electrode do the work of the other but use both types and with two different head-stages. The electrometer tube is essentially noisy because of low transconductance, and has poor high frequency response itself, even with compensation. However, it is the only good input coupler for the fluid pipettes because of its low grid current and high input impedance. This system is best used for intracellular work. The conventional high-transconductance cathode follower has little noise, but much grid current. However, if it is used only in the useful frequency range of metal electrodes by coupling the electrodes capacitatively and using a grid leak, the extracellular high-frequency signals, which cannot be seen with the fluid pipettes, come through beautifully.

[1] One should be cautious in trusting the fluid pipettes for measurement of dc potentials, however, on account of the liquid junction potential at the tip. It is true that when established in a narrow tube in simple solutions, this potential is reproducible and small; with the micropipette it can be large and variable, probably because the diffussion layer at the tip is sharp and unsteady, altering rapidly in structure and exposed to irregular hydrodynamic influences of great effect if any current is passed through. The fluid junction potential is also known to be much increased by the presence of polar colloids, particularly polyelectrolytes, which are capable of selectively reducing the mobilities and activities of ions by the strength of their ionic atmospheres.

Bibliography

[1] G. Svaetichin, "Low resistance microelectrodes," *Acta Physiol. Scan.*, vol. 24, suppl. 86, pp. 5–13; 1951.
[2] R. C. Gesteland and B. Howland, "Bridge for measuring impedance of microelectrodes," *Rev. Sci. Instr.*, vol. 30, pp. 262–264; April, 1959.
[3] R. M. Dowben and J. E. Rose, "A metal-filled microelectrode," *Science*, vol. 118, p. 22; July 3, 1953.
[4] E. Amatniek, "Measurement of bioelectric potentials with microelectrodes and neutralized input capacity amplifiers," IRE Trans. on Medical Electronics, vol. ME-10, pp. 3–14; March, 1958.
[5] J. Y. Lettvin, B. Howland, and R. C. Gesteland, "Footnotes on a headstage," IRE Trans. on Medical Electronics, vol. ME-10, pp. 26–28; March, 1958.
[6] J. Moore, private communication.
[7] G. Ling and R. W. Gerard, "Normal membrane potential of frog sartorius fibres," *J. Cell. Comp. Physiol.*, vol. 34, pp. 383–396; December, 1949.
[8] J. T. Alexander and W. L. Nastuk, "An instrument for the production of microelectrodes used in electrophysiological studies," *Rev. Sci. Instr.*, vol. 24, p. 528; July, 1953.
[9] D. W. Thompson, "Discussion on plateau," in "On Growth and Form," Oxford University Press, New York, N. Y., vol. 1, pp. 365–388; 1952.
[10] A. Strickholm, "Impedance changes during a single muscle fibre twitch," *Proc. Biophys. Soc. 1958 Meeting*, February 5–7, Cambridge, Mass., p. 18 F6.
[11] P. Wall, private communication.

Some Functions of Nerve Cells in Terms of an Equivalent Network*

WALTER H. FREYGANG, JR.†

Summary—**A distributed parameter equivalent network of a nerve cell is developed. The network is based upon the electrical constants of nervous tissue. Inserted in the network are electrically and chemically activated generators. Some experimental evidence is given for the properties of the network and the generators, as well as for the location of the generators in the network. The function of the neurons in the nervous system is discussed in terms of this network.**

* Original manuscript received by the IRE, June 8, 1959.
† Lab. of Neurophysiology, Natl. Institutes of Health, Bethesda, Md.

Introduction

The purpose of this article is to describe the electrical circuitry of a nerve cell, or neuron. The neuron is the fundamental unit of the nervous system. Thought, behavior, and consciousness are all, in some way, related to the interaction between neurons. This statement may be too mechanistic, but the number of neurons in the nervous system has been estimated as 10^{10}. After the functional potentiality of the neuron has been described in terms of an equivalent network and the extent of the connections between neurons is appreciated, the statement may seem more reasonable.

Most of the results described here have been obtained from experiments on the anterior horn cells of the cat's spinal cord. There are reasons to believe that the other neurons in the nervous system have similar mechanisms. By neglecting the relatively small impedance to the local flow of current in the tissues that surround the nerve cell and by placing generators at points in the circuit, one can draw a distributed parameter network to represent a nerve cell. The properties of the generators are special, however, and these will be defined. Some of the current questions with which neurophysiologists are concerned will be presented.

The Geometry

The geometric complexity of the nerve cell can be appreciated from the diagrammatic sketch of a neuron in Fig. 1(a). Although the demarcation between the dendrites, cell body, and axon is arbitrary, the neuron is divided into these portions for descriptive purposes.

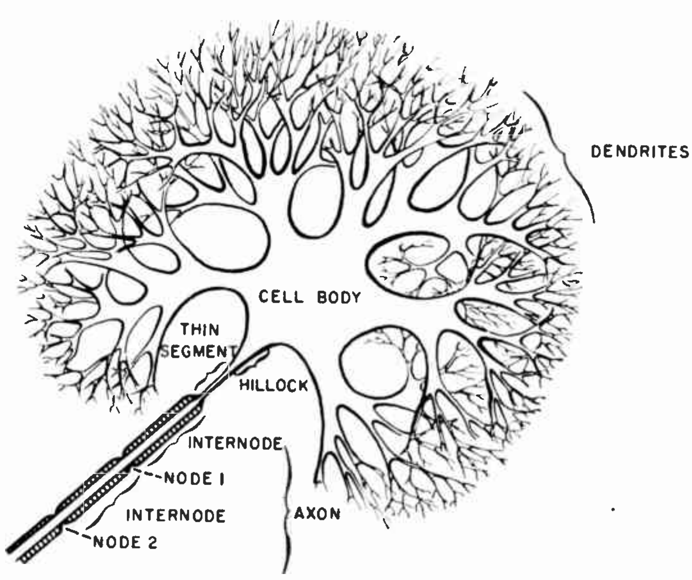
Cell Body: The cell body is the widest part of the neuron. Its diameter may be as much as 0.1 mm. The nucleus of the cell and much of the metabolic machinery is located in the cell body. At one region it tapers to form the hillock, from which the axon originates.

Dendrites: These are the long, tapering, branching structures that form the dendritic bush. The boundary line between the cell body and the dendrites is rather like that between the trunk and the branches of a tree. In any case, the fine arborizations are dendrites. They approach submicroscopic dimensions at their tips. The total surface area of the dendrites is very great because of their configuration.

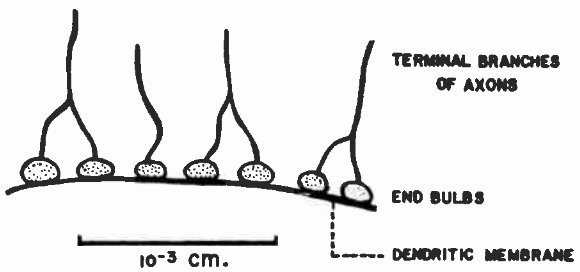
Axon: The axon begins with the constricted portion formed by the tapered end of the hillock, which is called the thin segment. Eventually the thin segment dilates and becomes covered with myelin, a fatty substance. The breaks in the myelin sheath of the axon are the nodes, and the myelin-covered portions between the nodes are the internodes. Internodes are 1–2 mm long.

The axon is the longest part of the neuron. It may extend for several feet before it branches and distributes end bulbs at the tips of its branches over the surface of the dendrites and cell bodies of other neurons. Axonal endings are not shown in Fig. 1(a) but are sketched in Fig. 1(b). The space between the end bulbs and the surface of the dendrites, or cell body, is only about 200 Å. Instead of ending on other neurons, an axon may supply terminal structures to effector organs, such as muscles, in order to activate them. The synapse is the junction of an end bulb with the neuronal surface beneath it. Interaction between neurons results from transmission across these junctions. It is believed that the end bulbs secrete substances that affect the membrane, or cell wall, of the dendrites and cell body. This effect will be described later as one type of generator.

The membrane, or surface coating of the neuron, is so thin (100 Å) that it can be considered to have an infinitesimal thickness. It is in contact with the tissues and fluids that surround the neuron, except at the internodes where it is covered with myelin. The membrane of the cell body and dendrites is largely covered by end bulbs. Fig. 2 is a photomicrograph of a large neuron. It shows a relatively large, elongated cell body and one long dendrite. The fine black dots on the surface of the neuron are end bulbs. Extensive interaction between neurons is made possible by the great number of end bulbs on each neuron.



(a)



(b)

Fig. 1—Diagrammatic sketches of a nerve cell. (a) Cell body, dendrites, and axon. (b) The endings of the terminal branches of an axon on the cell body and dendrites of another neuron.

The Electrical Constants

It is possible to insert either one or two very fine saline filled glass pipette electrodes into the cell body of a living, functioning neuron in the intact nervous system. The tips of these electrodes can be as small as 10^{-4} mm. Such electrodes have an appreciable resistance (5–50 megohms) at their tips and a distributed capacity (0.4 $\mu\mu$f/mm) across their walls which introduce some difficulties, but these problems are dealt with in another article in this issue[1] and will not be considered here. Po-

[1] R. C. Gesteland, *et al.*, "Comments on microelectrodes," this issue, p. 1856.

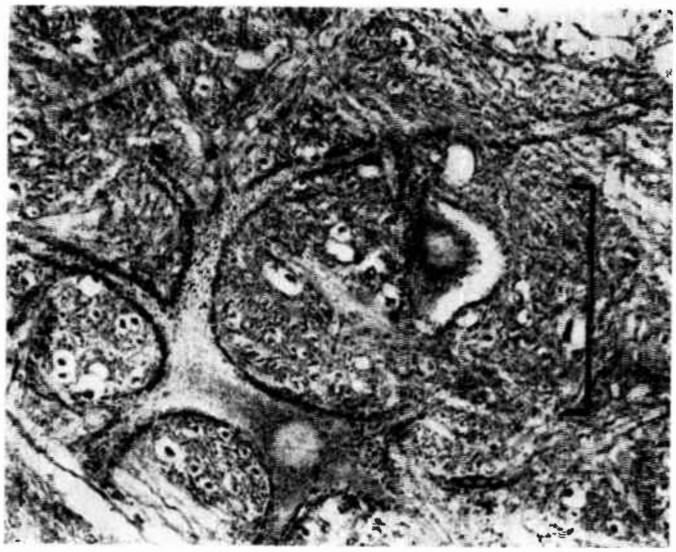


Fig. 2—Photomicrograph of part of a neuron. The small black dots on the cell body and dendrites are end bulbs. Bracket indicates 0.1 mm. (Courtesy of Dr. Grant L. Rasmussen.)

tential differences measured with such an intracellular electrode are with respect to a low resistance electrode at some point remote to the neuron. By means of a resistance-capacity-network analog of a neuron and the conducting volume that surrounds it, it has been shown that the resistance between the outer surface of the cell body and the remote electrode is so small, relative to the resistance of the membrane, that it can be neglected under most conditions. Also, results obtained with this analog show that because of the relatively large diameter of the cell body, and the low volume resistivity of the internal protoplasm (50 ohm cm), its intracellular space is practically equipotential for the biological voltage transients induced across parts of its membrane. The construction of the analog will be described later.

In accord with the above, if two pipettes are inserted in the cell body and a known amount of constant current is injected through one intracellular pipette, one can measure the potential difference between the inside of the cell body and the remote electrode by means of the second intracellular pipette. The resistance to current flow presented by the membrane of the cell body can be estimated from this type of experiment. Similarly, the time course of decay of the potential difference between the inside of the cell body and the remote electrode can be measured after the constant current is turned off. When corrected for the flow of current into the dendrites, it is found that the resistance of the membrane of the cell body is about 1 megohm and that the time constant for the exponential decay of potential difference is 4 msec.[2] (Because of its small diameter, and therefore its high input impedance, the flow of current into the thin segment is negligible.) From these data, the capacity of this membrane is calculated to be about 4000 $\mu\mu$f. Using 4×10^{-3} cm^2 as an estimate of the surface area of a cell body, the electrical constants for the flow of current across a 1 cm^2 area of membrane are 4000 ohm cm^2 and 1 μf cm^2. These values are in good agreement with the constants of invertebrate axons which, like the cell body and dendrites, are not covered with myelin sheaths.

Experiments on single myelinated axons have shown that the transverse resistance and capacity for a 1 mm length of myelinated internode are 290 megohms and 1.6 $\mu\mu$f. For a passive node these values are 40 megohms and 1.5 $\mu\mu$f. The internal longitudinal resistance of an internode is about 50 megohms. A description of how the electrical constants of axons have been estimated is beyond the scope of this article.[3,4]

The Equivalent Network

With these results in mind, one can draw an equivalent network to represent the neuron. Such a network is drawn in Fig. 3. It should be remembered that the resistances and capacities of the membrane at the internodes, thin segment and hillock, and of the dendritic membrane are continuously distributed along the core resistances of these structures. In addition, the branching of the dendrites has not been drawn schematically in Fig. 3 for the sake of simplicity. Since the dendrites taper and branch, it would be necessary to add many circuits having continuously increasing core resistances with distributed resistances and capacities along them. Note that the core resistance at the nodes has been neglected. This is permissible because a node is only 10^{-3} mm in width and therefore its longitudinal resistance is negligible. The internal resistance of the cell body has also been omitted for the reason that has been given already. With the generators, which are represented by the G's, all open-circuited, Fig. 3 is the equivalent circuit of the nerve cell that was employed to determine the electrical constants of the membrane of the cell body. In order to discuss how the neuron functions when it is not at rest, we must consider the properties of the two types of generators.

The Generators

One of the two types of generators are represented by G_1, G_2, G_A, and G_B in Fig. 3. These generate the nerve impulses that travel along the axon. Their properties are described more fully in an article in this issue by Moore[5] and will be described only briefly here.

The other type of generator is represented by G^* in Fig. 3. They act in response to transmitter substances liberated by the end bulbs during synaptic transmission.

[2] W. Rall, "Membrane time constant of motoneurons," *Science*, vol. 126, p. 454; September 6, 1957.

[3] I. Tasaki, "Nervous Transmission," Charles C Thomas, Springfield, Ill.; 1953.

[4] I. Tasaki, "New measurements of the capacity and the resistance of the myelin sheath and the nodal membrane of the isolated frog nerve fiber," *Am. J. Physiol.*, vol. 181, pp. 639–650; June, 1955.

[5] J. W. Moore, "Electronic control of some active bioelectric membranes," this issue, p. 1869.

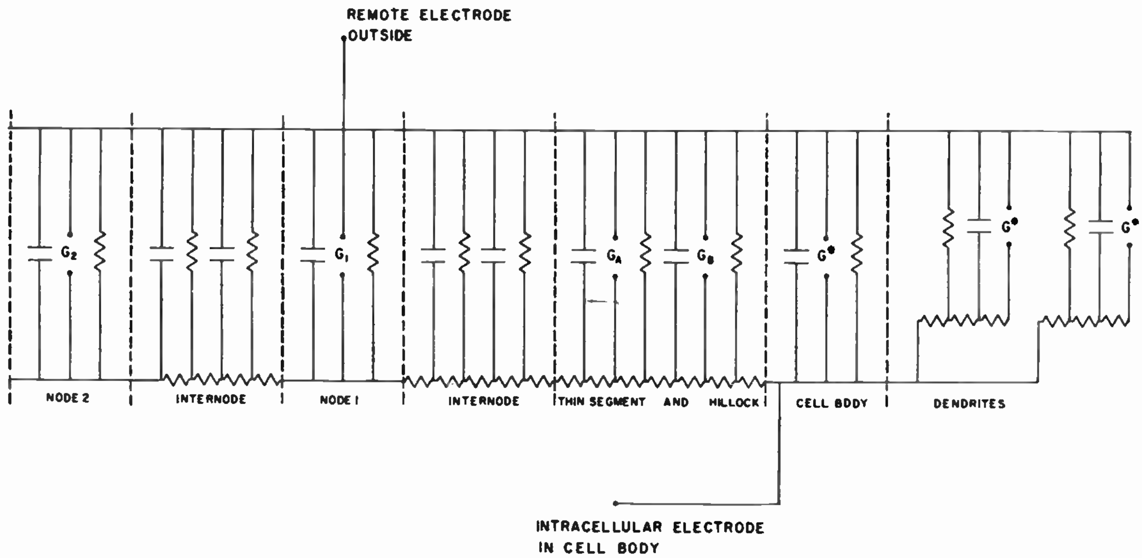


Fig. 3—Network equivalent of a nerve cell.

Nerve Impulse Generators[6,7] (G_1, G_2, G_A, G_B)

All of these generators can be represented by a circuit like Fig. 4. Resistances R_1 and R_2 vary with the potential difference across the membrane. When the membrane is not generating a nerve impulse, R_2 is much greater than R_1 and the potential difference across the membrane is almost 70 mv. In response to a sufficient change in potential difference across the membrane, about 10 mv in the direction inside less negative to outside, R_2 undergoes a large transient decrease in resistance and the potential difference across the membrane changes to inside 40 mv positive to outside. In less than a millisecond, R_2 returns to almost its initial high value. Resistance R_1 is also voltage sensitive and it decreases as the inside becomes less negative to the outside, but it lags in its change so that when R_2 has returned to its initial value, R_1 is still decreased. The outward flow of current, while R_1 is decreased, recharges the membrane capacity and the initial potential difference across the membrane is rapidly restored. When the inside of the membrane is no longer far from its initial value, R_1 increases and the cycle is complete. The voltage transient generated across the terminals of the generator is the nerve impulse, or action potential, and it has the form of Fig. 7(a). A few milliseconds must elapse before R_2 completely regains its ability to decrease again. This dead time is called the refractory period.

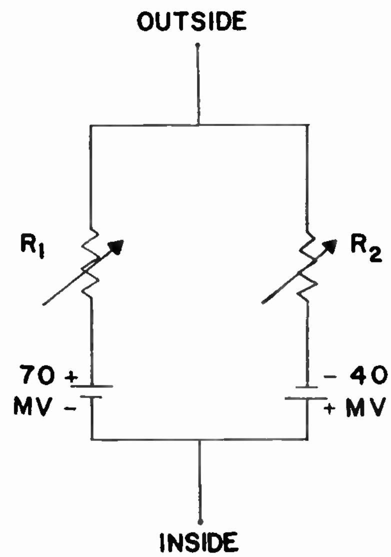


Fig. 4—Equivalent circuit of a nerve impulse generator.

In the foregoing description of the properties of the generators of nerve impulses, resistance R_2 decreased rapidly after the potential difference across the membrane changed to a critical voltage. The necessary initial change in voltage was in the direction of inside becoming less negative to outside. The subsequent decrease in R_2 changes the potential difference across the membrane further in the same direction which, in turn, produces a further decrease in R_2. In other words, the process is regenerative. If these effects occur fast enough, there is not sufficient time for R_1 to be significantly decreased. (A large decrease in R_1 initially would turn off the regenerative process by changing the potential difference in the direction of inside more negative to outside.) The critical voltage change necessary to initiate the regenerative process is the threshold.

Suppose that for some reason, generator G_A in Fig. 3 has the potential difference across it changed to its threshold level and the regenerative process is triggered. The peak effect is to make the inside of the cell about 110 mv more positive to the outside than it was before it was triggered. The outward current that will

[6] A. L. Hodgkin, "Ionic movements and electrical activity in giant nerve fibers," *Proc. Roy. Soc. B*, vol. 148, pp. 1–37; January, 1958.

[7] A. M. Shanes, "Electrochemical aspects of physiological and pharmacological action in excitable cells," *Pharmacol. Rev.*, vol. 10, pp. 59–273; March and June, 1958.

flow across the impedances in parallel with both G_B and G_1 will similarly trigger these generators by reducing the potential difference across them to the threshold voltage. Generator G_1 will trigger G_2, G_2 will trigger G_3, etc., and the nerve impulse will propagate along the axon to its end. This type of conduction is described as "saltatory," because it skips from node to node.

Nodal generators are either conducting impulses or they are at rest. Such behavior is the basis for the analogy between the nervous system and a digital computer since both systems are, at any instant, triggered or at rest, or binary. It will be seen later, however, that the neuron has other properties that limit the use of Boolean algebra in describing its function.

Chemically Activated Generators[8,9] (G^*)

These generators can be represented by a circuit like that in Fig. 5. Two variable resistances in this circuit, R_E and R_I, undergo decreases in resistance when they are acted upon by the chemical transmitter substances secreted by the end bulbs during synaptic transmission. Resistances R_E and R_I are decreased by specific transmitter substances; the substance that decreases R_E does not change R_I and vice versa. A decrease in R_E changes the potential difference across the membrane in the direction of inside less negative to outside, while a decrease in R_I changes this potential difference in the opposite direction. These resistances are not influenced by the potential difference across the membrane. The duration of the effect of transmitter substances on them is limited to the time necessary for these substances to diffuse away from the synaptic junction. An effective concentration lasts about one millisecond. If more transmitter substance is released by the end bulbs before the initial dose has diffused away, however, the effects will summate. Fig. 6(a) is a drawing of the transient in voltage that appears across the membrane where R_E is decreased by a transmitter substance, and Fig. 6(d) is a similar effect but of opposite polarity, induced by the appropriate transmitter substance on R_I.

Since resistances R_I and R_E are indifferent to the potential difference across the membrane, the transients produced by changes in them do not propagate like a nerve impulse. Instead, the chemically activated generators exert their influence on the generators of nerve impulses by the local flow of current across the impedances in parallel with them. If the summed result of the transmitter substances is to change the potential difference across the generators of nerve impulses to threshold, action potentials will be set off and will propagate along the axon away from the cell body. Therefore, synaptic action that makes the inside less negative to the outside is described as excitatory and the opposite action is called inhibitory.

[8] J. C. Eccles, "The Physiology of Nerve Cells," The Johns Hopkins University Press, Baltimore, Md.; 1957.

[9] H. Grundfest, "Electrical inexcitability of synapses and some consequences in the central nervous system," *Physiol. Rev.*, vol. 37, pp. 337–361; July, 1957.

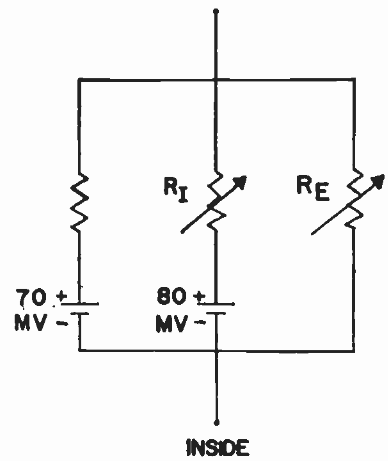


Fig. 5—Equivalent circuit of a chemically activated generator.

Some Evidence

As was mentioned above, it is possible to record the potential difference between the inside of the cell body and a remote electrode. The results of stimulating nerves that have single actions, *i.e.*, either excitatory or inhibitory, is drawn in Fig. 6. By stimulation of the appropriate nerve one can individually excite the chemically activated generators that change the inside either positive or negative to the remote electrode. In Figs. 6 and 7, a change in the direction inside less negative to outside is an upward deflection. This is an excitatory effect. Conversely, the action of the inhibitory chemically reactive generators produces a downward deflection in the figures; its signal makes the inside more negative to the remote electrode. The effect of stimulating a nerve that supplies the excitatory transmitter substance to the generator is drawn in Fig. 6(a). Applying increasingly stronger stimuli to such a nerve produces the family of curves in Fig. 6(b). As more axons in the nerve send propagated nerve impulses to the excitatory end bulbs at their ends, more excitatory transmitter substance is released. Therefore, the response of the chemically activated generators is enhanced and the signal increases. The result of applying two stimuli, a few milliseconds apart, is drawn in Fig. 6(c). It is additive. The converse of Fig. 6(a) is the drawing in Fig. 6(d), which describes the effect of applying a single stimulus to an inhibitory nerve. The effects of excitatory and inhibitory generators can summate, as is shown in Fig. 6(e). The excitatory stimulus has been applied a few milliseconds after the inhibitory one. Note that these effects decay with a 4 millisecond time constant. It is from experimental results like those shown in these drawings that the properties of the chemically activated generators are determined.

To study the properties of the generators of nerve impulses, another experimental technique is employed. With the very fine electrode still in the cell body, the far end of the axon is stimulated. Such a stimulus sends a propagated nerve impulse along the axon toward the

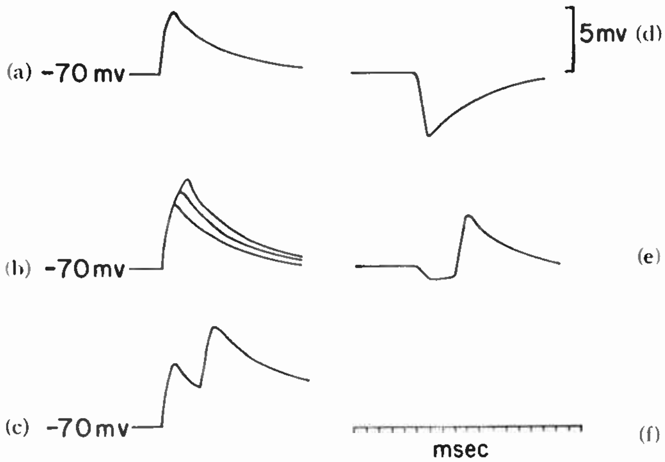


Fig. 6—Voltage transients appearing across the membrane of the cell body caused by chemically activated generators. (a) The effect of the excitatory generators. The family of curves in (b) shows the effect of increasing the excitatory stimulus. Two excitatory stimuli placed a few milliseconds apart produce records like the drawing in (c). (d) The effect of an inhibitory stimulus, and (e) the effect of applying an excitatory stimulus after an inhibitory one.

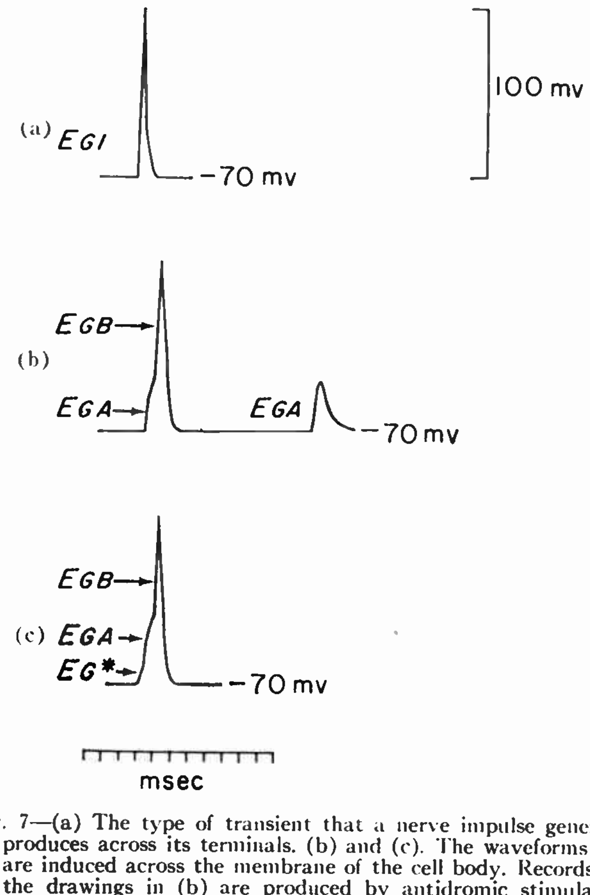


Fig. 7—(a) The type of transient that a nerve impulse generator produces across its terminals. (b) and (c). The waveforms that are induced across the membrane of the cell body. Records like the drawings in (b) are produced by antidromic stimulation, while those like (c) are produced by synaptic excitation.

cell body. This is opposite to the normal direction of propagation and is called antidromic conduction. Nevertheless, the generators of nerve impulses respond to the threshold voltage because the electrically responsive generators at a node do not distinguish between the effects of the generators at the nodes on either side. Indeed, their effects are identical when measured across the nodal membrane. The voltage transient that is generated across the electrically excitable membrane of a node has the form of the drawing in Fig. 7(a). When the recording electrode is in the cell body, some distance from node 1, the attenuation between node 1 and the cell body is so large that only a very small transient, if any, is visible when G_1 is triggered. G_A and G_B are closer to the cell body and the transients generated by G_A and G_B are not subject to the attenuation of an internode located between them and the cell body. The signal generated by G_A, E_{GA}, is somewhat attenuated by the thin segment. Its effect on generator G_B is sufficient to trigger it, and the signal generated by G_B, E_{GB}, appears across the membrane of the cell body in a relatively unattenuated form. The combined effects of the outward flow of current across the impedance of the cell body's membrane is the first voltage transient drawn in Fig. 7(b). Generator G_B must have a longer refractory period than G_A because when a second stimulus is applied to the axon a few milliseconds after the first, only generator G_A responds. Generator G_B is unable to respond until later, and only the effect of G_A, E_{GA}, is seen across the membrane of the cell body. The voltage transient that appears across the membrane of the cell body after sufficient synaptic excitation of the chemically-activated generators is drawn in Fig. 7(c). The portion of the voltage transient generated by the chemically responsive generators is labelled E^*. Generator G_B must have a higher threshold than G_A, because G_A is triggered before G_B. The transient generated by G_A that occurs across the membrane of G_B is sufficient to trigger G_B. The voltage developed across G_B by the chemically responsive generators alone was insufficient to trigger it. It has been assumed that the voltage transients generated by all generators of nerve impulses are the same when measured across the terminals of the generators. This has been done in order not to have to ascribe any more special properties to these generators. The justification of this assumption is mostly parsimony and the experimental evidence for other views is scanty.

In order to check two of the assumptions that were employed in the design of the equivalent circuit drawn in Fig. 3, a resistance-capacity-network analog of a neuron was constructed.[10] The two assumptions to be tested were: 1) that the space outside of the neuron was, for practical purposes, at the same potential as the remote electrode; and 2) that all points within the cell body were reasonably close to the same potential. It is known that the volume resistivity of the tissue outside the cell is about 222 ohm cm[11] and that of the protoplasm inside the cell is approximately 50 ohm cm. In

[10] A. J. McAlister, "Analog Study of a Single Neuron in a Volume Conductor," Naval Med. Res. Inst., Bethesda, Md., Naval Med. Inst. Rept., vol. 16, pp. 1011–1022; December 15, 1958.

[11] W. H. Freygang, Jr. and W. M. Landau, "Some relations between resistivity and electrical activity in the cerebral cortex of the cat," *J. Cellular Comp. Physiol.*, vol. 45, pp. 377–392; June, 1955.

constructing the analog it was assumed that the cell was axially symmetrical and had one large cylindrical dendrite. Such a model can be represented by square meshes of resistances that correspond to the radial and longitudinal resistances of a unit volume at any position inside or outside the cell. A unit volume in an axially symmetrical system will be a tube of unit length and thickness. Therefore, radial resistances in the square meshes will decrease logarithmically with radial distance from the axis and the longitudinal resistances will decrease in inverse ratio to the distance from the axis. The resistance and capacity of the membrane were inserted between the meshes at the location of the membrane in the axially symmetrical analog. Transients like the first one in Fig. 7(b) were imposed on a small area of the membrane of the cell body in the analog. It was found that both assumptions mentioned above did not introduce a serious error. In addition, it was noted that the time course of the very small potential differences that appeared between the outside surface of the membrane and the remote electrode was proportional to the time course of the membrane current.

This last observation has been employed experimentally to provide evidence that no nerve impulse generators are in the membrane of the cell body.[12,13] Two concentric pipettes, drawn in Fig. 8, were located so that the inner pipette of the assembly was inserted inside the cell body while the outer pipette remained close to the outside surface of the cell body. The voltage transients occurring at the tips of both pipettes were recorded simultaneously after an antidromically propagating nerve impulse was conducted along the axon. Since the voltage transient recorded by the larger pipette outside the cell is proportional to the flow of current through the membrane at the site of this electrode, the situation can be represented by the circuit drawn in Fig. 9. In this figure, $V_i(t)$ is the intracellularly-recorded voltage transient, $V_0(t)$ is the extracellularly-recorded voltage transient, and $I_m(t)$ is the time course of the current that flows through the membrane at the site of the extracellular recording. R_m and C_m are the resistance and capacity of this membrane, respectively. Their product, T_m, is equal to 4 msec. R_0 is the resistance between the tip of the extracellular electrode and the remote electrode. If α is defined as equal to $R_0/(R_m+R_0)$ the transfer function between V_0 and V_i in the p domain, where $p=d/dt$, is as follows:

$$\frac{V_0(p)}{V_i(p)} = \frac{p + \dfrac{1}{T_m}}{p + \dfrac{1}{\alpha T_m}}.$$

[12] W. H. Freygang, Jr., "An analysis of extracellular potentials from single neurons in the lateral geniculate nucleus of the cat," *J. Gen. Physiol.*, vol. 41, pp. 543–564; January, 1958.

[13] W. H. Freygang, Jr. and K. Frank, "Extracellular potentials from single spiral motoneurons," *J. Gen. Physiol.*, vol. 42, pp. 749–760; March, 1959.

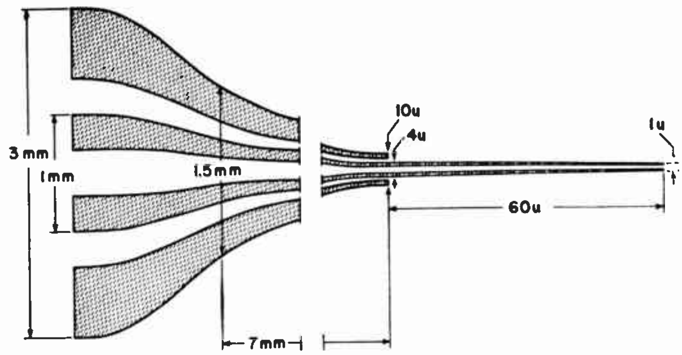


Fig. 8—Diagram of concentric pipette electrodes used for simultaneous recording from intracellular and extracellular locations

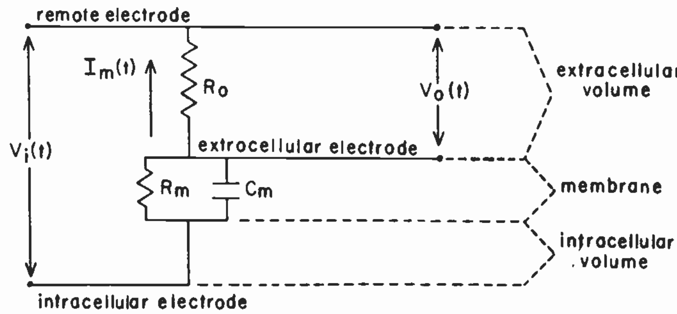


Fig. 9—Equivalent circuit of the recording situation for intracellular and extracellular location of electrodes. $V_i(t)$ is recorded from the intracellular location, $V_0(t)$ from the extracellular location. Membrane resistance and capacity are R_m and C_m. R_0 is the resistance from the tip of the extracellular electrode to the remote electrode.

Since this equation depends only on T_m and α, and since T_m is known, components may be chosen to build a circuit that will produce the same sized $V_0(t)$ as is actually recorded. The value of α found this way is 0.06. When the signal recorded from the electrode inside the cell, $V_i(t)$ is fed into this circuit, its output, $V_0(t)$ may be compared with the transient recorded by the electrode outside of the cell. The time courses of these transients are practically identical. This result is possible only if the membrane resistance of the cell body does not change during the time of the transients. Since it is known that the capacity of biological membranes is independent of voltage, the resistance of the membrane of the cell body must also be unchanged. This can be true only if there are no nerve impulse generators in this membrane.

The reader is cautioned that these interpretations of experimental results are not universally agreed upon by neurophysiologists. For example, it may be that, in the last experiment described, the electrode assembly has damaged the electrically responsive generators. If the reader reviews the experimental results that have been described and considers the assumptions upon which the interpretation of them is based, it will be evident that these arguments are not very rigorous and that the reasoning seems to approach being circular at times. These difficulites are partially specious because, in an article of this sort, only part of the data can be described. In some cases, however, the difficulties in interpretation are not completely overcome.

Some Consequences of the Equivalent Network

If the neuron does send a propagating nerve impulse along the axon to some other nerve cell or to some effector organ, the transient in voltage that is found across the membrane of the cell body has the form drawn in Fig. 7(c). In other words, the combined effects of the excitatory and inhibitory chemically activated generators has made the inside of the membrane containing generator A sufficiently positive to trigger this generator. The nodal generators are triggered subsequently. It may be noted that it is not necessary for generator B to be triggered; a nerve impulse is started at node 1 whether or not generator B is triggered.

The dendrites act like delay lines in their conduction of the excitatory and inhibitory voltage transients. As the distance from the cell body along a dendrite increases, the dendrite tapers and its core resistance per unit length becomes larger. Therefore its attenuating effect is enhanced, especially for rapid transients. One can appreciate the very complex summation of the voltage transients that are generated by the chemically activated excitatory and inhibitory generators in the dendrites. In general, those generators that are closer to the membrane of the cell body, or are in it, will be more influential in determining whether or not the nodal membranes are triggered.

It is possible, or maybe probable, that the dendrites function in other ways than have been indicated. Some dendrites are so thin and long that it is difficult to imagine how the generators at their tips could have any effect on the generators of nerve impulses. Possibly they have some sort of biological amplifier along their shafts. Very little is known about the dendritic function. Equally unclear is the mechanism by which nerve impulses in the axon lead to secretion of transmitter substances from end bulbs.

As has been mentioned, the number of neurons in the nervous system is about 10^{10} and the interaction between neurons is extensive. If the neuron is taken as the fundamental component in this system and the somewhat overworked analogy to an electronic computer is made, one can appreciate the capabilities of such a device, as well as how appropriate the analogy is.

Electronic Control of Some Active Bioelectric Membranes*

JOHN W. MOORE†, SENIOR MEMBER, IRE

Summary—**Special purpose real-time analog computers are used to measure and control nerve membrane potential or current in a squid axon or a single frog node. Under current control, the membrane potential has a region of discontinuity and an "action potential" rather similar to that observed in normal impulse propagation. With potential control, the current pattern is a continuous function of the potential, and a negative resistance is found in the region of potential discontinuity for the current-controlled membrane. The membrane's electrical characteristics may therefore be compared with some two-terminal transistor switching circuits.**

Introduction

Electronic control of the potential across the wall, or membrane, of electrically active cells is a powerful tool which has been developed to make possible a new realm of physiological experimentation. Such control provides more or less easily interpretable data on the properties of ionic currents across cell membranes. These data can then be used to test the validity of hypotheses concerning ionic movements through membranes as the basis of normal electrical activity.

In order to develop such a control system, it was necessary to consider and appreciate some of the normal membrane characteristics; these will be described first.

Bioelectric Membrane Characteristics

Passive Characteristics

In another paper in this issue, Schwan [1] has described in detail the passive electrical characteristics of living tissue. The passive electrical characteristics of active cell membranes, such as found in muscle, nerve, and in electric organs of fish and eels, will be briefly noted here. In addition to the rather commonly found one microfarad per square centimeter and a resistance from 1–10 kilo-ohms cm² (depending on the condition of the membrane), a steady or resting potential of between 70 and 100 mv, the inside negative with respect to the outside, is found to exist across the usual membrane

* Original manuscript received by the IRE, September 1, 1959.
† Laboratory of Biophysics, Natl. Institute of Neurological Diseases and Blindness, Natl. Institute of Health, Bethesda, Md.

[2]–[4]. Although this potential is a rather small working voltage for capacitors with which engineers are familiar, the dielectric is quite thin (the best estimate from electron micrographs [5] is about 75 angstroms) so that the field is in the order of 100 kv/cm. Experimentally, it has been found that the potential can be increased by at least 50 per cent before an apparent spark breakdown takes place [6]. This certainly puts the membrane in the dielectric strength region of oils and paraffin. It is also interesting to note that the membrane may recover completely after an incipient breakdown of a few milliseconds, or be in relatively good condition after a short breakdown pulse.

Hodgkin and Huxley [7] have studied squid axon by means of some of the techniques to be described here and have arrived at the electrical circuit model for the membrane shown in Fig. 1. The steady resting potential

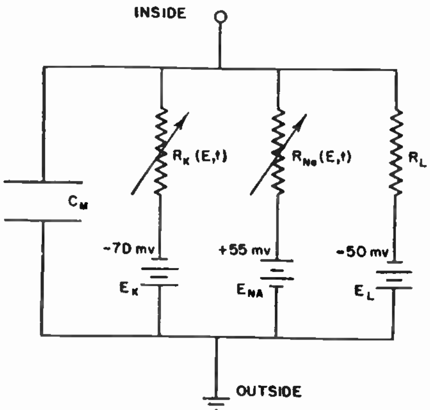


Fig. 1—Electrical equivalent circuit of squid axon membrane (slightly modified version of Hodgkin and Huxley [7]). E_K, E_{Na}, E_L refer to equilibrium potentials and their corresponding source resistances for potassium, sodium and a leakage ion (possibly chloride). R_K and R_{Na} are both functions of time and membrane potential.

across the membrane seems to depend mostly on the large difference in concentration of potassium ion [2], [3], where the concentration inside of the cell is some 40 or more times than that outside. The Nernst equation gives a potassium equilibrium potential of $E_K = RT/F \ln [K]_0/[K]$ [8]; while this concentration ratio makes the inside negative in the resting condition, electroneutrality and osmotic balance require that another cation have a low concentration inside. It has been found that while sodium in the extracellular space has a high concentration compared to all other cations, it is in quite low concentration inside these cells, usually in the order of 1/10 of that outside. Thus, there is also a sodium battery in the membrane with a positive inside potential.

Active Characteristics

When one of these tissues is excited by moving its potential from rest to somewhere in the neighborhood of −40 mv on the inside, the over-all membrane resistance rapidly decreases by a factor of about 100. This is apparently caused by a change in the resistance in series with the sodium battery from a value much higher than the potassium battery resistance to a value much lower than the potassium source resistance. Because of this, the potential across the membrane reverses to about 50 mv positive. The low resistance of the sodium battery is not maintained and, in nerves, the potassium resistance decreases and the potential reverts to its resting value in a millisecond or less, depending on the temperature. However, this potential may be maintained for several milliseconds in heart muscle and electric organs of marine animals [4]. It will be shown later, in the experimental section, that this resistance change can take place in a small fraction of a millisecond after a step change in potential and that the conductance can vary from the resting value of a fraction of a millimho to 100 millimhos per cm². Currents of up to 10 or 20 milliamps/cm² may be observed under such conditions. The capacity seems to remain constant at one microfarad/cm² [10] throughout the course of the resistance changes and apparently represents the maintained integrity of the majority of the membrane structure. In nerve and muscle, where the geometry is cylindrical, propagation along its length takes place as successive cross sections undergo the potential changes noted above. In contrast, in marine electric organs a more or less matchbox geometry is often found [4]. Here the active membrane is on one of the large sides and a low-resistance membrane with many undulations giving a large cross-sectional area and a lower current density is on the opposite side. A relatively large fraction of the 100–120-mv change across the active membrane appears between the external solutions on the two sides of the cell. Nature has given certain electric fish cells of this sort, stacked in series for high potentials needed in fresh water or arranged largely in parallel to give the very high current densities required for shocking prey in sea water.

Energy Considerations

With each impulse or action potential, energy is dissipated by both sodium and potassium. It is found that a few micromicromoles (per cm² of membrane) of sodium go from the high outside concentration to the low inside concentration, and a similar number of micromicromoles of potassium flow out of the cell [12]. In the giant axon of the squid, to be described later, there is a large volume-to-surface ratio and the percentage change in the concentration with each impulse is extremely small. In fact, thousands of action potentials can be fired with rather small changes in the concentration ratios. It appears that under normal conditions in life there is a continuous slow extrusion of sodium and a corresponding accumulation of potassium [13] in order to maintain, so to speak, a large balance of money in the bank to cover the very high withdrawal rates for short periods. This exchange of ions requires an energy that must in some way be provided by metabolism.

Special Case #1, Unmyelinated Axon

While all of the bioelectric tissues seem to have several features in common, there are some apparent differences in detail. Because nerve is the only one of these tissues so far studied in which the parameters were controlled at the option of the investigator, the discussion will now be limited to these. The problem with which we will concern ourselves here is the mechanism by which signals are transmitted over considerable lengths of nerve on their way to activating other nerves, muscles, or glands, etc. In another paper in this issue, Freygang [14] gives a description of what goes on at junctions between nerves.

The unmyelinated axon, one of the simplest forms of nerve in both evolutionary and structural senses, is a long thin cylinder stretching over many centimeters between its end junctions at sense cells, ganglia, or muscles, etc. This cylinder is rather like a sausage filled with a gel containing, as already noted, a high concentration of potassium, a low concentration of sodium [15], and rather large organic anions [16] to which the membrane is impermeable. The conductivity of the interior [17] is almost as good as that of the exterior plasma solution and all of the electrical activity takes place across the very thin membrane of the sausage. Nature was, for once, rather kind to the electrophysiologist in providing one type of nerve cell which was large enough for the insertion of wires. Apparently, during the course of evolution, a large number of small fibers fused to make each of these large fibers which occur in the squid commonly found along our Atlantic Coast. It is particularly useful for the squid to have the large fast conducting fibers to transmit signals to the muscles far distant from the head. The synchronous and powerful contraction of the whole mantle thus ejects sea water at the high velocity needed for effective jet propulsion. Two of these giant axons, with diameters up to 0.6-mm lengths of several centimeters, may be dissected out of one animal. Recent reports indicate that larger squid of a different species found off the coasts of Chile and Peru provide much longer axons, which are more than a millimeter in diameter.

The whole length of the cylindrical nerve acts as if it were a distributed pulse generator and shaper so that, once triggered at any point, an action potential moves the length of the nerve with a shape which is independent of both the position and initiation. The propagation of the impulse is described by the one-dimensional wave equation

$$\frac{\partial^2 V}{\partial x^2} = \frac{1}{\theta 2} \frac{\partial^2 V}{\partial t^2}$$

where θ is the conduction velocity.

Special Case #2, Myelinated Axon

The myelinated axon is a much more highly developed and efficient impulse transmitter than the simple unmyelinated axon. It appeared later in evolution and is widely used in vertebrates. It is very much smaller than the giant axon—usually a few micra in diameter—and has heavy myelin insulation except at small patches (about 1 micron long) called nodes. Only these exposed areas (about 2 mm apart) show electrical activity, and they appear to have essentially the same electrical properties as the unmyelinated fibers. An impulse initiated at one node passes enough depolarizing current through an adjacent one to fire it. Thus, a repeater-station type of transmission provides a fast propagation with a minimum of energy loss.

A large number of these small fibers are usually packaged in a bundle; the bullfrog sciatic is a convenient source for a length of such a nerve trunk. Considerable skill is required to dissect out single fibers and to expose single nodes as described later in the paper.

Squid Axon Membrane

Equation Simplification by Complex Experimental Arrangement

In Fig. 2 (top), the membrane current is separated into a capacitive charging current, $C(dV/dt)$, and an ionic current, i_i where the characteristics of the ionic element, indicated in the circle as in parallel with the capacity, are to be found. The equation for the system can be simplified by removing the longitudinal variation of V with internal and external short circuits as schematically indicated in Fig. 2 (center). This short circuiting can be accomplished by the insertion of an axial wire several millimeters long; the axon can survive for a few hours with a wire of diameter up to about one-fifth that of the axon. The presence of a low resistance of the wire does not alone insure a satisfactory short circuit unless the contact between the wire and the axoplasm (or gel) can be relatively low. The whole problem of surface impedance between metal and solutions or gels is rather complex. In the widely used silver-silver chloride (Ag-AgCl) electrode, chloride can move reversibly onto or off the surface. An even lower surface resistance may be obtained by electroplating platinum from a platinic chloride solution onto either platinum or another metal. This process causes a rather soft soot-like deposit on the surface which can be given mechanical protection by coating with gelatin or collodion. The greatly increased effective area is an important factor in lowering the surface resistance and increasing its shunt capacity. Such a surface-treated wire can effectively short circuit a length of membrane and require it to move together in potential [18].

Under the condition of *current* control of this area of longitudinally short-circuited membrane, there is still a region of instability giving rise to a rather characteristic action potential [19]. However, with the *potential* across the membrane controlled, the current is everywhere single valued and the region of negative resistance which gives rise to the instability may be studied.

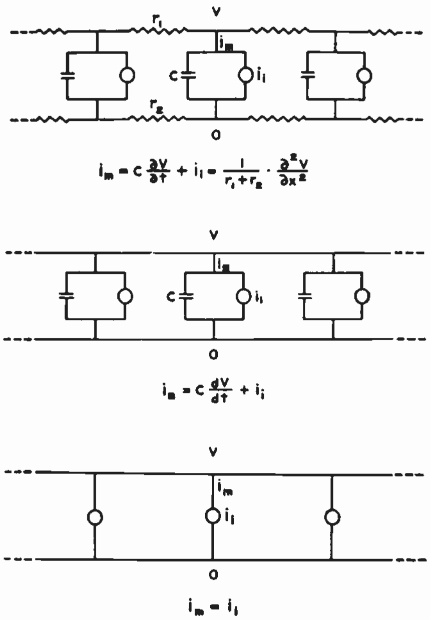


Fig. 2—Equivalent circuit of axon under various experimental conditions. Top: distributed circuit with membrane regions coupled by internal axoplasm and external solution resistances (r_1 and r_2, respectively). The three (ionic currents) of Fig. 1 are lumped into i_i. Center: axon longitudinally short-circuited by an axial wire (external resistance also made small by a closely placed electrode). Bottom: voltage-controlled axon after a step potential change. (Top of Fig. 2 reproduced from K. S. Cole, "Beyond Membrane Potentials," in "The Electrophysiology of the Heart," *Ann. N. Y. Acad. Science*, vol. 65, art. 6. Courtesy of the author and *Ann. N. Y. Acad. Science*.)

Electronic control of the membrane potential at the desired level, obtained by supplying the proper current to the longitudinal axial wire, is equivalent to a radial short circuit. Under such conditions, the capacitive current is zero except during the transient when the potential is pulsed from one level to another. When the voltage transient and corresponding capacitive charging current are limited to 50 μsec or less, the relatively slower ionic current may be observed subsequently. This temporal separation of the ionic and capacitive currents allows study of the ionic characteristics above. This is indicated in the bottom part of Fig. 2.

The concept of controlling the potential across the membrane and measuring the associated currents required during a short step of potential was first introduced for squid axon by Cole [20] in 1949. When the membrane is maintained at a potential near the resting value, its resistance is a few thousand ohms for a square centimeter and the potential across it is essentially equal to that between the internal current wire and the exterior. However, when the membrane is driven in a positive direction by 30 mv or more from its resting value, the resistance decreases so markedly that it approaches that of the network connecting it to the wires as shown in Fig. 3. Cole [20] originally used the potential between the inside and outside current electrodes as a measure of the membrane potential, and this introduced considerable error under large current conditions. Hodgkin, Huxley, and Katz [21], recognizing this difficulty, carried out a set of elegant experiments (to which they gave the name "voltage clamp") in which they added a second pair of electrodes to measure the potential. These electrodes gave a more accurate representation of the membrane potential because they did not include the potential drop across the electrode surface impedances. Still later, Moore and Cole (manuscript in preparation) obtained a more accurate measure of the potential across the desired element by using an electrolyte-filled micropipette [11] penetrating to just inside the membrane and a larger reference electrode just exterior to the membrane (see Fig. 4). Thus, the voltage drops across the external sea water and the internal axoplasm resistances, as well as the potential drops across the current-carrying electrodes, were eliminated.

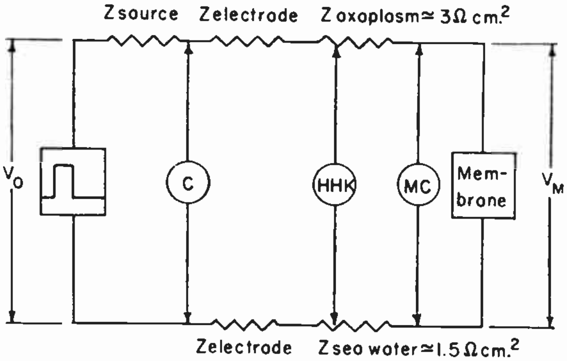


Fig. 3—Equivalent circuit showing resistances between the membrane and controlling potential source. Control of the potential between points at arrow heads has been applied by Cole [20], Hodgkin, Huxley, and Katz [21], and Moore and Cole.

The radial membrane current is divided between three separate external current electrodes (relatively large wires also electroplated for low impedance). Guard chambers on each side of the central measuring chamber insure that the measured current is normal to the surface and uncontaminated with end currents.

Experimental Preparation

Several centimeters of the nerve trunk controlling the mantle contraction of the squid are dissected out of the animal. The single giant fiber of about 0.5 mm diameter may be separated, under a microscope, from a bundle of very small adhering parallel nerve fibers.

After dissection, the cleaned axon is put into a trough with flowing sea water (or test solution) which is rather carefully temperature-regulated because the amplitude and rate of change of the ionic conductances are very temperature dependent. A fine wire usually 75 microns in diameter which has been carefully electroplated, generally with platinum, to give a low surface impedance is inserted through a cut near an end of the axon and is pushed along the axis to the position shown in Fig. 4.

Potential Measurement

A fine glass capillary tube pulled down to a tip of about 0.5 micron is inserted through the axon mem-

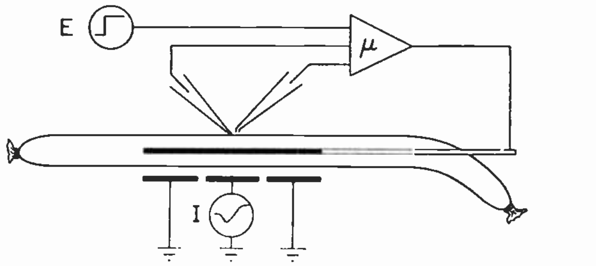


Fig. 4—Simplified schematic representation of experimental arrangement for voltage control of squid axon membrane. The potential between the capillary electrodes on the two sides of the membrane is forced to match the signal E by connecting the output of a high-gain amplifier to the axial wire. Lateral guard electrodes minimize end effects in the measured current flowing in the central electrode.

brane, and can remain, without apparent injury to the cell, over a period of several hours. This capillary is filled with 3M KCl solution to provide an electrolytic path between the interior of the axon and a calomel half cell for measurement of the inside potential. Such microelectrodes may have a resistance of 1–50 megohms, depending on the size of the tip. In addition, there is a capacity across the glass wall of approximately 1 picofarad per mm of length immersed in the sea water exterior to the membrane. The potential just outside the membrane is obtained from another capillary with a considerably larger tip and a lower resistance KCl-Agar bridge to another half cell.

The high-resistance electrolyte-filled microelectrodes are widely used in electrobiology, and a number of high-input impedance electrometer-type preamplifiers, with anticapacity compensation to match such an input, have been designed in the past few years for intracellular potential measurements with microelectrodes. A detailed discussion of electrodes and matching amplifiers is given in another paper in this issue [11]. A preamplifier of this type which has been chopper-stabilized against drift (descriptive manuscript in preparation), while maintaining an extremely high input impedance, measures the internal potential, V, with a gain of five, as shown in Fig. 5. The potential on the external reference electrode, V_R, is inverted in sign and appears at the output of the reference operational amplifier #2. Because this electrode has some resistance, a padder resistance is included in the feedback path to adjust the gain to unity. The output of the electrometer preamplifier, $+5V$, and the reference amplifier, $-V_R$, are added with the proper weighting factors by operational amplifier #3 to give $-V_M$, the negative of the voltage difference across the membrane; this is monitored on the CRO and X-Y plotter.

Control and Readout Circuits

A simple, special purpose analog computer measures and controls the desired parameters. For potential control of the membrane, the input switch for operational amplifier #4 is connected to $-V_M$. This amplifier simply inverts the polarity of its input to the proper phase for

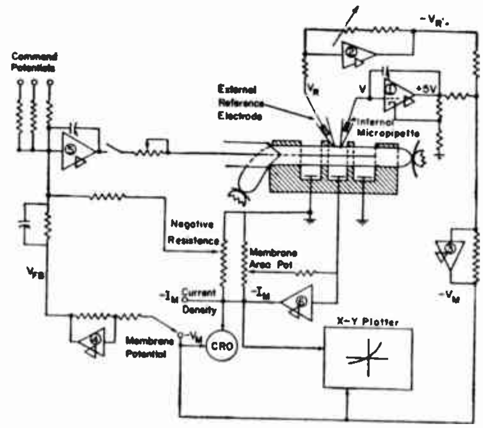


Fig. 5—Measurement and control circuit schematic diagram. The large triangles represent operational amplifiers; the small attached triangles indicate an associated chopper-stabilization amplifier. Refer to text for a description of the operation of the circuit.

addition to the command signals. If V_{FB} is not equal and opposite to the sum of the command signals, the output of the control amplifier #5 changes in potential with the proper direction and magnitude to bring about equality. A switch and a variable resistance have been inserted between the output of amplifier #5 and the axial wire to avoid damage to the axon during setup and to allow a gradual approach to complete control. The command signals normally consist of a dc level plus pulse potentials of various magnitudes, durations, and sequences. Slowly varying ramp potentials are also used to scan the membrane potential for plotting approximate steady-state current-voltage characteristics on the X-Y plotter.

The current-carrying electrodes in the guard chambers are connected to ground. The central measuring chamber electrode is connected directly to the summing point of operational amplifier #6 which holds its summing point very close to ground potential by feedback. The area of the membrane in the measuring section was usually less than the 0.1 cm² for an axon 550 microns in diameter. Therefore, the feedback resistor (for amplifier #6) was chosen to give a convenient scale factor in current density for 0.1 cm² for the membrane area potentiometer at maximum setting. Smaller measuring areas from smaller axons were conveniently entered on the dial of a 10-turn potentiometer so that the output of amplifier #6 was always proportional to the membrane current density. Compensation for any resistance between the potential electrodes and in series with the active membrane can be accomplished by picking off an appropriate signal from the potentiometer labeled "negative resistance."

Provision is also made to control membrane current instead of potential, by switching the input to amplifier #4 to the upper position. A current clamp is useful for studying action potentials and excitation thresholds of the membrane and its passive electrical characteristics, such as capacity and series resistance.

In common with all feedback systems, phase and amplitude relations must be controlled in the high-frequency cutoff range. The rise time of the microtip, with a usual resistance of about a megohm, and the electrometer-preamplifier, in response to a step input, is about 10 μsec and is a limiting factor in the over-all system response. Chopper-stabilized operational amplifiers with the fastest responses compatible with low noise, hum, and ripple are used throughout the remainder of the system; these have rise times of about a microsecond or less. The large capacity of the membrane to be controlled, about 0.3 μfd for the three chamber sections, is in itself an extreme load to put on the output of any amplifier. Its impedance decreases with increasing frequency and, because of the electrode and axoplasm resistance in series, the voltage across the membrane capacity lags the output of amplifier #5 increasingly with higher frequencies. Stability of the over-all feedback circuit against oscillations was obtained with the addition of some phase advance to V_{FB} and a small feedback condenser across amplifier #5.

Component and System Requirements

The steep membrane characteristic in the negative resistance region (see Fig. 9) sets the noise and drift limits at the equivalent of a fraction of a millivolt of membrane potential if widely scattering data are to be avoided. Still other specific requirements must be met by amplifiers #5 and #6. In particular, amplifier #6 must be very quiet (ripple and noise currents should be small compared to 10 μa) to give a good signal-to-noise ratio on low membrane currents and yet be fast enough to carry the condenser-charging currents. If it is not fast enough to faithfully reproduce the condenser-charging current (the rise time should be less than a microsecond with a 10-kilo-ohm feedback resistor), not only will it introduce an error into the current measurement, but also an undesired change in potential at the summing point. The gain amplifier, #5, must have a fractional ohm output impedance to properly drive the membrane capacity load already mentioned.

In addition, it must have a rather good output current capacity because it may have to supply at least 0.3 cm^2 of membrane plus areas beyond the end of each guard with ionic currents of up to 20 milliamperes/cm^2. If the over-all system can be made faster, the condenser-charging current, equal to CdV/dt, would, for instance, have to be as much as 50 ma in order to change the membrane potential by 150 millivolts in one microsecond.

Many commercially available operational amplifiers are designed for relatively slow computation with full-scale voltage range of up to ± 100 volts. Consequently, the need for reduction of ripple and noise below a few millivolts and bandwidths in excess of 100 kc is not felt. Although no one amplifier satisfies all the desired characteristics, we have usually been able to make fair compromises between wide bandwidth and low noise requirements by modifying commercial amplifiers and using low ripple (5 mv or less p.p.) power supplies. It is often possible to remove internal cutoff networks, designed to keep the amplifier stable under all gain and load conditions, and obtain much faster response by using minimum high-frequency cutoff external networks consistent with stability for the particular load and gain position in our circuit. All of the amplifiers used were chopper-stabilized and designed for low grid current. In addition, some had a positive feedback control provided to adjust the open loop gain to approximately infinity. However, there were times when it was desirable to use another control amplifier with less noise but with an open loop gain of less than 10^5. Because of this, particular attention was paid to the design of the summing network for amplifier #5; all of the summing resistors were made as high as possible compared to the feedback resistor. The error signal at the summing point of the control amplifier was therefore as large a fraction of the feedback function (V_m or I_m) as practical to give a maximum effective over-all loop gain of the system.

A reduction of the present over-all system response time is desirable particularly for experiments at the higher temperatures. As the temperature is raised, the ionic currents change rapidly, being about ten times as fast at 25°C as at 5°C, and start to change before the capacitive charging transient becomes negligible. If it is possible to make a faster microtip-preamplifier package by elimination of the few inches driven shield input cable (its capacity also slows the package response), still faster operational amplifiers would be required.

Control System Performance

Fig. 6 shows typical choices of response of the voltage across the membrane, V_m, to a step command and the corresponding condenser charging currents which are completed in 50 μsec or less.

The output impedance of the control circuit was measured by recording the voltage introduced across the membrane when an external pulse of current was injected into the axial wire. With a summing network already mentioned and a control amplifier of open loop gain limited to 3×10^4, a steady-state impedance of 0.25 ohm was obtained. This figure was still lower with an amplifier equipped with internal positive feedback.

The schematic representation of potential patterns at various points in the control circuit, as in Fig. 7, is a helpful aid in understanding the operation of the voltage clamp and the magnitude of the signals involved. A potential pulse requiring a large inward sodium current is chosen to demonstrate how large the potential variations can become. Starting at the bottom, with a peak sodium current of about 5 ma/cm^2, the potential of the reference electrode just outside the membrane may vary as much as 20 mv. In order for the difference of potential across the membrane, V_M, to be a step function, the potential at the microtip must be just the sum of the step and the reference potentials.

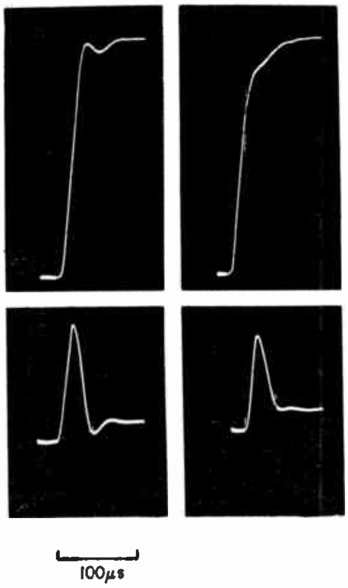


Fig. 6—Typical choices (left and right) of response of the over-all control system to a step command. Top: controlled-membrane potential, V_m; bottom: corresponding capacitive current, essentially completed in 50 μsec.

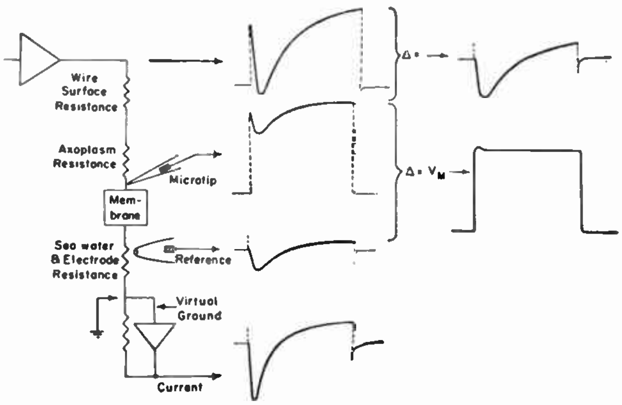


Fig. 7—Schematic representation of potentials at points around the control loop. All potentials are to same scale, $V_m = 50$ mv; current (lowest record) has peak of 5 ma/cm².

Since the potential drop across the axoplasm and needle impedance is approximately proportional to the current, the output of the control amplifier (connected to the axial wire) must move in the same sense as the inside of the membrane, but with wider swings. The output of the control amplifier is, of course, a direct measure of the error in the matching of membrane potential with the command signal. As the needle surface impedance increases, the output must move with larger amplitudes and the error in the membrane potential becomes larger.

Membrane Characteristics

A typical family of currents for various membrane potentials during the pulse is shown in Fig. 8 for a temperature of 10°C. The amplitude and time constants involved in both the peak- and steady-state currents are potential-dependent. The early peak in the current occurs with the turning on and then off of the sodium conductivity. The direction of the peak reverses as the potential goes through the thermodynamic equilibrium potential of the sodium battery. The potassium conductivity increases more slowly than the sodium and then remains constant almost indefinitely. The early peak- and steady-state currents are plotted as functions of the voltage during the pulse in Fig. 9. The positive slopes of the current-voltage lines (or conductances) appear to be approximately constant and equal to about 100 millimhos/cm² over a large range of potentials. There is also a region of steep negative slope, expressible as a low value of negative resistance where the sodium is being turned on. Although this is not a steady state, it is rather constant for a millisecond or so. At a temperature of 20°C, the currents and conductances are 50 to 100 per cent larger.

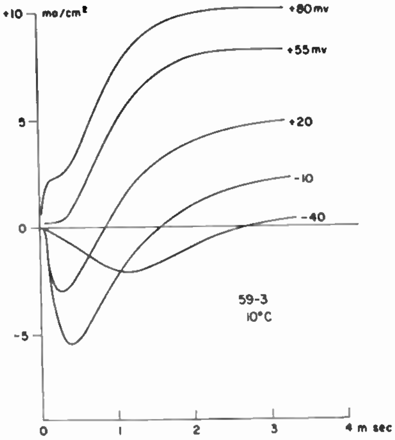


Fig. 8—Typical family of membrane current densities as a function of time after the membrane potential has been pulsed to the value given at the right end of each curve. The membrane potential was held at 70 mv between pulses. The fast transient capacitive currents of Fig. 6 have been omitted.

With a current clamp, measurements of the membrane capacity and the small series resistance may be obtained. For a step of current, the membrane potential will jump by an amount proportional to the series resistance and then the capacity will start to charge at a uniform rate, as shown in Fig. 10. While the value of the capacity (about 1 microfarad/cm²) is relatively easy to obtain from the slope, the speed of the circuit in applying the current step and the speed of the membrane potential recording system limit the accuracy of the series resistance determination and demonstrate the need for faster systems before an accurate value for the series resistance can be obtained.

Frog Node

Potential Measurement

The internal longitudinal resistance between nodes is about 50 megohms and the nodal membrane resistance varies from about 50 megohms at rest to 5 megohms or less in the active state [22]. Several approaches to the problem of obtaining reliable measurements of the

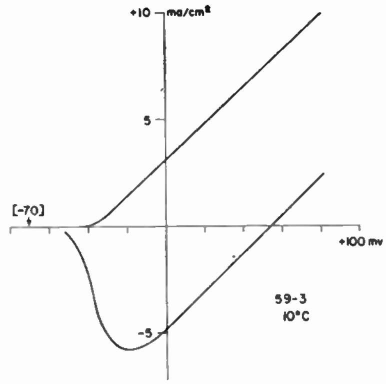


Fig. 9—Peak initial current density (sodium), lower curve, and the steady-state current density (potassium), upper curve, are plotted as a function of the potential during the pulse.

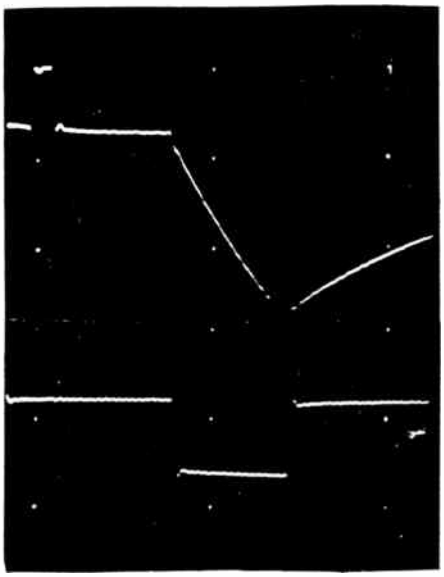


Fig. 10—Upper: record of membrane potential response to a step of current density, preceded by response to test pulse applied to input of microelectrode amplifier. Lower: record of current pulse of 250-μsec duration.

potential across this active membrane have been made. While the technique of insertion of submicron tip glass pipettes filled with potassium chloride solution [11] has served in a magnificent fashion for obtaining potentials from many types of cells in the past few years, it has been far from successful in measuring nodal membrane potentials. Apparent injury and a rapid decline of action potentials has been seen by Woodbury [23], even with very fine tips of up to 50 megohms resistance.

Other techniques, which essentially use the conductivity of the interior of the internodal region as a connection to the inside of the membrane, have been used with varying success by Tasaki and Frank [24] and Stämpfli [25]. In this case, the effective source resistance is the 5 to 50 megohms intrinsic in the nodal membrane plus the interior internodal resistance of another 50 megohms. Therefore, the problem is to make the measuring circuit and external leakage impedances very high compared to 100 megohms. Electrometer-type preamplifiers with fair bandwidth and feedback to cancel input capacitance, as developed for measurements with the high-resistance microelectrodes [11], are quite satisfactory for this application. However, an insulating gap of air is not perfect because of the small amount of physiological solution (containing about 0.1 N sodium chloride) adhering to the outside of, and probably permeating, the myelin between nodes. This exterior resistance of only 20 to 100 megohms shunts the source by a considerable amount. The higher values of this resistance are usually obtained by careful fanning with dry air and are not very stable. Flowing sucrose gaps [26] have given external resistances of 10 to 100 times that of the inside, but require a great deal of care in the preparation of the sugar solutions to assure very low conductivity.

Frankenhaeuser [27] developed an ingenious and elegant application of negative feedback to increase electronically the external leakage resistance path between two nodes by a large factor. It is similar in concept to an older experiment by Huxley and Stämpfli [28] in which they used multiple insulating seals and manually adjusted the current across one seal so that the longitudinal external current measured as a voltage across another seal was nulled. They balanced and thereby measured the resting and peak action potentials. Frankenhaeuser introduced automatic balancing with large negative feedback in a rather similar experimental arrangement to obtain both the resting and fast action potentials. Moore and del Castillo [29] have used two negative feedback configurations with results similar to those of Frankenhaeuser.

The concept of this method is illustrated by Fig. 11, in which negative feedback is used to increase the input resistance so that a signal potential with a high source resistance may be measured with fidelity. If Z and R represent the resistances across two insulating vaseline seals along the exterior of an internode region and R_S is the internal internode and membrane resistance, it can been seen from Fig. 12 that accurate measurements of the interior potential at a node N_0 may be made. Both nodes N_{-1} and N_{+1} are in isosmotic potassium chloride which reduces the resting potential to zero and reduces the nodal membrane resistance to about 1 megohm, and prevents action potentials. It is relatively easy to make vaseline insulating seals of 10 or more megohms resistance. Without feedback, therefore, about 20 per cent of the full action potential appears across the seal at the right. When the negative feedback loop is closed, the solution between the seals is driven so that a minimum voltage appears across the right seal and essentially all of the drop occurs across the center seal in parallel with the amplifier output. A stimulating current, with an extremely high source impedance, may be applied across the left seal and a relatively faithful reproduction of the resting and action potentials will be obtained at the output of the preamplifier. The output is independent of the actual value of the seal resistance, or changes of it, if the open-loop gain is made

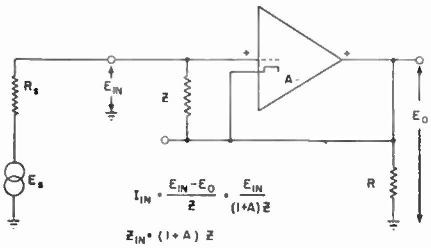


Fig. 11—Schematic diagram of conventional negative feedback method for increasing the effective input impedance of an amplifier.

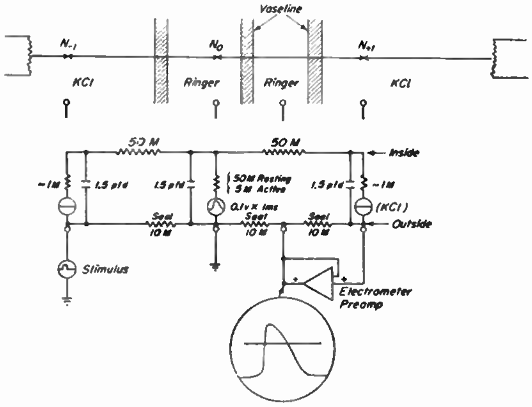


Fig. 12—Application of circuit in Fig. 11 to increase external internodal resistance by feedback. Nerve fiber and seals are shown above the equivalent circuit and connections.

large enough. It is possible to achieve a rather large open-loop gain for the preamplifier by means of positive feedback inside the loop with some sacrifice of bandwidth. The stimulus arrangement is not very practical because it requires a source impedance very high compared to 50 megohms.

A more practical feedback configuration, similar to that used by Frankenhaeuser, is shown in Fig. 13. In this case, the solution between the center and right seals is at ground and the potential across the right seal appears at the input to the preamplifier. The output of the preamplifier is further increased and inverted in sign by the operational amplifier. Its negative output is applied to the solution outside node N_0 with the proper phase and amplitude to keep the voltage across, and consequently the current through, the right seal at a minumum. Therefore, since the outside of node N_{+1} is near ground and since there is little or no potential drop across node N_{+1} in the isosmotic potassium chloride solution, the entire inside of the internode will also be essentially at ground potential. That is, the *outside* of node N_0 is electronically driven up and down in potential to maintain the inside of this node at ground potential. Open-loop gains of up to 5000 have been used routinely to reduce the external leakage current by the same factor. Thus, feedback gives a faithful full-scale reproduction of the nodal membrane potentials with inverted polarity. In this arrangement, the stimulus may be a potential pulse applied to node N_{-1}.

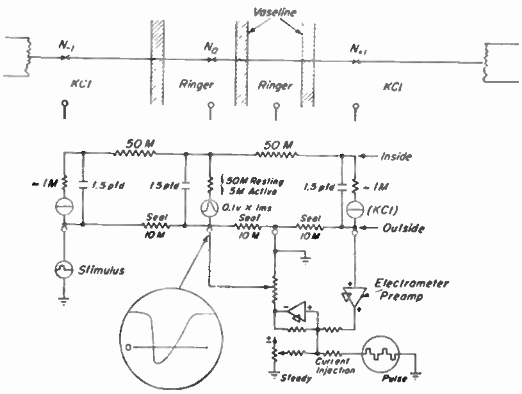


Fig. 13—More practical configuration of negative feedback to increase internodal resistance and give faithful reproduction (inverted) of the potentials across node N_0. This is equivalent to the circuit first described by Frankenhaeuser [27].

Because only a fraction of the potential across node N_0 is available, across the right gap, the effective feedback ratio is always

$$\leq \frac{10\text{ M}}{55\text{ M}} \geqq 0.2$$

and, therefore, the effective closed-loop gain is always ≥ 5. Amplifier noise and drift are indistinguishable from the input signal and so are also multiplied by the same closed-loop gain factor. This makes it necessary to use stable electrodes and places a premium upon quiet and stable amplifier performance and high seal resistances. The operational amplifier is chopper-stabilized by conventional techniques and we have recently chopper-stabilized the high input impedance electrometer preamplifier without introducing appreciable ripple. The stringent requirements may be somewhat relaxed when the feedback ratio is increased by carefully cutting the nerve in the far right chamber near the seal to reduce the 50 M internal internodal resistance. Because the use of feedback makes possible the equivalent of a very good contact at a point that is electrically and physically remote from the real electrodes, it has seemed appropriate to coin the term "electronic electrode" for this system.

Potential Control and Results

Having established the accuracy and validity of the measurement of potential across a node, Dodge and Frankenhaeuser [30] took the obvious next step of applying a potential clamp to this membrane (also del Castillo and Moore, unpublished). In one sense, this system is simpler than that of the squid axon because the active area is small and delineated and does not require a short-circuiting wire as was inserted into the squid axon. However, for this simplification, another difficulty is substituted in that the only obvious way to insert current is from the left node, N_{-1} as in Fig. 14. This path of current injection has the internal internode resistance of about 50 megohms, the effect of which is reduced by a factor of $2/(\mu+2)$ in a voltage clamp (see equivalent circuit in Appendix). This effective series re-

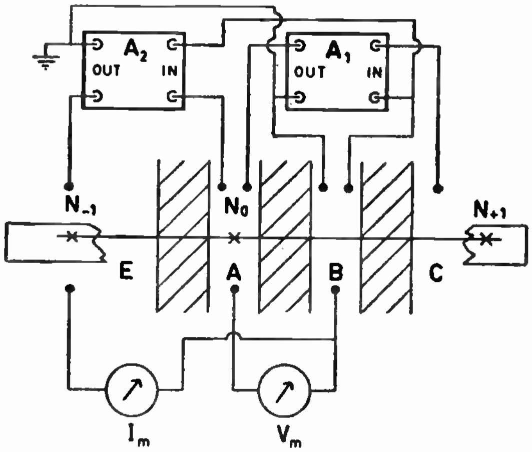


Fig. 14—Dodge and Frankenhaeuser [30] arrangement for clamping the nodal membrane voltage. Shaded areas are petroleum jelly seals around fiber. A_1 is a feedback amplifier for measuring the membrane potential of node N_0, and A_2 is a feedback amplifier for driving the nodal potential to the command value. Instruments for recording the membrane potential (V_m) and current (I_m) are indicated. (Figure reproduced, courtesy of authors and *J. Physiol.*)

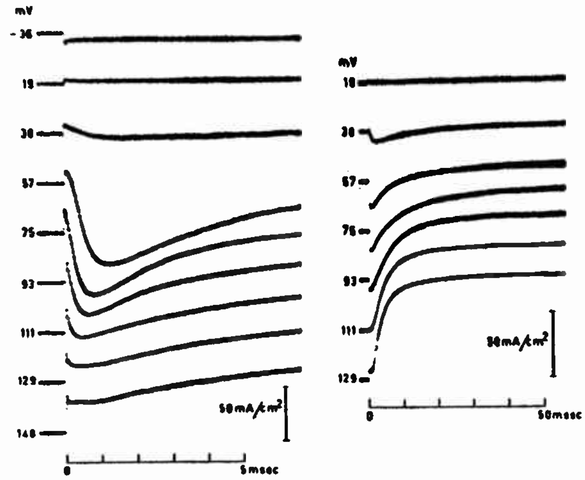


Fig. 15—Dodge and Frankenhaeuser [30] frog node membrane currents associated with step changes in potential. Left, membrane potential before pulse was minus about 15 mv beyond the resting potential; fast time base to show initial current. Right, experiment on another fiber with slow time base to show delayed outward current; potential before pulse was about 5 mv beyond the resting potential. Temp. 3°C. Capacity current is not visible in the records. (Figure reproduced, courtesy of authors and *J. Physiol.*)

sistance must be made small compared to the minimum nodal membrane resistance in order to obtain adequate and accurate control of the membrane potential. If the effective series resistance is to be one-tenth the minimum nodal resistance, which may be 1 megohm, a gain of 1000 is required. It is difficult to use large open-loop gains around such a circuit because of the lags introduced in the cable structure through which the current must be fed.

Dodge and Frankenhaeuser's experiments represent skill and achievement in face of difficult biological and electronic requirements. Their results are similar to those obtained in squid as is seen by inspection of Figs. 15 and 16.

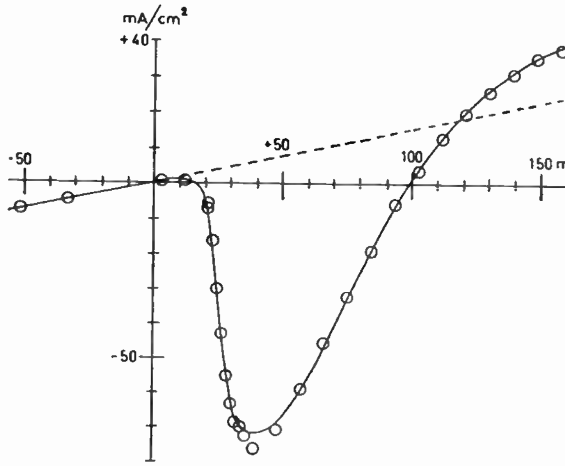


Fig. 16—Peak initial current density plotted against membrane potential during step. (Figure reproduced, courtesy of authors and *J. Physiol.*)

Discussion

Hodgkin and Huxley matched their experimental results on squid axon with a set of empirical equations in which the conductances in the potassium and sodium channels were nonlinear functions of voltage and had time constants of turning on and off that were also functions of voltage. Their formulation used linear first-order equations to generate nondimensional coefficients which were raised to powers as high as the fourth. These empirical functions were then substituted into the equations in Fig. 2 which describe the characteristics of the membrane when it is free to move in potential, as in a current clamp or for the case in which there is no short circuiting wire. Many phenomena of classical axonology, such as thresholds and action potentials, have been reproduced by computations [7, 31] using the Hodgkin-Huxley equations. In addition, excitation changes caused by variation of temperature and ionic medium, and even drastic shape changes such as very prolonged action potentials, have been reproduced by computations making use of these equations with rather straightforward variation of parameters (R. FitzHugh, in preparation). The accuracy and range of the predicted results are more than gratifying to even the most optimistic user of this technique.

What has been described here is a method of approach in which a very complex experimental technique has been substituted for classical physiological experiments with simple experimental arrangements and data requiring complex analysis. In return, rather direct and interpretable data are obtained which describe the ionic movements involved and which may give leads as to the underlying mechanisms.

It should be noted that the electrical characteristics of the nerve membrane giving rise to action potentials can be compared rather directly with those of some transistor switching circuits [32]. Both systems have a region of negative resistance bounded by regions of positive resistance. In the membrane case, the current is single-valued for any potential while variation of current may switch the potential back and forth around the negative resistance region. The negative resistance is quite transient in normal external sodium concentrations while it is rather stable with high external potassium [33] as seen in Fig. 17.

Possible Future Developments

It seems reasonable to expect, along with the advent of modern analog computation, that data of intermediate complexity between the classical electrophysiology and the voltage clamp may be taken and analog computers either working at slow time, on the recorded data, or at real time, while the experiment is in progress, may provide considerable (although not as complete) information on the individual ionic permeability changes. It also seems quite reasonable to expect the greatly increasing utilization of the "electronic electrode" technique for internal potential measurement by means of only external electrodes, because of its beautiful simplicity, accuracy, ease, and lack of injury to the cell under investigation.

Appendix

For purposes of illustration and deviation of an equivalent circuit, a simplified voltage clamp schematic is shown in Fig. 18. The feedback is from across an element whose potential is to be controlled. The impedance transformer is realizable in the form of an electrometer input preamplifier [11]. Applying Kirchhoff's law for the summing point, the following may be derived:

$$V = \frac{-E\mu}{\mu + 2} - \frac{R2I}{\mu + 2}.$$

An equivalent circuit is shown in Fig. 19 where the potential source is $E\mu/(\mu+2)$ with an effective series resistance of $R(2/(\mu+2))$. As the gain, μ, is increased, the voltage V approaches unity and the effective series resistance goes to zero. The resistance R', in series with the membrane to the ground, is outside the loop and, of course, is not reduced by the gain factor. The voltage across the membrane itself then deviates from V by the $R'I$ drop. If the value of R' is known, compensation for its voltage drop can be obtained by feeding back an additional signal to the summing point proportional to this error voltage (see potentiometer labeled negative resistance in Fig. 5).

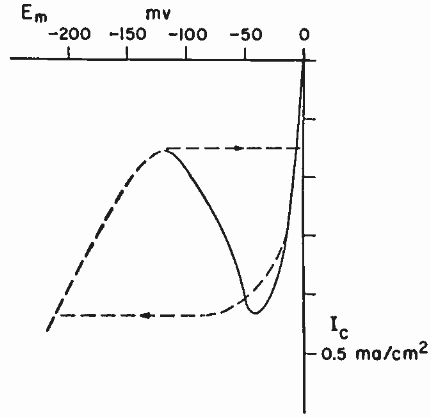


Fig. 17—Voltage-current characteristics of the squid axon membrane in 0.5 M potassium chloride. The continuous line was obtained when the potential was controlled and scanned slowly; the dashed sections were obtained when the current was swept. (Figure reproduced, courtesy of *Nature*.)

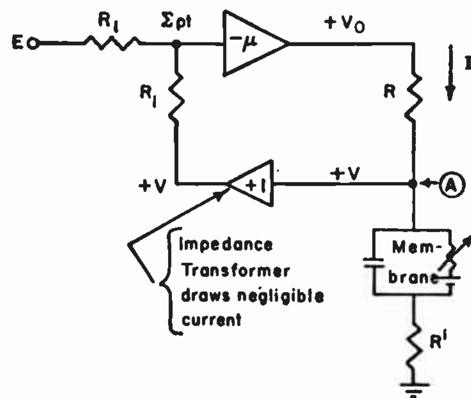


Fig. 18—Simplified and generalized circuit for control of the potential V by summing a command and feedback signal. R is a lumped resistance including the output impedance of the control amplifier, the electrode impedance, and the axoplasm impedance.

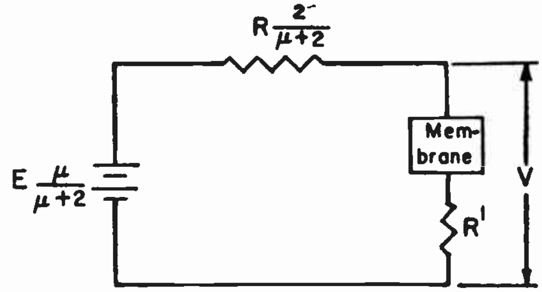


Fig. 19—Equivalent circuit of control system shown in Fig. 18. Because it is also the equivalent of a circuit where the command and feedback signals are compared by a differential amplifier, the sign of the source potential is omitted.

Acknowledgment

The author is indebted to Drs. K. S. Cole, J. del Castillo, W. J. Adelman, R. FitzHugh, R. E. Taylor, and Mr. J. Gebhart and appreciative of their collaboration, assistance, and discussions relative to the squid axon. The collaboration of Dr. J. del Castillo on the frog node as well as helpful discussions with Dr. B. Frankenhaeuser, Dr. J. Coombs and Mr. F. Dodge are gratefully acknowledged.

Bibliography

[1] H. P. Schwan, "Alternating current spectroscopy of biological substances," this issue, p. 1841.
[2] H. J. Curtis and K. S. Cole, "Membrane resting and action potentials from the squid giant axon," *J. Cellular Comp. Physiol.*, vol. 19, pp. 135–144; April, 1952.
[3] G. Ling and R. W. Gerard, "The membrane potential and metabolism of muscle fibres," *J. Cellular Comp. Physiol.*, vol. 34; pp. 413–438; December.
[4] R. D. Keynes and H. Martins-Ferreira, "Membrane potentials in the electroplates of the electric eel," *J. Physiol.*, vol. 119, pp. 315–351; 1953.

[5] J. D. Robertson, "New observations on the ultrastructure of the membranes of frog peripheral nerve fibers," *J. Biophys. Biochem. Cytol.*, vol. 3, pp. 1043–1047; November, 1957.
[6] R. Stämpfli and M. Willi, "Membrane potential of a Ranvier node measured after electrical destruction of its membrane," *Experientia*, vol. 13, p. 297; 1957.
[7] A. L. Hodgkin and A. L. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J. Physiol.*, vol. 117, pp. 500–544; 1952.
[8] A. L. Hodgkin, "The ionic basis of electrical activity in nerve and muscle," *Biol. Rev.*, vol. 26, pp. 339–409; November, 1951.
[9] L. A. Woodbury, J. W. Woodbury, and H. H. Hecht, "Membrane resting and action potentials of single cardiac muscle fibers," *Circulation*, vol. 1, p. 264; 1950.
[10] K. S. Cole and H. J. Curtis, "Electric impedance of the squid giant axon during activity," *J. Gen. Physiol.*, vol. 22, pp. 649–670; May, 1939.
[11] R. C. Gesteland, B. Howland, J. Y. Lettvin, and W. H. Pitts, "Comments on microelectrodes," this issue, p. 1856.
[12] R. D. Keynes, "The ionic movements during nervous activity," *J. Physiol.*, vol. 114; pp. 119–150; June, 1951.
[13] A. L. Hodgkin and R. D. Keynes, "Active transport of cations in giant axons from Sepia and Loligo," *J. Physiol.*, vol. 128, pp. 28–60; April, 1955.
[14] W. H. Freygang, "Some functions of nerve cells in terms of an equivalent network," this issue, p. 1862.
[15] R. D. Keynes and R. P. Lewis, "The sodium and potassium content of cephalopod nerve fibers," *J. Physiol.* vol. 114, pp. 151–182; June, 1951.
[16] B. A. Koechlin, "On the chemical composition of the axoplasm of squid giant nerve fibers with particular reference to its ion pattern," *J. Biophys. Biochem. Cytol.*, vol. 1, pp. 511–529; November, 1955.
[17] K. S. Cole and A. L. Hodgkin, "Membrane and protoplasm resistance in the squid giant axon," *J. Gen. Physiol.*, vol. 22, pp. 671–687; May, 1939.
[18] J. del Castillo and J. W. Moore, "On increasing the velocity of a nerve impulse," *J. Physiol.*, in press.
[19] G. Marmont, "Studies on the axon membrane," *J. Cellular Comp. Physiol.*, vol. 34, pp. 351–383; December, 1949.
[20] K. S. Cole, "Dynamic electrical characteristics of the squid axon membrane," *Arch. Sci. Physiol.*, vol. 3, p. 253; 1949.
[21] A. L. Hodgkin, A. F. Huxley, and B. Katz, "Measurement of current-voltage relations in the membrane of the giant axon of Loligo," *J. Physiol.*, vol. 116, pp. 424–448; April, 1952.
[22] I. Tasaki and W. H. Freygang, "The parallelism between the action potential, action current and membrane resistance at a node of Ranvier," *J. Gen. Phyisol.*, vol. 39, pp. 211–223; November, 1955.
[23] J. W. Woodbury, "Direct membrane resting and action potentials from single myelinated nerve fibers," *J. Cellular Comp. Physiol.*, vol. 39, pp. 323–339; April, 1952.
[24] I. Tasaki and K. Frank, "Measurement of the action potential of myelinated nerve fiber," *Amer. J. Physiol.*, vol. 182, pp. 572–578; September, 1955.
[25] R. Stämpfli, "Nouvelle méthode pour enregistrer le potentiel d'action d'un seul étranglement de Ranvier et sa modification par un brusque changement de la concentration du milieu extérieur," *J. de Physiol.*, vol. 48, pp. 710–714; 1956.
[26] R. Stämpfli, "A new method for measuring membrane potentials," *Proc. Internatl. Congress*, vol. 19, pp. 793–795; 1953.
[27] B. Frankenhaeuser, "A method for recording resting and action potentials in the isolated myelinated nerve fiber of the frog," *J. Physiol.*, vol. 135, pp. 550–559; 1957.
[28] A. F. Huxley and R. Stämpfli, "Direct determination of membrane resting potential and action potential in single myelinated nerve fibers," *J. Physiol.*, vol. 112, pp. 476–495; 1951.
[29] J. W. Moore and J. del Castillo, "An electronic electrode," 1959 IRE NATIONAL CONVENTION RECORD, pt. 9, pp. 47–50.
[30] F. A. Dodge and B. Frankenhaeuser, "Membrane currents in isolated frog nerve fiber under voltage clamp conditions," *J. Physiol.*, vol. 143, pp. 76–80; August, 1958.
[31] K. S. Cole, H. A. Antosiewicz, and P. Rabinowitz, "Automatic Computation of Nerve Excitation," *J. Soc. Indus. Appl. Math.*, vol. 3, pp. 153–172; September, 1955.
[32] Special Issue on Transistors, PROC. IRE, vol. 40, November, 1952.
[33] J. W. Moore, "Excitation of the squid axon membrane in isosmotic potassium chloride," *Nature*, vol. 183, pp. 265–266; January, 1959.

Measurement of Mechanical Properties of Muscle under Servo Control*

MARTIN LUBIN†, M.D.

***Summary*—Accurate measurement and analysis of the mechanical events in active muscle requires the use of high-speed equipment. A hydraulic servo-valve, controlled by analog units (integrators, adders, inverters), can be used to control the speed of shortening of muscle at rates as high as 1 mm per millisecond. The apparatus can be used for isometric, isotonic, and controlled release experiments. Both release and stretch, at high or low speeds, can be produced during a single contraction cycle. Force is measured by an unbonded strain gauge of high natural frequency and low compliance. To maintain constant force on the muscle, a signal proportional to measured force is fed into an error detector, whose output controls the servo-valve piston. The instrumentation described can provide the necessary and sufficient information to specify completely both transient and steady-state mechanical properties of muscle.**

* Original manuscript received by the IRE, September 2, 1959. The work described in this paper was supported by a grant from the Muscular Dystrophy Associations of America, Inc., and the U. S. Public Health Service (H-1498C and Senior Research Fellowship SF-83).

† Dept. of Pharmacology, Harvard Medical School, Boston, Mass.

INTRODUCTION

When a muscle is stimulated, it changes abruptly from a pliable, inactive body to a rigid, contractile structure. An active muscle will do work on an attached load, shortening at a rate which depends on the magnitude of the load. Under light loads, shortening is rapid; with heavy loads, shortening is slower. The nature of this adjustment of speed to load is a problem of major interest to physiologists. The mechanical information needed includes the steady-state relation between speed and load, and the rapidity with which one of these variables responds to a change in the other.

The purpose of this article is to discuss the use of a servo system for measurement of the mechanical properties of muscle. Some illustrative experiments on muscle are included, but details and interpretations will not be presented here.

Some Properties of Muscles

The requirements of measuring devices can be clarified by considering two well-studied muscles. 1) The sartorius muscle of the frog extends 3 to 4 cm from pelvic girdle to knee cap, and is composed of several thousand cylindrical cross-banded ("striated") cells which lie in parallel. This muscle, as a whole, is more a ribbon than a cylinder, and has cross-section dimensions of about 0.1 by 0.4 cm. 2) The anterior byssus retractor muscle of the edible mussel, *Mytilus edulis*, extends for 3 to 4 cm from the end of the shell to the byssus organ. From the byssus organ (*byssus*, L.: cotton) a fibrous tuft extends through the cleft in the shell and attaches the animal firmly to a rock, wharf, or another mussel. As in the frog sartorius, the cells in byssus retractor lie in parallel and extend the full length of the muscle. However, they are not cross-banded in appearance, but "smooth." A large specimen of this muscle has a cross section of about 0.1 cm², and develops a tension of nearly 200 grams.

Both sartorius and byssus retractor are unusual in this parallel arrangement of cells. In general, particularly in smooth muscle, cells are attached end to end.

A muscle isolated from an animal ordinarily remains in a resting state, but may be activated in a variety of ways, including electrical stimulation and the application of drugs. Attachment of the ends of muscle to a string, chain, or lever, necessary for any measurement of force, adds an inert elastic element which is in series with the muscle. By careful choice of materials, the compliance of the added attachments can be made small, and the inertial mass insignificant.

The residual component is the muscle, which will be considered only in its active state. The muscle is equivalent in its mechanical behavior to two elements arranged in series: 1) the series elastic element; and 2) the contractile component. The evidence which supports this subdivision is clear, and rests on the demonstration of a mechanical element whose stress is a function of strain, but is independent of the velocity of deformation. The properties of sartorius and byssus retractor, shown in Table I, determine the minimum requirements

TABLE I

Properties of Frog Sartorius and *Mytilus* Byssus Retractor Muscles

	Frog sartorius[1]	*Mytilus* byssus retractor[2]
Length	3 to 4 cm	3 to 4 cm
Maximum tension	50 to 100 grams	100 to 200 grams
Linearity of series elastic element	nonlinear	nonlinear
Approximate compliance (reciprocal elasticity)	0.05 cm/100 grams	0.2 cm/100 grams
Maximum velocity of contractile component	5 cm/sec at 0°C.	0.1 cm/sec at 10°C.

for any mechanical deformation and measuring system.

In the model shown in Fig. 1, the compliance of attachments to the muscle is lumped with the compliance of the muscle. The system is then composed of only two types of elements, elastic and contractile. The forces in each component are equal to each other and to the load P. The tension in the contractile component is then accurately measured by the tension in an attached measuring device, *e.g.*, the elastic element of a lever or strain gauge of sufficiently high-frequency response.

Stretch or release of the elastic element in frog sartorius at velocities greater than ten times the maximum velocity (5 cm per sec) of the contractile component will clearly separate the two components. The velocity of release required is 50 cm per sec. For a desired release of only 0.05 cm, which produces a large change in tension in the elastic element, this release must be over in 1 msec. For *Mytilus* byssus retractor, the requirements are less stringent. Release of 0.2 cm, at ten times the maximum velocity of the contractile component, requires a velocity of 0.2 cm in 200 msec [Figs. 1(D), 1(F), 8(b), 9(a), 9(b)].

In most muscles, including sartorius and byssus retractor, another elastic component lies in parallel with the contractile component. Fortunately, the tension exerted by the parallel elastic component becomes significant only at lengths greater than the natural length which the muscle has in the living animal. By restricting the investigation to lengths below this, the mechanical problem is simplified.

Muscles of widely divergent varieties, striated or smooth, taken from both vertebrate and invertebrate animals, have contractile components with very similar properties. For the contractile component of active muscle, force (or tension) is a function of the speed of shortening. The empirical steady-state relation between speed of shortening and the load is accurately described by Hill's equation:[3]

$$(P + a)(v + b) = b(P_0 + a). \qquad (1)$$

(P = force, v = speed of shortening, P_0 = *maximum* force which the muscle can develop at a given length, and a and b are constants.) At $P = 0$, the velocity of shortening is maximal, and is equal to $P_0 b/a$ (hence $a/P_0 = b/v_{max}$). Two qualifications must be added to Hill's equation: 1) The equation holds for muscle of a given length. If the muscle is stretched or permitted to shorten to lengths considerably different from that which it has in the intact animal, P_0 falls off significantly. At shorter lengths, Hill's relation will still accurately describe the behavior of muscle if the value of P_0 appropriate to each length is used.[4] 2) v is the steady-state value of velocity reached, corresponding to a given load. This value is

[1] B. R. Jewell and D. R. Wilkie, "An analysis of the mechanical components in frog's striated muscle," *J. Physiol.*, vol. 143, pp. 515–540; October, 1958.

[2] B. C. Abbott and J. Lowy, "Contraction in molluscan smooth muscle," *J. Physiol.*, vol. 141, pp. 385–397; May, 1958.

[3] A. V. Hill, "The heat of shortening and the dynamic constants of muscle," *Proc. Roy. Soc. (London) B*, vol. 126, pp. 136–195; October, 1938.

[4] J. M. Ritchie and D. R. Wilkie, "The dynamics of muscular contraction," *J. Physiol.*, vol. 143, pp. 104–118; August, 1958.

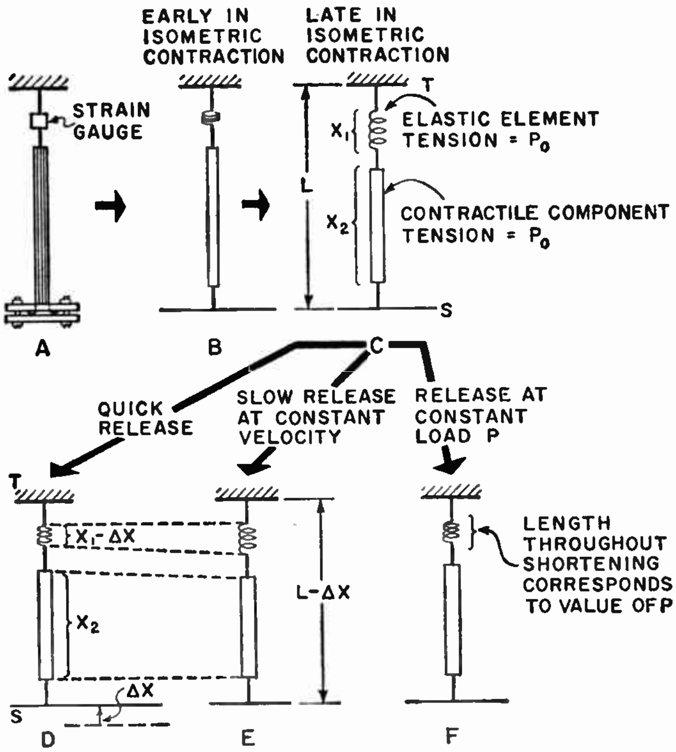


Fig. 1—Schematic picture of changes in length of elastic element and contractile component.

(A) *Mytilus* muscle is removed from animal and attached to rigid support at bottom and inflexible strain gauge at top. Muscle is made active by electrical stimulation (electrodes not shown).

(B) Contractile and inert elastic parts of muscle are lumped into two components. Elastic element shown includes attachments to the muscle (*e.g.*, string, tendinous tissue, elastic part of strain gauge), and elastic elements distributed throughout length of muscle.

(C) In isometric contraction, contractile component shortens while elastic element stretches an equivalent distance. Total length between S and T remains constant. At any instant, the velocity of shortening of contractile component depends on the load (see force-velocity relation shown in Fig. 2). When contractile component is attached to a spring, load is given by restoring force of spring, and increases as spring is stretched.

(D) Quick release produces shortening of elastic element only. Shortening of contractile component is too slow to produce appreciable change in length in a short time interval [also see Fig. 9(a)].

(E) Slow release, with shortening in both elastic element and contractile component [also see Fig. 9(c)].

(F) Under constant force P, elastic element rapidly changes to length corresponding to force, after which length of elastic element is constant. Slower contractile component shortens at speed determined by force-velocity relation [also see Figs. 2 and 8(b)].

reached very rapidly after stimulation, but not instantaneously.

The shape of the rectangular hyperbola (1) with asymptotes at $-a$ and $-b$ is determined by the constant a/P_0. In normalized form, the force-velocity relation becomes

$$\left(\frac{P}{P_0}+\frac{a}{P_0}\right)\left(\frac{v}{v_{\max}}+\frac{a}{P_0}\right)=\left(1+\frac{a}{P_0}\right)\frac{a}{P_0}. \quad (2)$$

For frog sartorius, a/P_0 and $b/v_{\max}$ have a value of about 0.25; for *Mytilus* byssus retractor, a/P_0 and $b/v_{\max}$ are about 0.2. Significant relations exist between the constants a and b, derived from measurements of

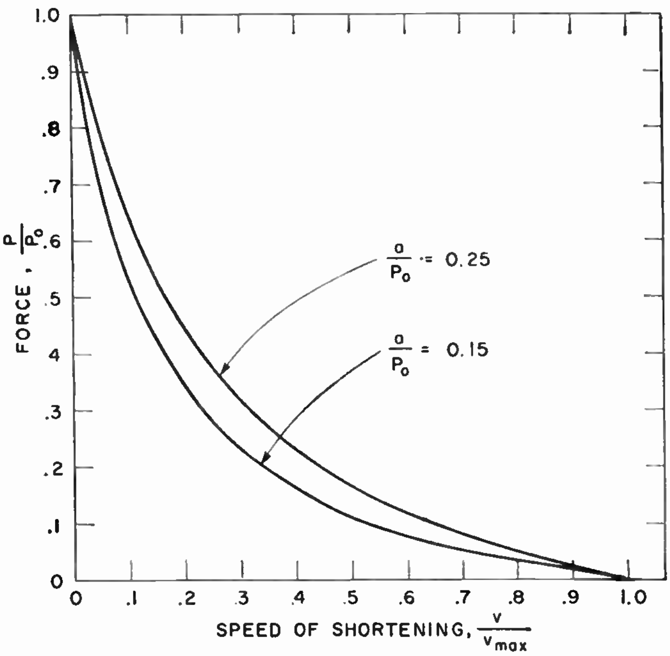


Fig. 2—Relation between load and speed of shortening. Curves drawn according to Hill's force-velocity relation with coordinates normalized [see (1)]

$$\left(\frac{P}{P_0}+\frac{a}{P_0}\right)\left(\frac{v}{v_{\max}}+\frac{a}{P_0}\right)=\left(1+\frac{a}{P_0}\right)\frac{a}{P_0}.$$

Two representative values of a/P_0 are chosen, to show the dependence of curvature on a/P_0.

the heat evolved during shortening. Further discussion of the force-velocity relation and many other vital aspects of the physiology of muscular contraction will not be considered in this paper, but are discussed in recent comprehensive reviews.[5–7] The data comprising the force-velocity relation, derived from experimental measurements, can be fitted as well by other functional forms; *e.g.*, a sum of exponentials. In several instances, these alternative forms have been derived from plausible models based on experimental data.[5,6,8]

General Methods of Measuring Mechanical Properties

There are two classes of control and measurement, either of which yield a complete description of the mechanical properties of a one-dimensional array of elements: 1) control of tension, with measurement of velocity of shortening; and 2) control of velocity of shortening, with measurement of tension.

A one-dimensional array of elements indicates that motion along only one axis of a three-coordinate system is considered. Bending or torsion of the muscle is not considered here, although there is reason to believe that

[5] D. R. Wilkie, "Facts and theories about muscle," in "Progress in Biophysics and Biophysical Chemistry," Academic Press, Inc., New York, N. Y., vol. 4, pp. 288–324; 1954.

[6] A. F. Huxley, "Muscle structure and theories of contraction," in "Progress in Biophysics and Biophysical Chemistry," Pergamon Press, London, Eng., vol. 7, pp. 255–318; 1957.

[7] "Physiology of voluntary muscle," *British Med. Bull.*, vol. 12, pp. 161–236; September, 1956.

[8] F. D. Carlson, "Kinematic studies on mechanical properties of muscle," in "Tissue Elasticity," Amer. Physiol. Soc., Washington, D. C., pp. 55–72; 1957.

significant information may be derived from a study of changes in torsional rigidity during contraction.[9]

Of the two methods, the first—control of tension—gives information on the force-velocity relation with simplicity. In the second method, the velocity controlled is the movement of one terminal of the whole system [the piston attached to point S in Fig. 1(C)]. The velocity of the contractile component, which is of prime interest, must be extracted by computation. In the classical method for determining the relation between force and velocity, active muscle is allowed to shorten under *constant* load ("isotonic load") and the velocity of shortening is measured. Precautions must be taken to reduce inertial forces to an insignificant level, and to prevent the load from stretching the resting muscle ("afterloading"). A series of experiments, each at a different load, determines the force-velocity relation.

Consider a muscle attached to a piston at one end, and a stiff strain gauge at the other (Fig. 1). The region between the two points (S, T) includes the elastic element of the strain gauge. If point S remains fixed, the system is of constant length ("isometric") between the fixed points (S, T). Displacements of the two components, elastic and contractile, are subject to the constraint $\Delta x_1+\Delta x_2=0$. For the elastic element, displacement is determined only by the force P; $\Delta x_1=cP$. For an elastic element which obeys Hooke's law, the compliance c is constant. In muscle, the value of c depends on P (nonlinear stress-strain curve).

Measurement of isometric tension is convenient, and often used by physiologists interested in comparing the maximum force developed by a muscle before and after the application of a treatment or drug. No direct information can be deduced about the force-velocity relation by isometric measurements alone, but any change observed in tension developed can be attributed to the contractile component if two assumptions are made: 1) the elastic elements remain unaltered during the experiment; and 2) the same number of cells in the muscle (usually all of the cells) are activated by stimulation.

If point S moves upward with a velocity V, the system now is neither isometric nor isotonic, but is described by the relation: V = velocity of shortening of elastic element + velocity of shortening of contractile component, or:

$$V = cdP/dt + v_{cc}. \tag{3}$$

The function v_{cc} represents an analytical description of the force-velocity relation of the *contractile component*. No assumption is made here about its nature, but in Hill's formulation for the steady-state relation, $v_{cc} = b(P_0-P)/(P+a)$.

If P is constant in an isotonic experiment, $cdP/dt=0$, and the velocity of point $S=v_{cc}$ [see Figs. 1(F), 8(b)]. This is analogous to the "voltage clamp" for the isolated nerve axon, described by Moore in this issue;[10] the total current across the nerve membrane I_m is related to the membrane capacity C, membrane potential E, and ionic current across the membrane I_i, by the relation:

$$I_m = C\frac{dE}{dt} + I_i. \tag{4}$$

If V is held constant in a "voltage clamp," $I_m=I_i$.

If the velocity of movement of point S [Figs. 1(E), 9(c)] is kept at a constant value V_0, the force-velocity relation may also be simply derived from (3), but with less accuracy than by a series of constant load experiments. The compliance c must be determined and dP/dt measured. The value of c is measured in a rapid release [see Fig. 9(a) and 9(b)].

The results of a series of experiments, each at a given load, or the results of a series of constant velocity releases, give identical results. A spurious difference between the two methods may exist unless the results which are compared refer to the contractile component *at the same length.*

In Fig. 8(b), control of force at a fixed level results in a rapid adjustment in length of the elastic element, followed by shortening of the contractile component at constant velocity. Fig. 9(c) illustrates the change in tension that occurs with a constant velocity release. In this case, a steady-state value of tension may not be reached for a considerable length of time, during which the contractile component shortens significantly [also see Fig. 1(E)].

An Electrical Analog

The mechanical changes in contraction may also be visualized by considering an equivalent electrical analog shown in Fig. 3. It is only one of several possible analogs.[11] The force exerted by the contractile component is, in general, a function of time, velocity of shortening, and length. Similarly, the element in the electrical analog containing the source voltage E and internal impedance Z has a voltage at its terminals which depends on time, current, and the amount of charge drawn from the system (see Table II on the next page).

Some Instruments Currently Used in Muscle Physiology for Mechanical Measurements

Methods used for mechanical measurement and control of muscle have reached an advanced level of refinement. Some are suitable for controlled velocity experiments, others for control of force. The Levin-Wyman ergometer, a device used for controlled velocity release or stretch, consists of a sturdy lever, pulled by a weight

[9] O. Sten-Knudsen, "Torsional elasticity of the isolated cross striated muscle fibre," *Acta Physiol. Scand.*, vol. 28, suppl. 104, pp. 1–240; 1953.

[10] J. W. Moore, "Electronic control of some active bioelectric membranes," this issue, p. 1869.

[11] M. F. Gardner and J. L. Barnes, "Transients in Linear Systems," John Wiley and Sons, Inc., New York, N. Y., vol. 1; 1942.

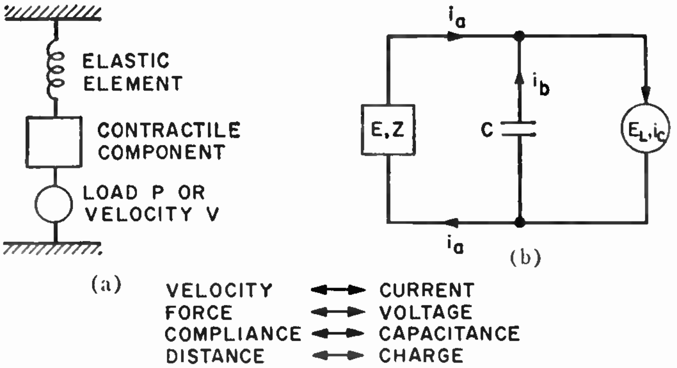


Fig. 3—(a) Mechanical equivalent of muscle, with only two elements, elastic and contractile. Either velocity controlled and force measured, or force controlled and velocity measured (see Figs. 5, 7). (b) Electrical analog of mechanical model. See Table II and text for discussion.

TABLE II
ELECTRICAL MODEL ANALOGOUS TO MUSCLE MODEL

Mechanical	Electrical
1) Contractile component.	1) Source voltage with internal impedance Z.
2) Elastic element with compliance c in series with contractile component.	2) Capacitance C in parallel with source.
3) Servo piston velocity V.	3) Current i_c.
4) Load P = force in elastic element.	4) Load voltage E_L = voltage across capacitance C.
5) Stretch of elastic element, $\Delta x = cP$.	5) Charge Q on capacitance $= CE_L$.
6) Velocity of stretch of elastic element $= cdP/dt$.	6) Current through capacitance, $i_b = CdE_L/dt$.
7) Velocity of contractile component, v_{cc}.	7) Current i_a.
8) $V = v_{cc} + cdP/dt$.	8) $i_c = i_a + CdE_L/dt$.
9) For P = constant (isotonic load), $V = v_{cc}$.	9) For E_L = constant, $i_c = i_a$.
10) For $V=0$ (isometric contraction), $v_{cc} = -cdP/dt$; *i.e.*, contractile component shortens as elastic element lengthens [see Fig. 1(B) and (C).	10) For $i_c=0$, $i_a = -CdE_L/dt$.

or spring, and damped by an oil-filled dashpot.[12,13] The velocity of motion can be changed by manual adjustment. In the hands of A. V. Hill, his colleagues and his students, it has provided a considerable amount of precise information on muscle mechanics. The force of contraction can be conveniently measured by a strain gauge or an anode transducer (RCA 5734). The problem of using an appropriate transducer for measurement of force will not be considered here, although it is an essential part of the instrumentation. In isotonic contraction, the velocity of shortening may be measured by the displacement of a lightweight vane mounted on the attachments to the muscle. Movement of the vane exposes more or less of a phototube to an illuminated source, and the phototube current is then proportional to the position of the vane.[13]

A description of apparatus capable of precise measurement of isometric and isotonic contraction in frog sartorius muscle, with provision for quick release, is described in recent papers by Wilkie, and Jewell and Wilkie.[14,1] In addition, Jewell and Wilkie discuss in detail refinements in the technique of mounting a muscle which significantly affect mechanical measurements. Carlson has successfully used a feedback-controlled galvanometer movement to provide constant force loading of muscle.[15] With refinements, this device can also be used for rapid and controlled shifts in the magnitude of load applied to the muscle. Many experiments, particularly those concerned with changes in rate or amplitude in slowly contracting (usually smooth) muscle, can be simply and satisfactorily performed with a lever and smoked drum!

On the other hand, some important mechanical properties of muscle cannot be conveniently determined with the methods referred to above. For example, the question of how rapidly muscle responds to sudden changes in either force or velocity has not been examined extensively. The answer bears vitally on the nature of the mechanism which underlies contraction. The instrumentation needed to examine this requires close control of velocity and direction of deformation of the muscle. A hydraulic servo-valve can provide the required transduction of a set of controlling electrical signals into mechanical motion.

HYDRAULIC SERVO-VALVE CONTROL OF FORCE OR VELOCITY

The servo-valve[16] and the method of mounting the muscle are shown in Fig. 4. The piston maintains a constant position for zero input voltage to the torque motor. Valves controlled by the torque motor regulate the flow of hydraulic fluid. A signal voltage to the torque motor results in a minute shift in valve position, and the change in pattern of fluid flow causes the piston to move swiftly to a new position. Position is detected by a signal derived from a differential transformer, whose core is rigidly coupled parallel to the piston. The position signal is fed back to the servo amplifier, and the loop gain adjusted for optimum performance. The circuit for the servo amplifier is nearly identical with that of a commercially available amplifier.[17] When control signal to torque motor is constant, piston "give" is only 0.007 mm per 100 grams of load.

The piston of the servo-valve has been machined down to a cross-sectional area of 0.35 square inch to reduce needed fluid flow. With 100 pounds per square inch of fluid pressure, the response to a pulse is a constant

[12] A. Levin and J. Wyman, "The viscous elastic properties of muscle," *Proc. Roy. Soc.* (*London*) *B*, vol. 101, pp. 218–243; April, 1927.

[13] A. V. Hill, "The series elastic component of muscle," *Proc. Roy. Soc.* (*London*) *B*, vol. 137, pp. 273–280; July, 1950.

[14] D. R. Wilkie, "The mechanical properties of muscle," *British Med. Bull.*, vol. 12, pp. 177–182; September, 1956.

[15] Carlson, *op. cit.*; also, unpublished data.

[16] The hydraulic servo-valve used for experiments illustrated in this paper is similar to units made by Midwestern Instruments, Tulsa, Okla. Strain gauge is model PT45-10 (natural frequency unloaded, 3000 cps); Dynamic Instrument Co., Cambridge, Mass.

[17] The servo amplifier used is nearly identical to Model 12 amplifier, Midwestern Instruments. A signal proportional to piston position was picked up in the feedback loop of the servo amplifier. In the Model 12 amplifier, almost 10 per cent of the command input signal appears at this point, but a cathode follower inserted in the loop provides sufficient isolation of the position signal.

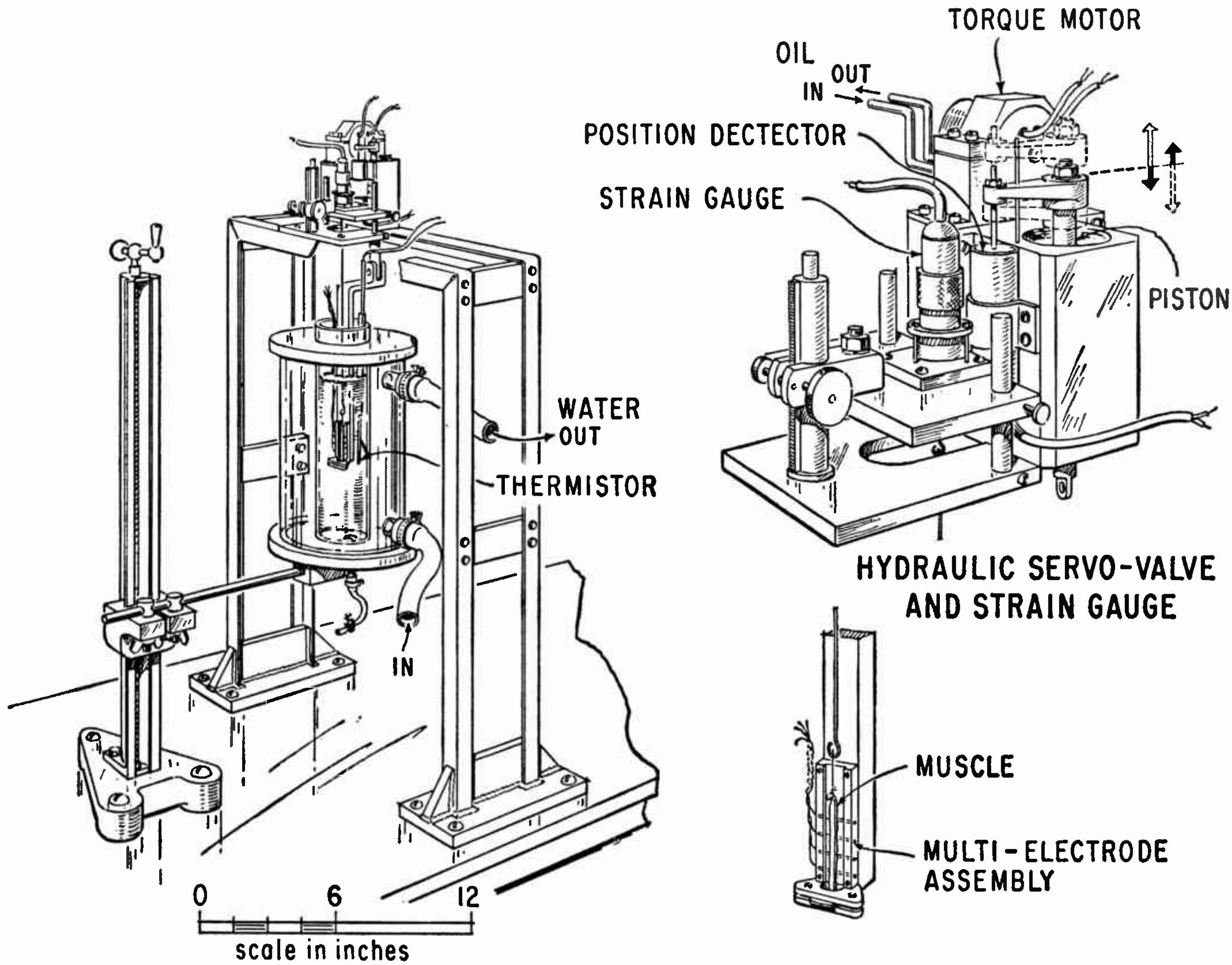


Fig. 4—Apparatus for measurement of mechanical properties of muscle. Oil pump and lines to servo-valve not shown. Electrodes for stimulation of muscle shown in "multi-electrode assembly" consisting of silver wires embedded in plastic. Inner lucite cylinder surrounding muscle contains Ringer solution. Outer jacket holds temperature-controlled water.

velocity movement of about 1 cm in 10 msec. In the unit now in use, optimal gain adjustment for large excursions results in sluggish response for movements of less than 0.05 cm. Improved models of servo-valves now commercially available have considerably improved stability and frequency response.[18]

Vibration caused by the pump (not shown in Fig. 4), which drives the fluid under pressure, has not been a problem. The plate holding strain gauge and servo-valve are rigidly coupled. Oil leak may be a nuisance to muscle physiologists. It cannot be entirely prevented, and the problem is solved by the artful use of small drip trays. The hydraulic equipment will not be described; pumps and pressure regulators can be obtained to suit commercially available servo-valves.

Controlling signals were derived from Tektronix[19] waveform generators (series 161, 162). To obtain signals in tandem from a series of 162 generators, a transformer was used to change triggering signal polarity. The tail end of the "gate out" signal of one generator was used to trigger the next. Square pulses from the Tektronix 162 generators were integrated on analog units manufactured by G. A. Philbrick Researches, Inc.[20] In addition, Philbrick modules were used to construct plug-in adders, inverters, gain controls, voltage crossing-detectors, and other functional units as needed.

Integration of a square wave, which results in a ramp (see Fig. 5), is ideally suited for constant velocity release of muscle (Fig. 9). A saw-tooth generator will not do. Rapid stretch of an active muscle will often cause irreversible changes in the muscle and may promptly terminate an experiment long in preparation. With a ramp, on the other hand, the muscle remains at a new short length after a constant velocity release, and can be slowly restretched after activity in the muscle has

[18] Servo-valves and actuators may be purchased from a number of companies. Among those in the U. S. are: a) Midwestern Instruments, Tulsa, Okla.; b) Pesco Products Div., Borg Warner Corp., Bedford, Ohio; c) Moog Valve Co., East Aurora, N. Y.

[19] Tektronix, Inc., Portland, Ore.

[20] "Applications Manual for Computing Amplifiers," Eds., George A. Philbrick Researches, Inc., Boston, Mass.; 6th printing, 1959.

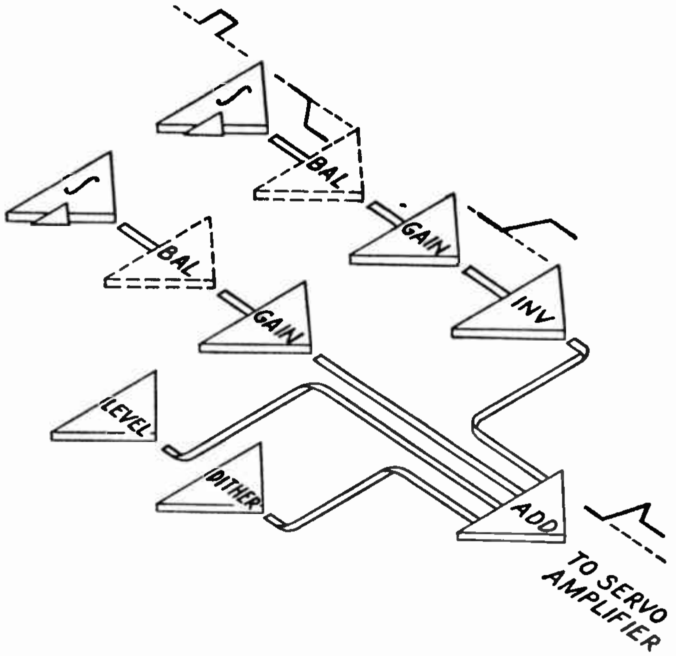


Fig. 5—Arrangement of operational amplifiers for constant velocity release or stretch. Operational amplifiers shown are K2W Units (G. A. Philbrick Researches, Inc.). Integrator uses chopper-stabilized K2P in tandem with K2W (double triangle indicates stabilization). Units are modular, with ten amplifiers in one chassis. Balance, shown by dotted lines, is optional, and is used to offset voltage of electronic switch which shorts integrator (see Fig. 6). Balance unit is a cathode follower, using Philbrick amplifier; offset voltage produced by bias control on amplifier. Functional units are shown separately, although some can be combined (*e.g.*, gain control and balance). Dither to servo-valve may be part of some commercial servo amplifiers; here 400-cps dither is fed in separately. Level unit supplies dc signals for changing initial length of muscle. All units except cathode follower invert signal in addition to performing specific functions. Synchronization of signals discussed in text.

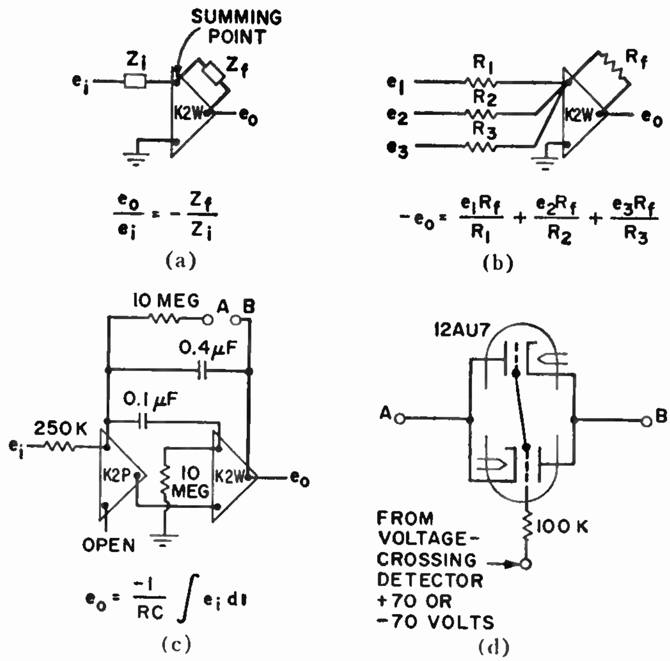


Fig. 6—Philbrick circuits. For further applications, see the literature.[20] (a) Operational amplifier, generalized. (b) Adder, inverter, or gain control. (c) Stabilized integrator. (d) Electronic switch.

subsided. The time of restretch is set by another Tektronix 162 generator, which controls a voltage-crossing detector.[20] The detector operates an electronic relay which shorts the capacitance (0.4 μF) of the integrator (Fig. 6). A resistance of 10 megohms limits the speed with which the output of the integrator, and hence the piston, returns to its initial state.[21]

Any desired combination of stretch or release may be obtained by feeding suitably-timed ramp signals into an adder (Fig. 5). Polarity of any ramp is changed by feeding the signal through an inverter (adders invert, as do the integrators and gain controls used). The electronic switch used for shorting the integrator output has no unusual virtues except speed and low cost (Fig. 6). On closed circuit, the switch contributes a half-volt bias which must be "bucked out." High-speed mechanical switches may be preferred for many applications. In addition, the use of a switch to keep the voltage output of the integrator at zero except during a release prevents very slow changes in piston position (and muscle length) which might occur if the integrator output slowly drifts.

The arrangement for controlled force involves a servo loop (Fig. 7). A signal from the strain gauge is

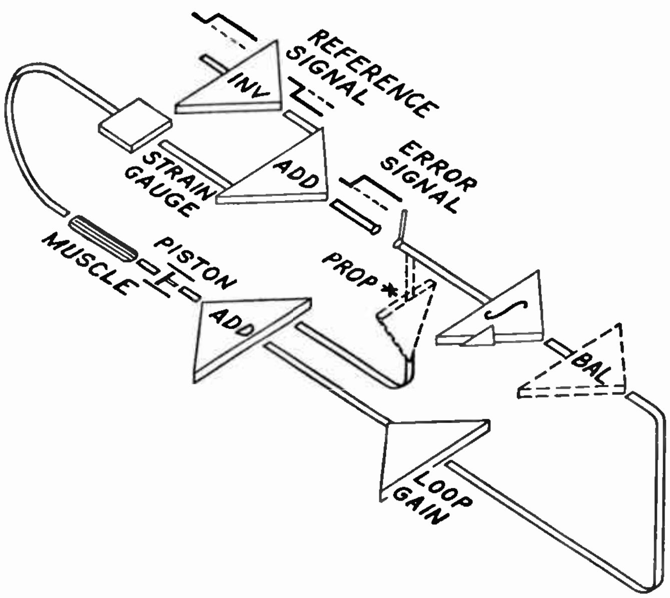


Fig. 7—Arrangement of operational amplifiers for constant force. See legend to Fig. 5. "Prop*" indicates optional unit with output proportional to error signal; polarity must be the same as signal from integrator loop.

fed into an adder which also receives a reference signal (which may be of any polarity). The sum, or "error," is integrated. If the error signal has any value other than zero, the integrated signal causes the piston to move. This results in a change in signal from the strain gauge, and a reduction in error signal. The reference signals are derived from Tektronix 162 generators, and added and synchronized at will. Signals proportional to the error signal and its time derivative may also be needed for optimum performance with fast muscle.[22]

[21] For experiments on smooth muscle, the time scale of the Tektronix 162 waveform generators may be extended beyond the maximum of ten seconds by simple modifications.

[22] "A Palimpset on the Electronic Analog Art," H. M. Paynter, Ed., George A. Philbrick Researches, Inc., Boston, Mass.; 1955.

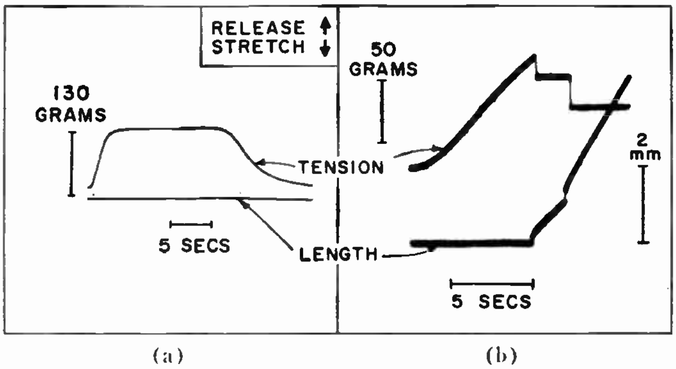


Fig. 8—Experiments on *Mytilus* muscle. Photographs of oscilloscope tracings. (a) Development of isometric tension following electrical stimulation. Stimuli 4 msec long, 5 per second, for about 20 seconds. Temperature 10°C. (b) Shortening under constant load in active muscle. System at first is isometric while tension develops. Load then set at constant level (see Fig. 7). Lower curve shows quick adjustment in length of elastic component, followed by nearly constant velocity of shortening. After two seconds, load shifts to lower value. After another adjustment of length in elastic element, velocity of shortening of contractile component is even higher [see Fig. 1(F)]. Note the rapid adjustment of velocity to force (for a similar result, see Jewell and Wilkie[1]). Photograph retouched over faint areas.

Fig. 8 shows the results of a controlled force (isotonic) experiment, performed on a muscle kept in an active state by repeated electrical stimulation. The rapidity with which the magnitude of force may be changed from one level to another without overshoot has not been fully determined for this system. It is rapid enough for *Mytilus* muscle. Refinements in servo-control, using analog computer and control units, may be necessary if these methods are to be used for frog sartorius muscle.

Other Applications to Muscle Physiology

In addition to the applications outlined above, several other experiments can be performed with servo-valve equipment.

A. Measurement of the Duration of the Active State After a Single Shock

A single electrical stimulus, applied to frog sartorius muscle, evokes a transient activation which has been called the "active state."[23] In the electrical analog of Fig. 3, this is equivalent to brief activation of the source voltage E. The voltage across the capacitance C does not reach the value E unless activation is repeated. In muscle, repeated fused activation is called a "tetanus." After a single stimulus, which causes a "twitch" response, the active state is too short, and the velocity of shortening of the contractile component too slow, to stretch the elastic element a sufficient distance for the full value of P_0 to appear in the strain gauge element. Measurement of the duration of the active state has in the past been performed in at least two ways—by quick stretch shortly after the stimulus,[23] and by release at various times after the stimulus.[24] With a servo-valve, stretch or release can be readily accomplished by the methods indicated earlier in this paper.

[23] A. V. Hill, "The 'plateau' of full activity during a muscle twitch," *Proc. Roy. Soc. (London) B*, vol. 141, pp. 498–503; September, 1953.

[24] J. M. Ritchie, "The effect of nitrate on the active state of muscle," *J. Physiol.*, vol. 126, pp. 155–168; October, 1954.

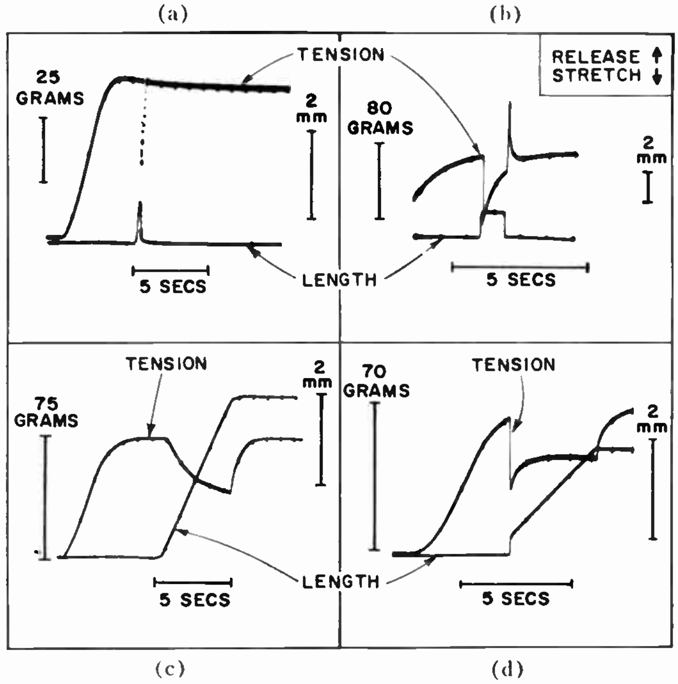


Fig. 9—*Mytilus* muscle.

(a) Tension develops as in Fig. 8(a) as a result of repeated electrical stimulation. With quick release and immediate restretch, elastic element shortens and lengthens without appreciable change in contractile component. Darker dots indicate a 1-second time marker applied to intensity control of oscilloscope. Most of the compliance is in the muscle, and only a small amount in the attachments to the muscle.

(b) Quick release causes drop in tension, followed by redevelopment of tension as muscle is held isometrically at shorter length. A quick stretch then causes tension rise in elastic element. Tension now exceeds maximum tension (P_0) which contractile component can sustain. As a result, contractile component lengthens while elastic element shortens (note appreciable time required for drop of tension). Also note that changes in tension accompanying quick release and quick stretch are about equal, indicating approximately linear compliance of elastic element in this range of stress-strain curve. Photograph retouched over faint areas in Fig. 9(b) and (d).

(c) Slow release after muscle has developed full tension isometrically. Shape of curve depends on force-velocity relation and stress-strain curve of elastic element.[4] At shorter length, system is again isometric and tension redevelops.

(d) Three changes in piston position are produced after isometric contraction: 1) quick release causes drop in tension; 2) slower, constant velocity release which follows results in constant force in muscle, according to force-velocity relation; and 3) very small, quick restretch (barely visible on tracings) then pulls on elastic element, and is equivalent to increasing the load. Further shortening of contractile component when muscle is held isometrically results in slight rise in tension to maximum value. Other experiments show that with an appropriate quick stretch, tension rises up to and remains at maximum value (P_0) without overshoot.

B. Change of Length of Resting Muscle

A rack and pinion, to which the strain gauge is fixed, has been used for length adjustment on mounting the muscle. During an experiment, it is convenient to change the initial length of a muscle reproducibly. An additional voltage applied to the adder controlling the servo-valve causes a change in piston position and hence in muscle length. A set of stable voltages is conveniently derived from the +300, −300 power supplies of the Philbrick operational manifold, and a lag introduced to produce very slow (up to a time constant of 1 second) stretch or release, by adding a capacitance in parallel with the feedback resistor of an adder.[20]

C. Measurement of Work Done by Muscle

If an active muscle is loaded isotonically, it will do work on the load. The muscle relaxes when the stimulus is turned off, returning to its resting state gradually. During relaxation, a falling load does work on the muscle. By "catching" the load at the peak of its traverse, no work will be done on the muscle during relaxation. In a servo system with appropriate control signals, the piston can be held at the final position reached, and restretch delayed until the muscle returns to its inactive state.

D. After-Loading

For isotonic measurements with appreciable loads, a resting muscle may be stretched appreciably by the load. Ordinarily, this is prevented by supporting the load until the muscle becomes active and develops sufficient tension to lift it. With a servo-valve, the load can be prevented from stretching resting muscle by controlling the time at which a constant force is applied [as in Fig. 8(b)]. Alternatively, the servo-piston can be constrained to move unidirectionally when the signal to it exceeds a reference value.[20]

E. Measurement of Contractile Component when Parallel Elastic Component Contributes Significant Force

The principles of measurement are the same as outlined; *i.e.*, for a one-dimensional array of components, measurements of instantaneous forces and velocities are necessary and sufficient to specify completely the mechanical description of the system.

Although a range of lengths of muscle can be found in which the parallel elastic element is negligible but the series elastic element is significant, in general the converse is not true. At lengths considerably greater than the natural length of the muscle in the body, the parallel elastic element is under considerable tension. The series elastic element is probably distributed in a complex way along the muscle, and at longer lengths the series elements can no longer be simply lumped.

A simplification would occur if the series elastic elements were so rigid that they could be considered as stiff connectors. The electrical equivalent then would consist of a tension generator E with internal impedance Z [as pictured in Fig. 3(b)] but with a capacitance in *series* with E, Z, which represents the parallel elastic element.

The measured force, in such a mechanical system, is given by the sum of the parallel parts, elastic and contractile: $P_{\text{total}} = P_{\text{elastic}} + P_v$, in which P_v represents the force-velocity relation of the muscle. A sudden change in velocity of the piston, for example from zero to some finite velocity of release, will produce a change in P_v. If the force in the contractile component follows a change in velocity rapidly, the drop in measured force will be sharp, and followed by a further, more gradual change in force as the tension in the parallel elastic component decreases. The shape of the curve describing the changes in force will be similar to the displacement tracing of Fig. 8(b).

The Nature of the Contractile Component and General Comments

Mechanical measurements lead to a relation between force and velocity in the contractile component. By considering the slope of this relation at any point, an equivalent mechanical rectilinear resistance may be formally calculated. Similarly, in the electrical analog of Fig. 3, the voltage-current relation of the source will have a shape determined by a nonlinear internal resistance. It is important to recognize one misleading quality of this electrical analogy, in which electrical resistance is analogous to mechanical viscosity. There is considerable evidence, based on thermal measurements, against the idea that the force-velocity relation derives its shape from an internal viscosity.[25,26] The force-velocity relation is more likely determined by an incompletely understood mechano-biochemical coupling between energy sources in the cell and the polymeric contractile proteins. Much information is available on these biochemical processes, in addition to thermal measurements and structural analyses, the interpretation of which must ultimately be consistent with the mechanical properties of muscle.

The usefulness of the servo system described in this paper lies in its versatility. All methods needed for measuring mechanical properties of muscle, which have in the past required the use of a variety of instruments, can be performed by one device, the servo-valve and actuator system, controlled by electrical signals. For experiments on some striated muscles in which shortening is very rapid, servo systems of the highest frequency response available must be used. This system has the additional virtue of applicability to the investigation of the contractile mechanism in its transient state, which has received only partial study by physiologists because of the difficulties of instrumentation.

Acknowledgment

The author is grateful to J. O. Silvey of the Massachusetts Institute of Technology Servomechanisms Laboratory for considerable assistance with the servo-valve assembly and loan of equipment. D. N. Sheingold of G. A. Philbrick Researches, Inc., made a number of suggestions on use of analog units. The author also is indebted to Dr. F. D. Carlson for reading the manuscript.

[25] B. C. Abbott, X. M. Aubert, and A. V. Hill, "The absorption of work by a muscle stretched during a single twitch or a short tetanus," *Proc. Roy. Soc. (London) B*, vol. 139, pp. 86–104; December, 1951.
[26] B. C. Abbott and X. M. Aubert, "Changes of energy in a muscle during very slow stretches," *Proc. Roy. Soc. (London) B*, vol. 139, pp. 104–117; December, 1951.

Scanning Microscopy in Medicine and Biology*

LESLIE E. FLORY†, FELLOW, IRE

Summary—**Scanning microscopy provides the means for extending the range of usefulness of the light microscope in several directions. In addition to the convenience of viewing a large bright image on the television type monitor, it can also enhance the contrast of faintly visible specimens and can extend the convenience of direct observation into the infrared and ultraviolet. By electrical processing of the video signal, a great deal of quantitative information can be extracted. This method has been used to determine the number, size, and size distribution of particles in a field and to quantitate absorption of biological materials for visible and ultraviolet light.**

Introduction

EVER since the invention of the microscope men have been trying to extend its usefulness and overcome its limitations. In its early use, the microscope was limited to a visual observation of the specimen. Several limitations in this utilization are apparent. The observation is limited to one viewer, the illumination is limited to the visual regions of the spectrum, there is no means for quantitation of the absorption of the specimen except by purely subjective estimation, and one has to be satisfied with the contrast which was present in the original specimen.

Through the years many things have been done to remove some of these limitations. Each of these improvements, while it obviously extended the utility of the microscope, usually introduced a new set of limitations of its own. For example, to enable more than one observer to view a microscope image, a microprojector was used. In order to get a picture large and bright enough, high light intensities were required, resulting in heating of the specimen. Photography permitted, through densitometry, a quantitation of the absorption, and at the same time allowed the use of light outside the visible spectrum. Contrast enhancement was also possible in the photographic process. However, the time element in photography was a limitation.

In recent years, a new tool has appeared which greatly enhances the usefulness of the microscope. This tool is electronic scanning and image reproduction. Electronic scanning, which is called scanning microscopy when used with the microscope, overcomes many of the limitations of the visual microscope and, at the same time, adds a new dimension to the information obtained from the specimen.

By means of electronic amplification and image reproduction it is possible to produce without high light intensities on the specimen, a picture as large and bright as desired for visual observation. This greatly facilitates the observation by groups of people and is of particular advantage in teaching.

* Original manuscript received by the IRE, August 10, 1959.
† RCA Laboratories, Princeton, N. J.

Photosensitive devices can be made responsive to a wide range of wavelengths from the far ultraviolet to the far infrared. This permits the instantaneous observation of specimens with a wide spectrum of illumination, and, by means of highly selective systems, even the observation at two or more wavelengths simultaneously. Where a photographic record is desired, the higher sensitivity of photo-electric processes permits, through the electronic system, photographic recording with less light exposure than direct photography. This is particularly important in the ultraviolet.

Because any degree of electronic amplification may be used, limited only by signal-to-noise ratio, a high degree of contrast enhancement may be accomplished, permitting the study of specimens of low contrast without staining.

Finally, the scanning process introduces a completely new dimension into the picture, that of amplitude or intensity vs time. This gives an instantaneous quantitative measure of the light intensity at each successive scanning point of the image, presented in a time sequence. This facility opens up untold possibilities in the processing of the information present in the specimen and permits analyses not possible to obtain by any other direct means.

The concept of the use of television-type scanning techniques to extend the usefulness of the light microscope was first reported by Zworykin in 1933.[1] Many years passed, however, before the state of the art in photosensitive pickup devices had advanced to a point to permit the construction of a useful instrument. It was not until 1951 that reports of actual uses began to come in.[2–5]

Types of Scanning Microscopes

By this time it was also obvious that there were two approaches to scanning microscopy, not considering the scanning electron microscope.[6] One of these approaches is similar to the conventional television camera in which the image from the microscope is projected on the target of the camera tube and the electrical signal extracted by scanning the target surface. In this system, the area of the specimen being viewed is illuminated continuously

[1] V. K. Zworykin, "Electric microscope," *Primo Congr. Intern. Electroradio biologia*, vol. 1, pp. 672–686; September, 1934.

[2] J. Z. Young and F. Roberts, "A flying spot microscope," *Nature*, vol. 167, p. 231; February 10, 1951.

[3] L. E. Flory, "The television microscope," *Cold Spring Harbor Symp. on Quantitative Biology*, vol. 16; 1951.

[4] A. K. Parpart, "Television microscopy in biological research," *Science*, vol. 113, pp. 483–484; April 27, 1951.

[5] V. K. Zworykin and L. E. Flory, "Television as an educational and scientific tool," *Science*, vol. 113, p. 483; April 27, 1951.

[6] V. K. Zworykin, J. Hillier, and R. L. Snyder, "A scanning electron microscope," *ASTM Bull.*, vol. 117, pp. 15–23; August, 1942.

by an illuminator. While the scanning beam is on a particular area of the target for only a very short time during each frame, the information in the image is being integrated continuously by the storage properties of the pickup tube so that the illumination is effective all of the time. Fig. 1(a) illustrates this type of scanning microscope. One of the greatest advantages to this approach is the flexibility provided by its ability to use any conventional type of illumination in any desired region of the spectrum to which the pickup tube is responsive.

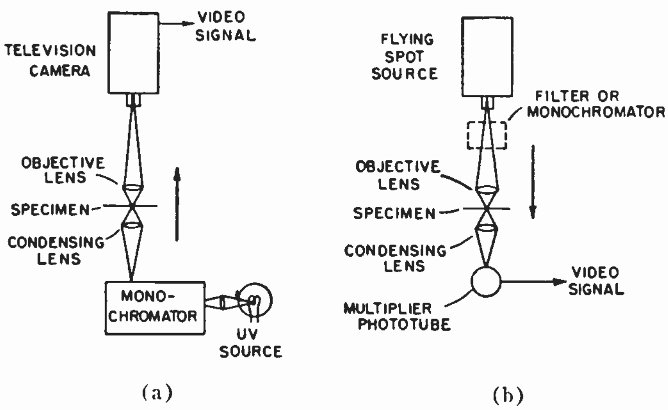


Fig. 1—The two common types of scanning microscope. (a) The camera tube type; (b) the flying spot type.

In Fig. 1(b) we illustrate the second approach to a scanning microscope, known as the flying spot microscope. As can be seen, the optics and general arrangement are similar to the camera tube microscope. The chief difference lies in the method of illumination and scanning. In the flying spot system, the light is generated in the phosphor of a cathode ray tube. This phosphor is scanned by a cathode-ray beam in a raster exactly like that of a conventional picture tube in a television receiver, except that in this case the beam is unmodulated. An image of the scanning raster is projected upon the specimen by the objective lens of the microscope. This constitutes the illumination of the specimen. The light transmitted by the specimen is collected by the microscope condensing lens and projected on the sensitive area of a multiplier phototube where the resulting photocurrent is amplified to constitute the video signal. This signal will be identical in character to that from the camera amplifier in the previous system. The over-all systems, it can be seen, are quite similar; the light sources and detectors are essentially interchanged. There are, however, some important differences in operation.

In the flying spot system, the specimen is illuminated by the tiny moving spot of light so that each elementary area is under illumination only during the time the signal is being generated. In the camera tube system, the speciman is illuminated continuously. However, in the camera tube system the image is being stored on each elemental-area throughout the time during which the beam is not discharging that particular area. Assuming ideal systems which generate no noise in themselves, the limit of sensitivity would, in both cases, be set by noise in the primary photo emission. Under this assumption, the *average* photocurrents and, consequently, the average illuminations are the same. This requires that the intensity of the flying spot must, for the same average current, be greater by the number of picture elements (or spot areas) in the raster than the illumination on the camera tube, since in the latter case the illumination is continuous and hence is already averaged over the scanning time.

Neither system can be ideal, of course, and many practical factors enter into the situation. A discussion of the relative merits of the two systems has appeared in the literature,[7] so it will suffice here to mention only some of the most important advantages of each.

In the flying spot microscope, the light is generated by electron bombardment of the phosphor in the cathode ray tube. A particular wavelength band may be chosen by interposing a filter or monochromator somewhere in the light path. Light generated by this method is precious compared to that from conventional light sources such as those used with the camera tube microscope, and with present phosphors it is difficult to obtain sufficient light to utilize narrow spectral bands, especially in the ultraviolet or infrared.

Spurious signals are less a problem in the flying spot microscope and the response is more linear than most camera tubes so that it offers some advantage in making precise measurements.

The amount of illumination which can be used from any source with living material depends upon the killing power of the light itself. While, as was pointed out, in ideal systems the average illumination would be the same with either system, there is some evidence that reciprocity does not hold and that intermittent light of the same average intensity is less lethal than continuous light. This appears to be due in part to partial recovery of the cell material during the interval between exposures in the flying spot system.[8] The same effect may be obtained in a different way by flash illumination of a storage type system, as used by Williams for time lapse photography.[9] Also, slow-speed scanning may be used with either system to reduce the over-all exposure.

Uses of Scanning Microscope

Aside from the utility of the scanning microscope in producing a bright, easily-viewed picture of practically any size, its greatest advantages lie in its ability to convert the invisible image in an ultraviolet or infrared microscope to a visible one and in the flexibility afforded

[7] E. G. Ramberg, "A theoretical analysis of the operation of flying spot and camera tube microscopes in the ultraviolet," IRE Trans. on Medical Electronics, vol. ME-12, pp. 58–64; December, 1958.

[8] P. O'B. Montgomery and W. A. Bonner, "The ultraviolet flying spot television microscope," IRE Trans. on Medical Electronics, vol. ME-6, pp. 143–146; September, 1959.

[9] G. Z. Williams, "Time lapse ultraviolet television microscopy instrumentation and biological applications," IRE Trans. on Medical Electronics, vol. ME-6, pp. 68–74; June, 1959.

by processing the electrical signal obtained by conversion of the three dimensions of the original image (two spatial dimensions plus intensity) into a time-variant amplitude-modulated wave. The two facilities are not mutually exclusive and many of the later users make use of both these useful properties.

The biological importance of ultraviolet in the study of cell structure has long been known. It can be seen very graphically by a glance at the absorption curves of some cellular materials as a function of wavelength in Fig. 2. It will be noted that none of these materials shows absorptions at wavelengths longer than 3000 A. This means that they are all transparent to visible light. However, they all show absorption in the region 2200 to 2800 A.

An ultraviolet television microscope with useful sensitivity was first described in 1950.[3] This instrument used a vidicon pickup tube with an ultraviolet transmitting face plate and a specially sensitized photoconductor. Such a microscope uses a standard closed-circuit television system and an ultraviolet monochromator. A typical response curve of an ultraviolet-sensitive vidicon is shown in Fig. 3. This curve indicates good sensitivity in the region of biological interest.

This type of microscope, because of its efficiency in the ultraviolet and its flexibility, has been used in a number of laboratories. Special mention should be made of the work of Dr. George Z. Williams who has worked with equipment of this type since its introduction.[9] Williams' interest is in the study of metabolic processes in living cells. Heretofore, the study of the orientation and localization of reactions within individual living cells has been limited by difficulties in methods. The ultraviolet spectrum has been used by many investigators to detect specific materials in cells, but available methods have required rather long exposures to ultraviolet, rapidly damaging fresh cells. The flexibility and amplification provided by electronic devices make it possible to examine living cells by low-intensity ultraviolet microscopy with less injury, and to utilize absorbing chemical indicators for observation of intracellular reactions. One group of such indicators are the tetrazolium salts which are relatively nontoxic in low concentrations. This substance enters the cell, is reduced to an insoluable formazan which precipitates immediately *in situ* and is detectable in very small amounts by ultraviolet absorption. For this work, Williams says the ultraviolet television microscope possesses the following advantages:

1) The facility for searching and focusing with visible light to avoid cell injury. After focusing, a photographic record of the television image may be made with a very short pulse of ultraviolet.
2) The use of the ultraviolet-sensitive vidicon permits pulses as short as 1/100 second. The contrast of the absorption image may be electronically enhanced to intensify faintly visible structure.

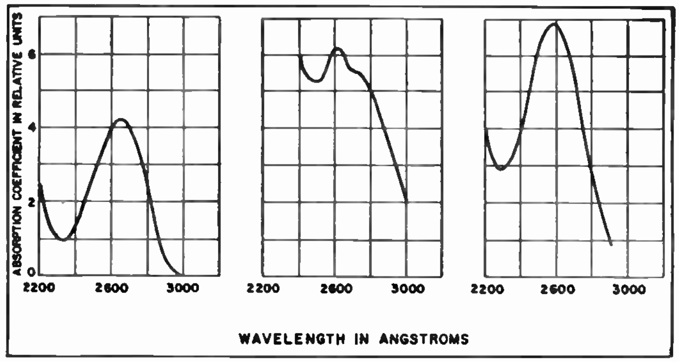


Fig. 2—Typical ultraviolet absorption curves of some biological materials (from Glasser, "Medical Physics").

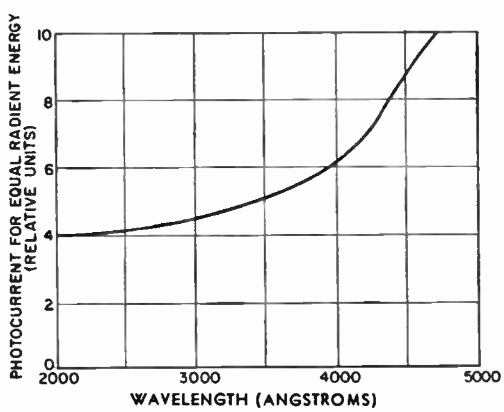


Fig. 3—Typical spectral response curve of ultraviolet vidicon.

3) Approximate values of reduction rates may be determined from records of deflections of a single TV line traversing the cell image as depicted on a line-selecting oscilloscope.

By properly synchronizing the television scan and the application of light pulses, the progress of reactions in living cells has been observed for several hours in this time-lapse manner. A block diagram of Williams' equipment is seen in Fig. 4. A trace from his line selector oscilloscope showing three levels of absorption is shown in Fig. 5.

Dr. P. O'B. Montgomery, with the same objective of studying living cells under ultraviolet illumination, is using the flying spot approach.[8] Because of the superior efficiency of the photomultiplier tube in the flying spot scanner over photographic emulsions, an absorption image is obtained by photographing the television monitor using considerably less dosage than that necessary to obtain a comparable image with a photographic plate.

Since the specimen is illuminated by a scanning spot of light rather than constant total illumination employed in photography, the specimen is irradiated only one picture element at a time, and the same picture element is not re-exposed until a second picture is required. By employing this sequential picture element irradia-

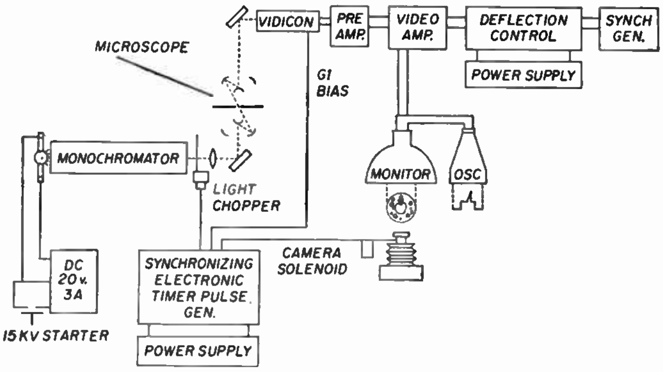


Fig. 4—Block diagram of ultraviolet television microscope for time lapse photography. (Courtesy Dr. G. Z. Williams.)

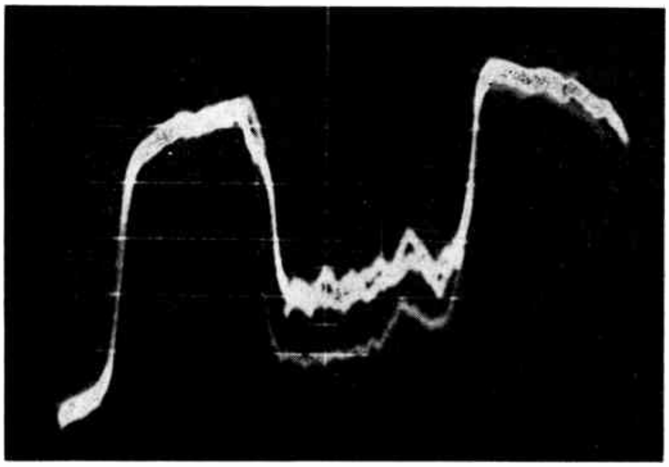


Fig. 5—Three superimposed oscilloscope tracings of one liver cell at 5, 10 and 30 minutes after adding tetrazolium. The successively lower deflections indicate progressive increase of formazan formation.

tion and using the slowest possible rate of repetition, the cell is allowed the maximum opportunity for recovery even though the total energy introduction may be large.

By proper control of the scanning process, any rate of time lapse can be obtained with exposures from 1/20 second to two and a half hours and intervals between exposures from 1/20 second to 25 hours.

By means of a gating circuit, it is possible to brighten any portion of the scanning raster down to one-micron square. The nucleus of a living cell could thus be damaged or destroyed by intensely irradiating it alone while normal ultraviolet absorption images of the remainder of the cell were simultaneously being obtained. This provides a means for obtaining time-lapse motion pictures of a living cell during the process of nuclear damage or destruction. The converse of this technique may be applied by protecting any portion of the field down to one micron while the remainder is intensely irradiated.

Fig. 6 is a block diagram of Montgomery's flying spot microscope. A special flying spot tube was developed for this microscope. An emission curve for this tube is shown in Fig. 7. It should be noted that the peak of emission occurs in the region of biological interest.

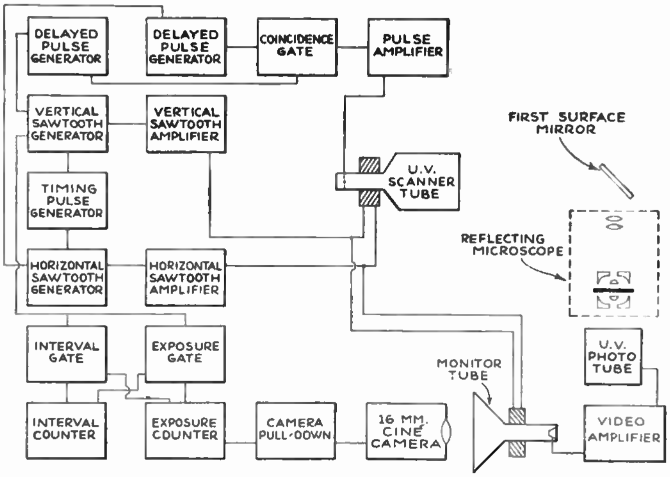


Fig. 6—Block diagram of variable speed flying spot microscope. (Courtesy Dr. P. O'B. Montgomery and W. A. Bonner.)

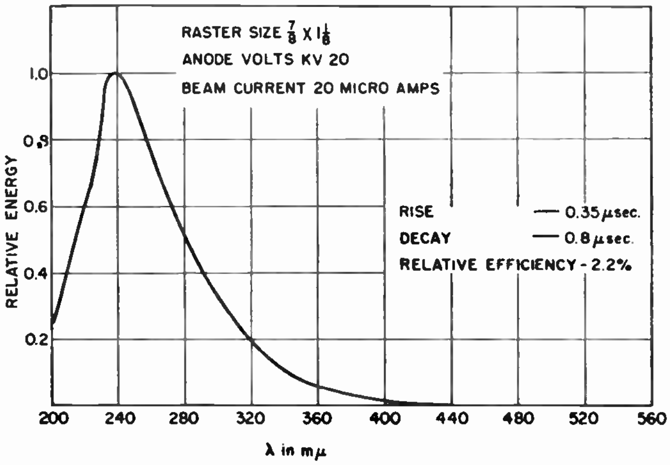


Fig. 7—Emission curve of special cathode ray tube for flying spot microscope. (Courtesy Lansdale Tube Co.)

With this equipment, HeLa cells and salamander cells have been observed to pass through their entire cycle of mitosis with no damage even though subjected to continuous ultraviolet scanning. Fig. 8 shows an ultraviolet absorption image of a HeLa cell. The bright square is a one micron area of intensified irradiation.

Signal Processing

The point-by-point scanning of the microscope image provides, in time sequence, complete information concerning the light intensity of every point in the field. Any object appearing in the field is completely described by the electrical signal, which will contain information giving the size, location, shape and light distribution of the object. The technique of line selection, long known in standard television techniques, has already been mentioned as a means of obtaining a cross section of the absorption pattern of a cell.

The entire scanning raster carries much more information, however, and considerable work has been done to devise equipment to extract this information in useable form. For example, it is obvious that information

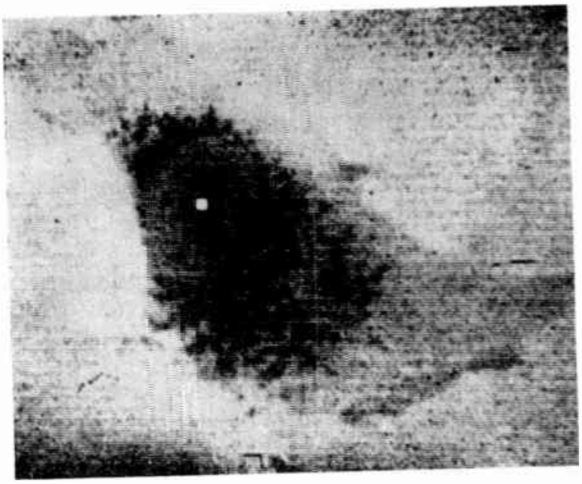


Fig. 8—Ultraviolet absorption image of a living HeLa cell. Bright square over nucleus is a one micron square area of focal irradiation for the purpose of producing highly selective damage.

giving the total number of particles, blood cells for example, exists in the signal. A simple pulse counter, however, will not give a true count since each particle is scanned by several lines and the number of scans per particle is a function of the size and shape of the particle. Young and Roberts[10] early proposed a particle counting system using the flying spot microscope which has proved quite successful. In this system, the field is scanned by not one but two spots, one line apart, the light from the two spots being received by separate phototubes. The output from the amplifiers is fed to an anti-coincidence circuit which is arranged to give an output only when one of the spots is obscured but not the other. This will obviously occur twice during the scanning of any once particle, one as the scans first reach a particle and again when they leave it. The circuit is arranged to pass one of these pulses to a counter, thus recording the true number of particles.

A block diagram of the flying spot microscope arranged for double spot particle counting is shown in Fig. 9. Each particle as seen on the monitor picture is marked by a brightening pulse to indicate that it is being counted.

Another approach to particle counting, particularly adapted to the counting of regular particles such as red blood cells, is the sanguinometer.[11] This device makes use of the camera tube microscope (although it could be adapted to the flying spot system) and corrects for multiple scans by measuring the average pulse length, then dividing the total count by this figure in a simple analog computation. This method, of course, is limited to regular-shaped particles of restricted size range but has given very good results on red blood cells. A block diagram of the equipment is shown in Fig. 10.

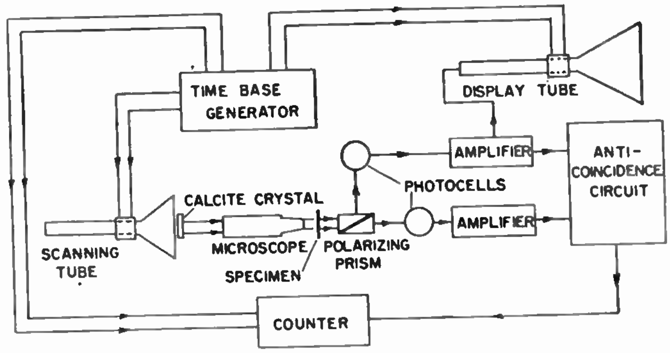


Fig. 9—Flying spot microscope arranged for double spot particle counting system. (Courtesy Dr. J. Z. Young.)

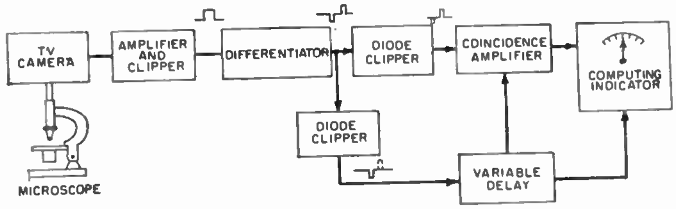


Fig. 10—Block diagram of sanguinometer for counting red blood cells. Compensation for size is by measurement of average pulse length (diameter of cell) and correcting the total count accordingly.

The waveform of the electrical signal from a scanning microscope contains information giving not only the total number of particles but also complete information regarding the size and shape of each. Further work has been done by Taylor on obtaining size and size distribution measurements with the flying spot microscope.[12,13] Taylor uses the two spot system of Young and Roberts with the modification that a single spot is made to follow a double line path in alternating steps. If the switching is done at a frequency at least double the highest video frequency, the result is as if there were two scanning spots.

Size discrimination is achieved by circuits that modify the video signals obtained from the two lines before they are applied to the counter. The circuits simply delay the signal from the leading edge of the particle by a variable amount *D*. When *D* is such that the delayed leading edge would fall beyond the trailing edge, the particle is rejected with the result that the modified video signal corresponds to a "remainder particle" bounded by the delayed leading edge and the trailing edge. It will now be apparent that when *D* is equal to or greater than the size of the particle in the direction of scan, there will be no remainder particle and no count. A size histogram is developed by taking the differences of successive counts as *D* is increased from zero to the value at which the count is just zero. By rotating the specimen and re

[10] J. Z. Young and F. Roberts, "High speed counting with the flying spot microscope," *Nature*, vol. 169, p. 963; June, 1952.

[11] L. E. Flory and W. S. Pike, "Particle counting by television techniques," *RCA Rev.*, vol. 14, pp. 546–556; December, 1953.

[12] W. K. Taylor, "An automatic system for obtaining particl size distribution with the aid of the flying spot microscope," *Bri J. Appl. Phys.*, suppl. no. 3, pp. 173–175; 1954.

[13] W. K. Taylor, "Flying spot microscope," to appear in "Ne Approaches in Cytology," Academic Press, Inc., New York, N. Y 1959.

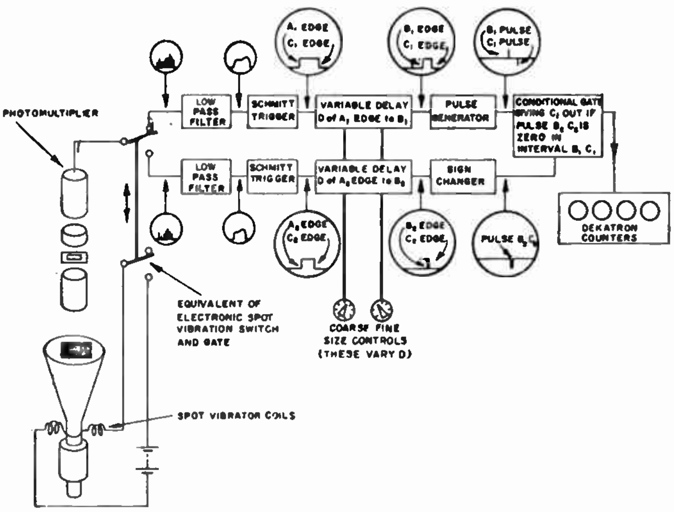


Fig. 11—Block diagram of flying spot microscope with size discriminating circuits. (Courtesy Dr. W. K. Taylor.)

peating the process, the effect of particle, shapes can be studied. A block diagram of Taylor's size discriminating system is shown in Fig. 11.

Use of Color

Techniques of color television can be applied to the scanning microscope. Recent work with small transistorized color television systems has resulted in an instrument particularly adapted for use with a microscope to reproduce natural or stained slides in full color. Pathologists have been particularly interested in color microscopy. A color TV microscope is now in use in the pathology laboratory of Walter Reed Army Hospital in Washington, D. C. The reproduced picture is visible not only to the pathologist but to the surgeon in the operating room, permitting an immediate exchange of information in case of a biopsy, although some distance separates the operating room and the laboratory.

The possibilities of color go beyond this use, however. Color translation is a process whereby three visible wavelengths of light (*i.e.*, ultraviolet) are used to produce visible images which are then presented as three primary visible colors to produce a full color display. This process was first used with photography as the translation means.[14]

The same translation can be obtained instantaneously by television.[15] If, in a television microscope, the specimen is illuminated during successive fields with light of three different wavelengths in the ultraviolet, there are generated three sets of television signals, one for each wavelength, the three sets occurring in sequence. By keying a color television monitor in synchronism with the sequencing mechanism, a color display representing the ultraviolet absorption spectrum is produced. Such a translation microscope has been built at Rockefeller Institute for Medical Research in New York City.[16] It is now in operation in medical and biological investigations.

Resolution

A word should be said about the resolution limitations of a system of scanning microscopy. A television type system will always have a limit in resolution. This limit is set in part by the number of scanning lines in the system, the bandwidth of the amplifying system, the spot sizes in the pickup, flying spot, and reproducing tubes and by spurious signals such as amplifier noise, grain size of phosphors and leakage effects in certain types of pickup tubes. However, by using sufficient optical magnification it is always possible to work within the resolution limits with no sacrifice except field, while realizing the full resolving capabilities of the optical microscope.

Acknowledgment

Such a review as this must of necessity cover the work of many people. Many important contributions known to the author could not be included because of lack of space. No doubt a greater number of equally important contributions are unknown to the author. Acknowledgment is hereby made of the cooperation of Dr. P. O'B. Montgomery and W. Bonner of the Southwestern Medical School of the University of Texas, Dallas; of Dr. G. Z. Williams, chief of the Clinical Pathology Department, National Institutes of Health, Bethesda, Md.; of Prof. J. Z. Young, and R. Roberts of the Anatomy Department, University College, London, England; and of Dr. W. K. Taylor, also of University College, London.

[14] E. H. Land, E. R. Blout, D. S. Grey, M. S. Flower, H. Husek, R. C. Jones, C. H. Matz, and D. P. Merrill, "Color translating ultraviolet microscope," *Science*, vol. 109, pp. 371–374; April 15, 1948.

[15] V. K. Zworykin, L. E. Flory, and R. E. Shrader, "Ultraviolet television microscopy," *Electronics*, vol. 10, pp. 150–152; September, 1952.

[16] V. K. Zworykin and F. L. Hatke, "Ultraviolet color translation microscope," *Science*, vol. 126, pp. 805–810; October 25, 1957.

Instrumentation for Automatically Prescreening Cytological Smears*

R. C. BOSTROM†, MEMBER, IRE, H. S. SAWYER†, MEMBER, IRE, AND W. E. TOLLES†, SENIOR MEMBER, IRE

Summary—**Mass-screening application of the cytological smear for the detection of cervical cancer has been limited by a lack of technicians to screen the smears. By using an instrument to identify automatically those smears which are clearly negative, the effectiveness of the technician could be greatly increased. A quantitative analysis of a large number of smears showed that positive smears usually had a small number of cells with abnormally large and intensely stained nuclei that did not appear on negative smears. On the basis of this analysis, an experimental instrument—called the Cytoanalyzer—has been constructed.**

The Cytoanalyzer scans a smear, measures the size and light absorption of approximately 10,000 cells on the smear, classifies each cell normal or abnormal according to its nucleus size and absorption, and totals the number of cells falling into each classification. The smear is then classified normal or abnormal depending on the fraction of cells having abnormal characteristics.

Preliminary tests with the Cytoanalyzer have been very promising. In a test of approximately 1000 smears, 65 per cent of the premenopause smears and 35 per cent of the postmenopause smears were properly identified. Plans are now underway to make a more thorough test of the screening capabilities of the instrument. If the test is successful, development of a clinical instrument will be started.

Introduction

In the past decade the value of the cytological smear in the early detection of cancer has become well established. In particular, the cytological smear when used according to the techniques established by Papanicolaou[1] has become the most important method for the early detection of one of the most prevalent types of cancer—cancer of the cervix. Cervical cancer can be controlled if it is detected in an early stage. Early detection, however, dictates that smears be taken periodically from the entire female population susceptible to the disease. Unfortunately, the problems of examining the large volume of smears obtained in such a routine mass-screening program has limited the use of this detection method. There is not a sufficient number of technicians available to screen smears in large quantities; the training of new technicians is expensive and time consuming; the task of screening a large volume of smears, only a small fraction of which are suspicious, is extremely tedious and conducive to technician error. Furthermore, the high cost associated with screening the entire population has presented a serious economic problem.

Mellors and Silver[2] first suggested the use of an automatic prescreening instrument to eliminate those smears which were clearly negative. The technicians could then devote all their attention to screening the smears having some degree of suspicion. Even if only 50 per cent of the smears were so eliminated, the effective screening rate of each technician would be doubled. Because such an instrument appeared to offer potentially great benefits, the development of an automatic prescreening instrument was started. The development program led to the construction of an experimental instrument called the Cytoanalyzer. The instrument is currently undergoing evaluation in a series of field tests.

This paper describes three phases in the development program: 1) the quantitative measurements of cells in cytological smears to establish a basis for distinguishing between negative and positive smears, 2) the modification of the standard Papanicolaou smear to make it more suitable for instrument analysis, and 3) the development of an experimental instrument to prove the feasibility of automatic prescreening. A fourth phase—developing a clinically usable instrument—will be started after its feasibility has been conclusively demonstrated.

Quantitative Studies of Cytological Smears

An important characteristic used by the cytologist in identifying cancer cells is that the cancer cells have abnormally large and intensely stained nuclei.[3] To determine whether quantitative measurements of this characteristic could be used for instrument screening, measurements were made on a representative sample of cells from a large number of negative and positive smears.[4] A manual microphotometer was used to measure the average diameter and light absorption of the stained nuclei. As was expected, the positive slides nearly always had a small number of cells that had abnormally large nuclei with relatively high light absorption.

Fig. 1 shows the appearance typical of the types of epithelial cells found on smears prepared from cervical specimens. On the normal smears, most of the cells have a nucleus diameter between 5 and 11 microns. The light

* Original manuscript received by the IRE, August 10, 1959. This work is being supported by the National Cancer Institute, Bethesda, Md., and the American Cancer Society, New York, N. Y.

† Dept. Medical and Biological Physics, Airborne Instruments Lab., Division of Cutler-Hammer, Inc., Mineola, N. Y.

[1] G. N. Papanicolaou and H. F. Traut, "Diagnosis of Uterine Cancer by the Vaginal Smear," Harvard University Press, Cambridge, Mass.; 1943.

[2] R. C. Mellors and R. Silver, "A microfluorometric scanner for the differential detection of cells: application to exfoliative cytology," *Science*, vol. 114, pp. 356–360; October, 1951.

[3] G. N. Papanicolaou, "Atlas of Exfoliative Cytology" (published for the Commonwealth Fund), Harvard University Press, Cambridge, Mass.; 1954.

[4] W. J. Horvath, W. E. Tolles, and R. C. Bostrom, "A study of the quantitative characteristics of exfoliated cells from the female genital tract" (in press).

absorption of the nucleus decreases as the size of the nucleus increases. This suggests that the total quantity of stain bound to the nucleus remains nearly constant. In addition to the normal-cell population there are on the positive smears abnormal cells with nuclei up to 18 microns in diameter and with absorption levels that indicate a basic increase in the quantity of stain bound to the nucleus.

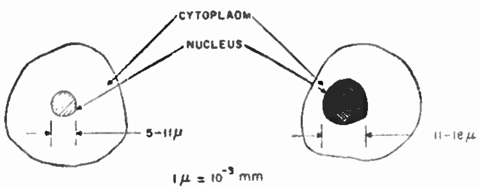


Fig. 1—Typical normal (*left*) and abnormal (*right*) cells on a cytological smear.

The cumulative distribution of the percentage of abnormal cells for a group of smears from premenopause subjects is shown in Fig. 2. These smears were collected by cervical swabs and prepared by the Papanicolaou technique. In this example, a cell was defined as abnormal if its nucleus absorption was greater than 40 per cent and its nucleus diameter was greater than 11 microns. The abscissa is the percentage of abnormal cells on a smear; the ordinate is the cumulative percentage of smears having an abnormal-cell percentage equal to or exceeding the value shown on the abscissa. Separate curves were plotted for both the negative and positive smears. From this plot it can be seen that over 90 per cent of the positive smears have greater than 0.5 per cent abnormal cells, whereas less than 35 per cent of the negative smears have that many abnormal cells. Therefore, if 0.5 per cent abnormal cells is selected as the screening criterion, approximately 65 per cent of the negative smears will be correctly classified negative but nearly 10 per cent of the positives will be falsely classified. The screening performance for other values of the screening criterion was calculated in a similar manner; the results are plotted in Fig. 3. A screening-performance curve for a group of postmenopause smears is also plotted.

The screening-performance curves show that it is impossible to obtain perfect screening using the cervical swab type of smear because there are always some positive smears that do not have a sufficiently large percentage of abnormal cells. The curves also show that the performance with the postmenopause smears is not as good as with the premenopause smears. The inferior performance of the postmenopause group is caused by an increased percentage of abnormal cells on the negative smears. These abnormal cells are the result of atrophic changes in the cervix which often occur after menopause.

Even without perfect performance, instrument prescreening was considered worthwhile if at least one-half of the negative smears could be screened without missing more than 5 to 10 per cent of the positive smears. The manual photometer data indicated that this level of performance could be achieved. Additional data showed that nearly perfect performance could be obtained by using a different collection method known as cervical scraping,[4] because positive smears collected by this method have a very high percentage of abnormal cells.

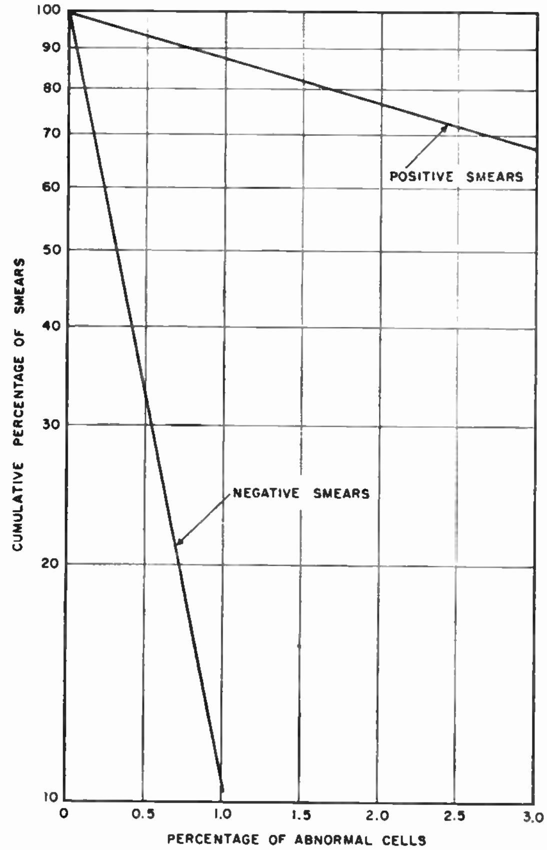


Fig. 2—The frequency of occurrence of abnormal cells on positive and negative premenopause smears. All cells with nuclei exceeding 11-micron diameter and 40 per cent absorption are defined as abnormal.

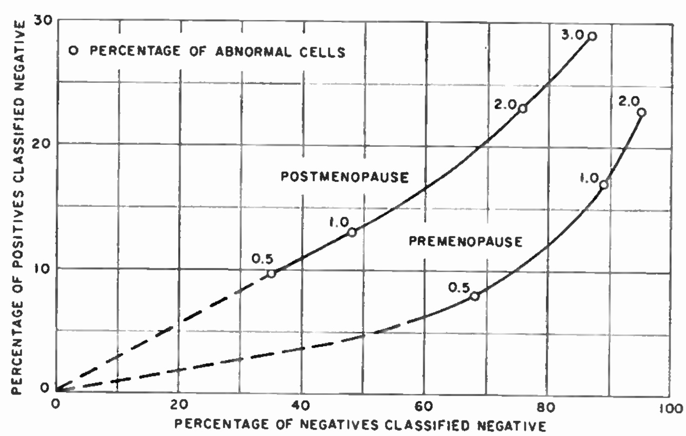
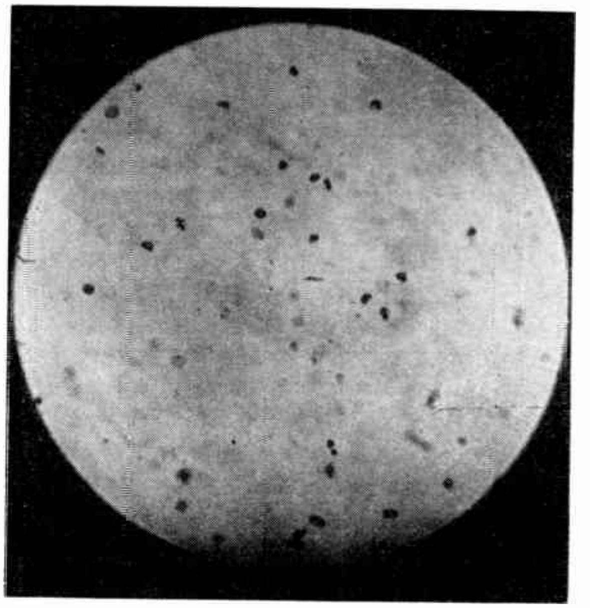


Fig. 3—Screening performance curves for premenopause and postmenopause smears. The curves are derived from cell data measured by manual methods. Screening-criterion values, expressed as a percentage of abnormal cells, are shown as points on the curves.

(a)



(b)

Fig. 4—Typical microphotographs of smear prepared by (a) the conventional Papanicolaou technique for technician screeners and (b) a modified technique for the Cytoanalyzer.

Although cervical scrapings would be somewhat more difficult to use in a mass-screening program, it is probable that a collection method patterned after cervical scraping can be eventually adapted to mass screening. On the basis of this analysis, it was concluded that a prescreening instrument which would satisfy the requirements of a mass-screening program was theoretically feasible.

Smear Modification

The conventional Papanicolaou smear, although quite satisfactory for human screening, is not suitable for use in an automatic instrument. The conventional smear contains appreciable amounts of dense cytoplasmic material that obscures the cell nuclei. In addition, the cells tend to remain in clumps, causing overlapping of the nuclei. Cytoplasm, cellular debris, and artifacts also frequently take forms similar to cell nuclei. The human screener has little difficulty in recognizing cell nuclei from other particles, but the development of an equal capability in an automatic instrument would be extremely difficult.

Accordingly, Pruitt and his co-workers[5] at the National Cancer Institute developed a special smear for the Cytoanalyzer. This smear is prepared by first sieving the suspended collection specimen to remove large cell clumps. The sieved material is sprayed in a dilute solution onto the microscope slide. The smear is then stained by the Feulgen method, which stains only the nucleus.

[5] J. C. Pruitt, S. C. Ingraham II, R. F. Kaiser, and A. W. Hilberg, "Preparation of vaginal-cervical material for automatic scanning—preliminary results with the Cytoanalyzer," *Trans. Sixth Annual Meeting, Inter-Society Cytology Council;* 1958 (in press).

Fig. 4 shows microphotographs of typical fields from a conventional and a modified smear. The major differences between the smears are the more uniform cell dispersion and the lack of cytoplasm staining on the modified smear. Since screening is performed on the basis of nuclear measurement alone, the elimination of the cytoplasmic stain has no adverse affect on the screening capability of the instrument.

Cytoanalyzer Design

The functions performed by the Cytoanalyzer are analogous to the manual procedures used in obtaining quantitative data on the smear. Fig. 5 is a functional block diagram of the system. A microscanner systematically converts a magnified optical image of the smear into a video signal. Each particle scanned produces a series of pulses in the video signal. A cell-measuring unit receives the video signal and from it measures the size and absorption of those particles on the smear exceeding a minimum absorption threshold. In an ideal smear, all such particles would be cell nuclei; in practice, the instrument will also measure a few non-nuclear particles on each smear. An association unit determines which of the pulses in the video signal are associated with the same particle. The output of the cell-measuring unit is sent to a cell sorter, which classifies the measured particles into normal and abnormal categories. Selection circuits apply rules based on cell characteristics. This function is a rudimentary type of cell recognition. Those particles that satisfy the rules are gated out to the totalizers, which count the particles falling into each category. After a representative sample of the cells on the smear has been measured, scanning is

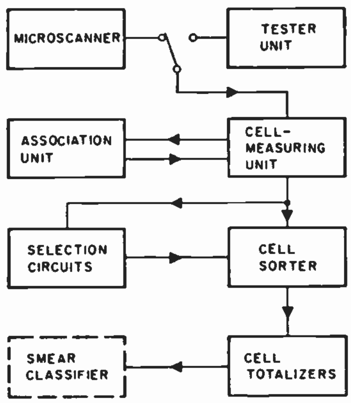


Fig. 5—Functional block diagram of the Cytoanalyzer.

stopped. If the percentage of abnormal counts is less than the screening criterion, the smear is called normal. A tester unit is provided which can be connected in place of the scanner to check if the instrument is functioning properly.

The primary purpose of the present model of the Cytoanalyzer is to demonstrate that an instrument can be used to sort smears accurately. Consequently, only the minimum instrumentation necessary to accomplish this aim has been constructed. Automatic functions required in a clinical instrument which are now performed by the operator include inserting and removing the slide, recording the data stored in the totalizers, and separating the normal and abnormal smears. The clinical instrument would also require a faster sorting rate. The design has been further simplified by permitting the instrument to measure and sort only one particle at a time. Consequently, when several particles fall on the same scan line, all but one of the particles will be ignored. Ignoring particles increases the screening time but does not significantly affect the screening performance because a representative sample of the cells on the smear will still be obtained.

Microscanner

The Cytoanalyzer microscanner[6] is a mechanical scanner using a Nipkow disk to scan a microscopically magnified image of the smear. The diameter of the scanning holes in the disk, when referred to the object plane, is 2 microns. The velocity of the holes is 0.25 micron per microsecond. A mechanical stage moves the smear at a uniform speed to produce an extended raster 0.01 centimeter wide by 5 centimeters long. The area scanning rate is 0.25 square centimeter per minute. The scanner operates until 10,000 normal cells have been counted; thus the total area scanned depends upon the amount of material on the slide. Obtaining 10,000 counts usually requires scanning about 2 square centimeters.

The smear is kept in focus by a mechanical system that uses a spring-loaded stage to press the smear against two feet secured to the microscope objective. The feet are adjusted to place the top surface of the microscope slide in the focal plane of the microscope objective. This focusing method has been found capable of keeping 95 per cent of the cells within ±2 microns of the exact focal plane. A 2-micron focusing error causes negligible error in the cell measurement.

Cell-Measuring Unit

The cell-measuring unit[7] measures the chord length, area and absorption of the nucleus, and the absorption of the cytoplasm surrounding the nucleus of each cell selected for measurement. Scanning the entire nucleus produces a series of video pulses—a pulse each time a scanning hole traverses a chord of the nucleus. The duration of each pulse is proportional to the length of the nucleus chord, and the amplitude of the pulse is proportional to the absorption. Analog circuits are used to measure and store each of these quantities. The accuracy of the chord measurement is ±0.5 micron; of the absorption measurement, ±5 per cent.

The area of the nucleus is used for sorting in preference to diameter because the shape in the nucleus is frequently quite elliptical. Area is measured by summing all the chord measurements associated with the same nucleus. The average error in the area measurement, when compared with the optically measured area, is about 10 per cent for nuclei with a mean diameter of 11 microns. This accuracy is equal to the accuracy used in obtaining the manual photometer data and should give performance results comparable with the predicted performance. It is not yet known how much the screening performance could be improved by increasing the measurement accuracy. If it proves worthwhile, the accuracy can be increased by using higher resolution in the scanner. Higher resolution, however, requires considerable reduction in the scanning rate.

The cytoplasm absorption is measured by sampling the video level at a point 3 microns before and after each nucleus chord. This measurement is used in the selection circuits to reject chords that are not caused by true cells. Because the cytoplasm is not stained, practically no absorption is measured. Thus, the presence of absorbing material adjacent to the chord pulse shows a condition that is not characteristic of the true cell.

Association Unit

The association unit indicates, to other parts of the computer, which chords are associated with a single nucleus. The unit uses a magnetic-core shift register to store the chords of one scan line. Chords of the incoming scan line are then compared with chords in the stored

[6] H. S. Sawyer and R. C. Bostrom, "A new Nipkow-disk scanner for accurate cytological measurement," 1958 IRE National Convention Record, pt. 9, pp. 37–42.

[7] W. E. Tolles, R. C. Bostrom, and H. S. Sawyer, "The application of automatic, high-speed measurement techniques to cytology," 1956 IRE Convention Record, pt. 9, pp. 17–23.

line. Chords are considered to be associated with the same particle if there is time coincidence between the stored and incoming chords.

The shift register was chosen for scan-line storage because 1) it is simple to program, 2) the storage time can be easily adjusted to the length of the active scan line, and 3) the shift register can be readily synchronized with the scanner. Chords must be quantized to enter the shift register. Quantizing may produce an error in the stored chord length of as much as 1.3 microns (the length of one bit in the shift register). The quantizing error does not affect the ability of the instrument to recognize associated chords, but it does limit the ability of the instrument to recognize shapes.

Selection Circuits

To distinguish between cell nuclei and other particles on the smear, selection circuits are used to establish rules which particles must satisfy in order to be accepted as nuclei. As soon as a particle fails to satisfy any one of the rules, measurement of the particle is discontinued and all the circuits are cleared for the next measurement. One set of rules sets limits on the size of the particle by limiting the chord length and the number of chords per particle. A second set of rules requires that the nucleus and cytoplasm absorption fall within specified limits. A third rule requires the shape of the particle to be such that the difference between adjacent chord terminations be less than 3 microns. A fourth set of rules rejects all particles lying on the edge of the scanning raster.

The selection circuits eliminate a substantial number of particles which would otherwise cause errors. However, on a negative smear having 10,000 normal counts, there will still be from 10 to 50 erroneous counts in the abnormal zone which are caused by particles with characteristics very similar to those of a cell nuclei.

Cell Sorter and Totalizers

The cell sorter classifies accepted particles into various categories according to the measured values of absorption and area. Fig. 6 is a zone diagram which shows the categories used in a preliminary test of the instrument. For ease of interpretation, the size boundaries are expressed in terms of mean diameter. Five zones, designated Z1 through Z5, are shown. White blood cells are classified in Zone 1, normal cells in Zone 2, and abnormal cells in Zones 3, 4, and 5. The zone configuration shown is still tentative and can be easily changed. When more operational data has been gathered a configuration will be chosen to optimize the screening efficiency of the instrument. A totalizer, or counter, is connected to each zone. When the normal-zone count reaches 10,000, the instrument stops and the operator records the counts stored in each zone.

Tester Unit

The tester unit consists of several pulse and sweep generators synchronized to produce a repeating 12-line video raster. Appropriate pulses can be inserted into this raster to simulate a rectangular particle of any required absorption and area. Additional pulses can be inserted to vary the shape and cytoplasmic absorption of the test particle.

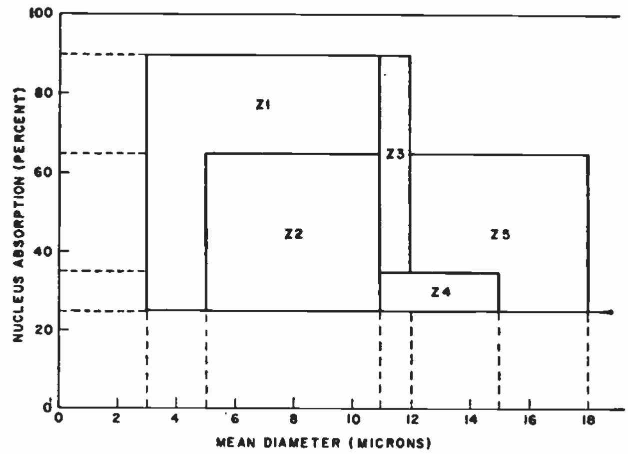


Fig. 6—Zone diagram of cell classification categories. Particles entered into Z1 are called white blood cells, those into Z2 are called normal cells, and those into Z3, Z4, and Z5 are called abnormal cells.

Prescreening Performance

A preliminary test on the prescreening performance of the Cytoanalyzer has been completed by using modified smears.[5] One thousand smears were collected from a group of normal subjects; about one-third of the subjects were postmenopause. All of these smears were negative for cancer. Obtaining positive smears prior to the application of any therapy procedures to the patients proved difficult. During the course of the test only 20 positive smears were obtained—7 from premenopause patients and 13 from postmenopause patients. The classification applied by the cytologist after a visual examination of a duplicate smear prepared by the conventional Papanicolaou method was used as the classification of the modified smear. Each smear was screened three times by the Cytoanalyzer, and the mean value of the abnormal-cell percentage was used in computing screening performance.

The screening-performance curves for the Cytoanalyzer data were constructed in the same manner as that used for the manual-photometer data. The results for the premenopause and postmenopause smears are shown in Fig. 7. Because of the small number of positive smears, a high level of statistical significance cannot be placed on the results. However, it is interesting to note that about 65 per cent of the premenopause smears and 35 per cent of the postmenopause smears were correctly classified negative when the falsely classified positives are limited to 5 per cent. This is roughly equivalent to the performance predicted by the manual-photometer data.

The nonrepeatability of the zone-count data obtained when screening a smear was a matter of concern in

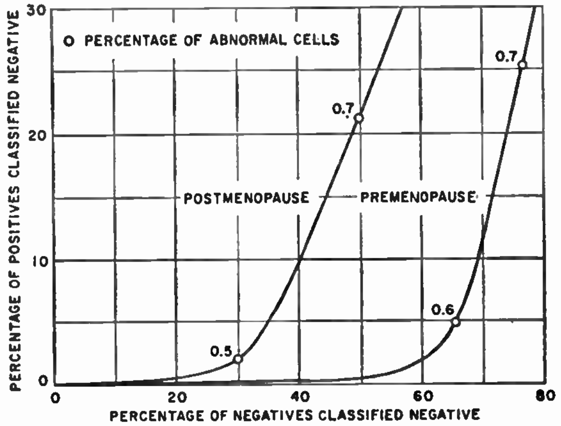


Fig. 7—Screening performance obtained in a preliminary test of the Cytoanalyzer. Screening-criterion values, expressed as a percentage of abnormal cells, are shown as points on the curves.

evaluating the screening performance. The main factors contributing to the variation in the zone counts were 1) the random selection of only one nucleus for measurement when two or more nuclei were on the same scan line, and 2) the noise in the video signal. When the same area of a smear was scanned several times, the standard deviation in the normal-zone count was about 2 per cent when the mean normal-zone count was 10,000. On a smear with an abnormal-zone count of 60 (a value approximately equal to the screening criterion), the standard deviation was about 10 per cent. The variations in the zone counts will have only a small effect on the over-all screening performance when a large number of smears are screened.

However, when only a small number of smears are used in an evaluation test, variation in the abnormal-zone count can cause considerable error in computing the screening performance.

The percentage of abnormal entries in the Cytoanalyzer data was often higher than expected. Cell-by-cell analysis of some of the smears showed that the selection circuits were allowing a few erroneous counts. About 80 per cent of the erroneous counts were caused by overlapped nuclei. Most of the remaining errors were caused by debris and artifacts superimposed on a cell nucleus. It is probably impossible to eliminate all of the erroneous counts because there are some configurations that are almost identical with true nuclei, but methods are being investigated to reduce the errors further.

Another difficulty that degrades the screening performance of the instrument is the erroneous counting of white blood cells as normal epithelial cells. In the region of small nucleus diameters (4 to 6 microns), both types of cells have the same size and absorption. The zone boundaries have been adjusted so that most of the cells are correctly classified, but there are always some cells that fall into the incorrect zones. A few positive smears have tremendous concentrations of white blood cells which may outnumber the normal epithelial cells by a hundredfold. If only a few per cent of the white blood cells are erroneously counted in the normal-cell zone, the apparent percentage of abnormal cells will be greatly decreased. As a result, positive smears are occasionally falsely classified negative. Methods of eliminating the white blood cells in the preparation of the smear by the use of enzymes are being investigated. If the procedure is not successful, it is believed that the white blood cells can be recognized by their small cytoplasm-to-nucleus-diameter ratio.

Significance of Results

The number of positive cases in the preliminary test were too small to provide a statistically significant evaluation. However, provided the results of this test continue to be substantiated by further tests, the performance does show that instrument prescreening is feasible. With 65 per cent of the premenopause negative smears correctly screened, the effectiveness of the cytological technician is greatly increased. On the postmenopause smears, with only 35 per cent of the smears correctly screened, the gain to the cytologist is considerably less. This suggests that, on the postmenopause subjects, smears collected by the cervical scraping method should be used to take advantage of the greatly improved performance predicted by the manual-photometer data for that type of collection.

Economically, instrument prescreening is undoubtedly advantageous. It appears entirely feasible to decrease the screening time to one minute per smear by increasing the scanning rate, counting all of the scanned particles, and reducing the average number of cells required to classify a smear. At this rate, the over-all screening cost for the premenopause smears would be less than one-half the screening cost when using only cytological technicians. The cost of prescreening the postmenopause smears is also less than technician screening, even though two-thirds of the smears must be subsequently re-examined.

On the basis of the promising results of the preliminary test a final evaluation is scheduled which will use at least 10,000 smears from a mass-screening clinic. If the test proves successful, development will begin on a high-speed completely automatic Cytoanalyzer for clinical use. This instrument should make mass screening possible on a truly large scale, and could be the means of finally bringing cervical cancer under control.

A Magnetic Flowmeter for Recording Cardiac Output*

HAMPTON W. SHIRER†, ASSOCIATE MEMBER, IRE, RICHARD B. SHACKELFORD†, ASSOCIATE MEMBER, IRE, AND KENNETH E. JOCHIM†

Summary—**The cardiac output flow pulse (less coronary flow) can be recorded with a magnetic flowmeter applied to the unopened ascending aorta, provided: 1) the extremely large EKG potentials in this region are rejected, 2) the flowmeter is phase-sensitive, and 3) the over-all instrument response is uniform from zero to 100 cps. These requirements are fulfilled by the instrument described through the use of the square-wave method recently introduced by A. B. Denison, and by using a high switching frequency (480 cps), an input high-pass filter, and a double-balanced demodulator. Two output channels of suitable response provide for simultaneous recording of instantaneous and mean flow. Each pickup sleeve is calibrated *in vitro* in terms of microvolts per flow rate. An electrical series calibrator provides the operational calibration. Past methods are reviewed and circuit details and design considerations are discussed.**

Introduction

THE ideal method for measuring and recording instantaneous rate of blood flow, such as cardiac output, is not yet at hand. Such a method, among other things, would allow the determination to be made on the unoperated, unanesthetized, intact animal. However, for use on the anesthetized, experimental animal in which the ascending aorta or any other vessel may be exposed, the magnetic flowmeter approaches the ideal for the following reasons: 1) the vessel need not be opened nor any foreign object inserted in the flowing stream; 2) calibration is linear and independent of the character of flow, be it turbulent or streamline; 3) calibration is independent of temperature, pressure, density, and viscosity; 4) response is essentially instantaneous; and 5) the instrument can be made to show direction of flow. It is the purpose of this paper to review the developments in magnetic flowmeter design and to describe an instrument employing some of the newer techniques.

The principle of operation of the induction or magnetic flowmeter is shown in Fig. 1, and is based on Faraday's law of electromagnetic induction given by:

$$E = BL\bar{V}10^{-8} \text{ volts}, \quad (1)$$

where

E is the potential in volts developed at the ends of a conductor moving through a uniform field,
B is the magnetic flux density in gauss,
L is the length in centimeters of the conductor at right angles to the field flux lines,
$\bar{V}$ is the velocity in centimeters per second of the conductor moving at right angles to the field.

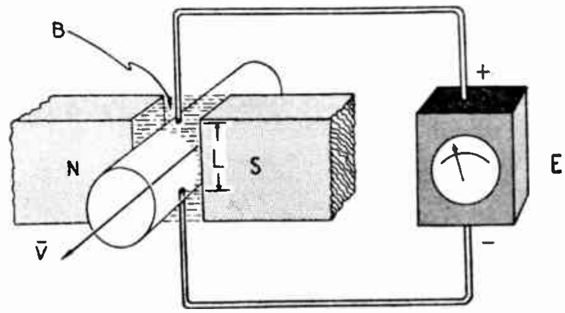


Fig. 1—Principle of the magnetic flowmeter. B, the magnetic flux density; L, the diameter of the moving stream; $\bar{V}$, the mean cross-sectional flow velocity; and E, the flow-induced potential.

In the case where the moving conductor is a liquid such as mercury, salt water, or blood flowing in a pipe, L becomes the internal diameter of the pipe at right angles to the flux lines and $\bar{V}$ the mean cross-sectional velocity of flow. If the flux density and the pipe diameter are held constant, the potential developed at the surface of the flowing stream is then a simple linear function of the volume rate of flow. In order to measure this potential in nonconducting pipes, electrodes must be placed on the inner surface of the pipe wall. For flow through pipes with conducting walls, such as blood vessels, the flow-induced potential can be picked off by simply placing electrodes on the outer surface.

Review of Developments

The first attempt to measure liquid flow by electromagnetic induction was made by Faraday in 1832.[1] He attempted, unsuccessfully, to determine the potential developed between two electrodes placed in the River Thames due to its flow through the earth's magnetic field. Lack of success was probably caused by spurious potentials resulting from electrode polarization. In 1920, however, Young and others clearly recorded potentials due to the ebb and flow of tides through the use of both towed and fixed electrodes.[2] Ten years later, the distribution of velocities in a copper sulphate solution flowing through a magnetic field was studied with exploring electrodes by Williams.[3] The first suggestion

* Original manuscript received by the IRE, June 8, 1959. Supported by a grant from the Natl. Heart Inst. USPHS.
† Dept. of Physiology, University of Kansas, Lawrence, Kans.

[1] T. Martin, "Faraday's Diary," G. Bell and Son, London, Eng., vol. 1, pp. 409–411; 1932.
[2] F. B. Young, H. Gerrard, and W. Jevons, "On electrical disturbances due to tides and waves," *Phil. Mag.*, vol. 40, pp. 149–159; July, 1920.
[3] E. J. Williams, "The induction of electromotive forces in a moving liquid by a magnetic field, and its application to an investigation of the flow of liquids," *Proc. Phys. Soc.* (*London*), vol. 42, pp. 466–478; August, 1930.

that the flow of blood might be determined by the magnetic method was made by Fabre in 1932.[4]

The electromagnetic flowmeter became a practical, quantitative instrument for the measurement of blood flow through the independent developments of Kolin[5–7] and Wetterer[8,9] in 1936 and 1937. Their findings and theoretical analyses were later confirmed by Einhorn[10] and Thürlemann.[11] Since its introduction, this instrument has received increasing attention by both physiologists and industrial users.[12–16]

The development of the magnetic flowmeter by various workers since its introduction has led to various attempts to eliminate background potentials that bear no relationship to rate of flow. Unfortunately, such background potentials are usually quite variable and many times the amplitude of the flow-induced potential. All of the early flowmeters were of the dc type,[5–9,11,17–21] that is, the magnetic field was induced by either a permanent magnet or an electromagnet excited by direct current. The flow-induced potential, being dc and only in the range of a few tens of microvolts at best, requires the use of nonpolarizable electrodes and a high degree of amplification. Even the best nonpolarizable electrodes develop variable potentials of the same order of magnitude as the flow signal. The problem was further complicated by the necessity of using high-gain dc amplifiers with their notoriously high drift rate. To some extent, this latter complication has been reduced by the use of mechanical choppers to convert the dc signal to ac, which can then be amplified by drift-free, capacitance-coupled amplifiers,[6,7,18] or through the use of direct-coupled amplifiers of improved design.[19,20,21] The electrodes still remain a problem, however, due to their variable potential and to their bulk. As the method was extended to a greater variety of blood-flow measurements, it became increasingly important to reduce the size of the electrode assembly.

Soon after Kolin's original publication, he suggested that, by use of an alternating magnetic field, the flow signal would also be alternating and thus eliminate the need for bulky nonpolarizable electrodes and direct-current amplification.[7] Practical and theoretical considerations of such carrier-operated flowmeters were later described by Kolin,[22–25] and in considerable detail by Einhorn,[10] James,[14,16] and Kolin.[26] The use of the alternating magnetic field, however, introduced a new problem. The potential induced in the pickup electrode circuit consists of two components: that due to flow and that due to the alternating field (the so-called "transformer" component); it is given by Kolin[26] as:

$$V = [(\mu H_0 \sin \omega t)dv - (A_0 \cos \omega t)]10^{-8} \text{ volts},$$

where

V is the instantaneous voltage induced in the pickup electrode circuit,
$\mu H_0 \sin \omega t$ is the instantaneous magnetic flux density,
d is the diameter of the flowing stream in cm,
v is the mean cross-sectional velocity of flow in cm/sec,
$(\mu H_0 \sin \omega t)dv$ is the instantaneous flow component,
$A_0 \cos \omega t$ is the instantaneous transformer component.

From this it can be seen that the flow component is in phase with the magnetic field, while the transformer component is in quadrature. James[14] further stressed, in his expressions shown below, that for a given set of conditions the magnitude of the flow component is proportional to the strength of the magnetic field, but the magnitude of the transformer component is, in addition, directly proportional to the rate of change or the frequency of the magnetic field:

[4] P. Fabre, "Utilisation des forces électromotrices d'induction pour l'enregistrement des variations de vitesse des loquides conducteurs: un nouvel hémodromographie sans palette dans le sang," *C. R. Acad. Sci.*, vol. 194, pp. 1097–1098; March 21, 1932.

[5] A. Kolin, "An electromagnetic flowmeter. Principles of the method and its application to blood flow measurements," *Proc. Soc. Exper. Biol. Med.*, vol. 35, pp. 53–56; October, 1936.

[6] A. Kolin, "An electromagnetic recording flowmeter," *Amer. J. Physiol.*, vol. 119, pp. 355–356; June, 1937.

[7] L. N. Katz and A. Kolin, "The flow of blood in the carotid artery of the dog under various circumstances as determined with the electromagnetic flowmeter," *Amer. J. Physiol.*, vol. 122, pp. 788–804; June, 1938.

[8] E. Wetterer, "Eine neue Methode zur Registrierung der Blutströmungsgeschwindigkeit am uneröffneten Gefäss," *Z. Biol.*, vol. 98, pp. 26–36; January, 1937.

[9] E. Wetterer, "Der Induktionstachograph," *Z. Biol.*, vol. 99, pp. 158–162; March, 1938.

[10] H. D. Einhorn, "Electromagnetic induction in water," *Trans. Roy. Soc. S. Africa*, vol. 28, pp. 143–160; June, 1941.

[11] B. Thürlemann, "Methode zur elektrischen Geschwindigkeitsmessung von Flüssigkeiten," *Helv. Phys. Acta*, vol. 14, pp. 373–419; October, 1941.

[12] A. Kolin, "An alternating field induction flowmeter of high sensitivity," *Rev. Sci. Instr.*, vol. 16, pp. 109–116; May, 1945.

[13] G. Remenieras, "Sur la possibilité de transformer directement en énergid électroque une partie de l'énergie d'une viene fluide," *Soc. Hydrotech. France*; November, 1947.

[14] W. G. James, "An induction flowmeter design suitable for radioactive liquids," *Rev. Sci. Instr.*, vol. 22, pp. 989–1002; December, 1951.

[15] J. S. Arnold, "An electromagnetic flowmeter for transient flow studies," *Rev. Sci. Instr.*, vol. 22, pp. 43–47; January, 1951.

[16] W. G. James, "An ac induction flowmeter," *Instruments*, vol. 25, pp. 473–478; April, 1952.

[17] K. E. Jochim, "Some improvements on the electromagnetic flowmeter," *Amer. J. Physiol.*, vol. 126, pp. 547–548; July, 1939.

[18] L. N. Katz and K. E. Jochim, "Electromagnetic flowmeter," in "Medical Physics," O. Glasser, Ed., Yearbook Publishers, Chicago, Ill., pp. 377–379; 1944.

[19] K. E. Jochim, "Electromagnetic flowmeter," in "Methods in Medical Research," V. R. Potter, Ed., Yearbook Publishers, Chicago, Ill., vol. 1, pp. 108–115; 1948.

[20] K. E. Jochim, "Circulatory system: methods, electromagnetic flowmeter," in "Medical Physics," O. Glasser, Ed., Yearbook Publishers, Chicago, Ill., vol. 2, pp. 225–228; 1950.

[21] T. G. Richards and T. D. Williams, "Velocity changes in the carotid and femoral arteries of dogs during the cardiac cycle," *J. Physiol.*, vol. 120, pp. 257–266; May, 1953.

[22] A. Kolin, "An a.c. induction flowmeter for measurement of blood flow in intact blood vessels," *Proc. Soc. Exper. Biol. Med.*, vol. 46, pp. 235–239; February, 1941.

[23] A. Kolin, J. L. Weissberg, and L. Gerber, "Electromagnetic measurement of blood flow and sphygmomanometry in the intact animal," *Proc. Soc. Exper. Biol. Med.*, vol. 47, pp. 324–329; June, 1941.

[24] A. Kolin, "Electromagnetic velometry. I. A method for the determination of fluid velocity distribution in space and time," *J. Appl. Phys.*, vol. 15, pp. 150–164; February, 1944.

[25] A. Kolin, "An alternating field induction flowmeter of high sensitivity," *Rev. Sci. Instr.*, vol. 16, pp. 109–116; May, 1945.

[26] A. Kolin, "Improved apparatus and technique for electromagnetic determination of blood flow," *Rev. Sci. Instr.*, vol. 23, pp. 235–242; May, 1952.

$$e_s = K'B \sin \omega t$$
$$e_b = 2\pi KBf \cos \omega t,$$

where

e_s is the flow component,
K' is a proportionality constant,
B is the amplitude of the magnetic flux density,
e_b is the transformer component,
K is another proportionality constant,
f is the frequency of the applied field.

It can be seen that A_0 in Kolin's equation above is equivalent to $2\pi KBf/10^{-8}$ in James' terminology. Thus, by using an alternating magnetic field and capacitance-coupled amplifiers, the wandering, unpredictable electrode potentials and drift of the dc method are replaced by the transformer component. This has the advantage, at least in theory, of being perfectly stable and therefore capable of cancellation.

The several methods of elimination of the transformer component prior to amplification and demodulation are shown schematically in Fig. 2. They are based either

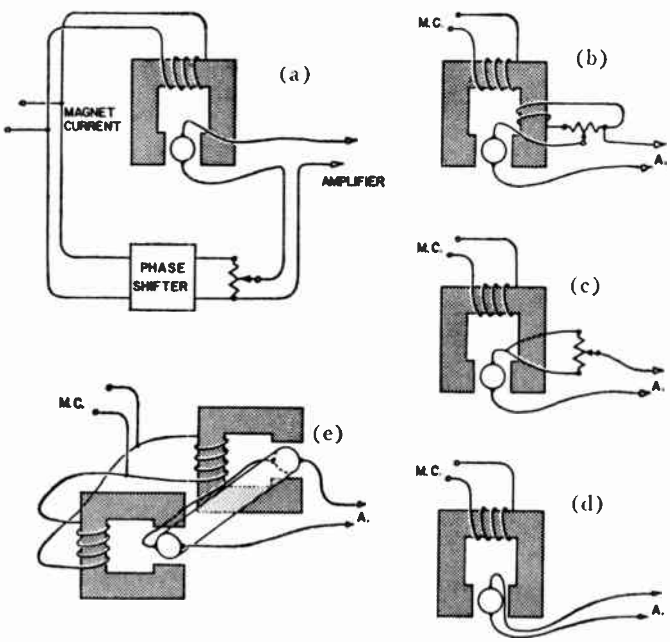


Fig. 2—Several methods used to eliminate the transformer component when an alternating magnetic field is used. (a), bucking voltage derived from the magnet-current source; (b) bucking voltage derived from the magnetic field; (c), split lead method; (d), lead orientation to enclose zero flux; and (e), two fields of opposite phase and two pickup electrode pairs connected, opposing for transformer component but aiding for flow signal.

on the addition to the input signal of a potential of exactly the same amplitude but opposite phase to the transformer component, or on the use of noninductive lead orientation. When this is done, the potential appearing at the amplifier input is due only to flow and is zero when the flow is zero. While careful orientation of the leads from the electrodes to form a noninductive loop, shown in Fig. 2(d), was the original method suggested by Kolin,[7] for some time he favored the addition of the proper cancellation voltage derived from the magnet current source, as shown in Fig. 2(a).[22–26] Indeed, he designed a special variable-phase transformer for just this purpose.[27] The difficulty encountered in this method is due to waveform distortion brought about by nonlinearity in the magnet-core material, making complete cancellation impossible. In order to overcome this problem, many have preferred to derive the cancellation voltage from the magnet field itself by a pickup coil as shown in Fig. 2(b).[7,28–30] The voltage induced in the auxiliary coil is then adjusted in amplitude by a voltage divider and added in series with the pickup electrodes to effect cancellation. By careful adjustment of the orientation of the pickup leads, they can be made noninductive with respect to the magnetic field. The method for doing this electrically, devised by Einhorn[10] and later by Denison,[31–34] is shown in Fig. 2(c). Here one electrode lead is split and carried away from the magnet on either side of the core and rejoined through a potentiometer. By adjusting the arm of the potentiometer, the lead that would form a noninductive loop, that is, enclose zero net flux, can be simulated. Fig. 2(d) shows the same scheme carried out by careful orientation of the inner lead, so that the loop formed by the two leads and the pipe diameter lies in a plane parallel to the magnetic flux lines and therefore encloses zero net flux.[14,16,35] A rather involved method of cancellation was used by Arnold,[15] shown in Fig. 2(e), in which two magnetic fields of opposite phase and two pairs of electrodes were placed along the axis of flow. The electrodes were then connected series-aiding for the flow-induced signal, which results in their being series-opposing for the transformer component.

Until recently, the use of phase discrimination to reduce or eliminate the transformer component has received little consideration. Indeed, the use of phase-

[27] A. Kolin, "A variable phase transformer and its use as an a.c interference eliminator," *Rev. Sci. Instr.*, vol. 12, p. 555; November 1941.

[28] J. W. Clark and J. E. Randall, "An electromagnetic blood flow meter," *Rev. Sci. Instr.*, vol. 20, pp. 951–954; December, 1949.

[29] A. W. Richardson, J. E. Randall, and H. M. Hines, "A newly developed electromagnetic flow meter," *J. Lab. Clin. Med.*, vol. 34, pp. 1706–1713; December, 1949.

[30] A. W. Richardson, A. B. Denison, and H. D. Green, "A newly modified electromagnetic blood flowmeter capable of high fidelity flow registration," *Circulation*, vol. 5, pp. 430–436; March, 1952.

[31] A. B. Denison, M. P. Spencer, and H. D. Green, "A square-wave electromagnetic flowmeter for application to intact blood vessels," *Circ. Res.*, vol. 3, pp. 39–46; January, 1955.

[32] A. B. Denison and M. P. Spencer, "Factors involved in intact vessel electromagnetic flow recording," *Federation Proc.*, vol. 15, p. 46; March, 1956.

[33] A. B. Denison and M. P. Spencer, "Square-wave electromagnetic flowmeter design," *Rev. Sci. Instr.*, vol. 27, pp. 707–711; September, 1956.

[34] N. P. Spencer, A. B. Denison, and C. A. Barefoot, "Continuous measurement of cardiac output in conscious dogs by means of an indwelling magnet and the square-wave magnetic flowmeter," *Federation Proc.*, vol. 17, p. 154; March, 1958.

[35] E. Abbott, N. Assali, G. Herrold, and A. Kolin, "The Present State of Development of the Electromagnetic Blood Flow Meter," presented at Biophys. Soc. Meeting, Cambridge, Mass.; February, 1958.

sensitive demodulation techniques to preserve the phase information as to direction of flow has been largely ignored. Most instruments employed simple half- or full-wave rectifiers which cannot discriminate between forward and reverse flow (a 180° phase difference in the carrier), or between the flow signal and the transformer component (a 90° phase difference). Kolin used a form of phase-discriminating demodulation in his method of recording from the screen of an oscilloscope. The sweep of the oscilloscope was synchronized with the magnet-current source (usually the 60-cycle line), allowing the presentation of one stationary cycle of flow-modulated carrier. A slotted mask was placed over the screen and its horizontal position adjusted so that only a short segment of the sine-wave peak could be seen. At zero flow, this segment fell at the middle of the screen, for forward flow it was deflected upward, and for reverse flow it was deflected downward from the center and was so recorded on moving film. The transformer component, being 90° out of phase with the flow component, caused no deflection of this segment.[22–26,36] James suggested the use of phase discrimination to eliminate the transformer component and to preserve the flow direction information, but did not employ it in the instrument he described.[11,16] Some simply recorded the undemodulated carrier and observed the carrier phase to determine flow direction.[10,15] Others used the transformer component as a bias, so that the flow signal increased the rectified voltage on forward flow and decreased it on reverse flow.[30,37]

This latter method can introduce a high degree of nonlinearity, since the rectified output is the vector sum of the transformer and flow components rather than the algebraic sum, due to the 90° phase difference between the two components.

Practical difficulties in the elimination of the transformer component, either by cancellation in the input circuit or by phase discrimination at the demodulator, led Denison and Spencer to a unique approach to the problem. They pointed out that, although the transformer component could be eliminated quite simply in principle, stray fields from the magnet induced potentials that are not stable as to either amplitude or phase into the surrounding tissues and then into the input circuit. Such instability is brought about mainly by changes in the position of the magnet assembly with respect to the tissue, and is particularly severe when flow determinations are made in the regions of the heart and lungs. They further reasoned that, since it is the changing field that induces these troublesome potentials, they could be eliminated by using a steady field that is rapidly switched from one polarity to the other periodically, *i.e.*, a square rather than a sinusoidally-varying field. The transformer component then becomes a very brief switching transient which can be eliminated simply by gating off the amplifier momentarily during the switching phase. With the switching transient thus removed, the remaining signal is a rectangular wave of amplitude proportional to flow rate and phase according to direction. This method is illustrated in Fig. 3. It can

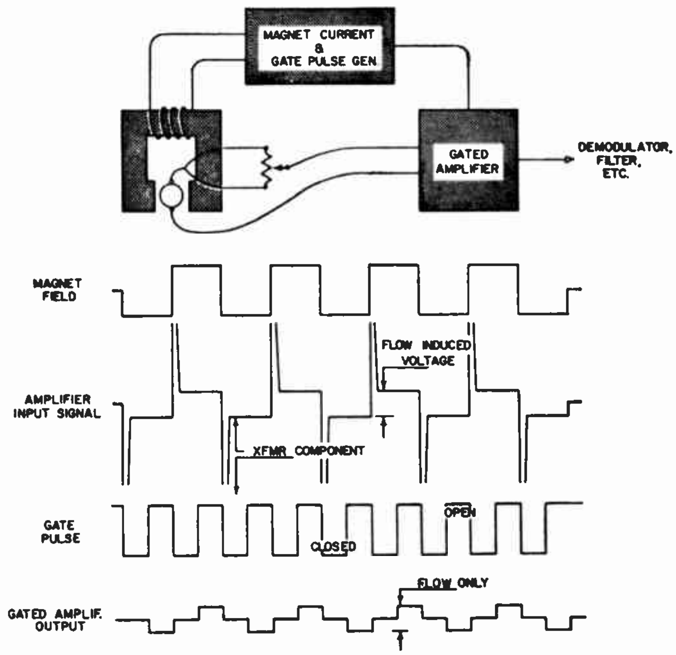


Fig. 3—The Denison and Spencer square-wave method.

be seen that this method is a form of phase discrimination with respect to elimination of the transformer component, but differs from the sine-wave case in that the gating phase does not have to be maintained with absolute precision. In practice, the gate-off phase is made longer than the period of the switching transient. The relative positions of the gate-off phase and switching transient are not critical, providing that switching transient is contained entirely within the gate-off phase. In the sine-wave case, however, the phase of the demodulator carrier must be maintained precisely, and any deviation of the transformer component phase from the demodulator carrier phase will cause an output voltage proportional to the sine function of the deviation.

The frequency of the exciting magnetic field, *i.e.*, the carrier frequency, sets the upper limit on the net frequency response of the flowmeter. In principle, the carrier frequency must be twice the highest frequency component of the flow signal to be recorded, and in practice, a bit higher. It is generally agreed that to record undistorted pressure and flow waveforms in the cardiovascular system, particularly in vessels near the heart, a uniform frequency response and linear phase response from zero to 100 cps is desirable. This would then require the use of carriers of at least 200 cps for

[36] A. Kolin, "A method for adjustment of the zero setting of an electromagnetic flowmeter without interruption of flow," *Rev. Sci. Instr.*, vol. 24, pp. 178–179; February, 1953.

[37] F. Olmsted, "Continuous registration of phasic cardiac output in the active dog: the chronically implanted electromagnetic flowmeter," Abstracts of the 11th Annual Conf. on Electrical Techniques in Medicine and Biology, Minneapolis, Minn., pp. 24–25; November, 1958.

flowmeters capable of following the most rapid changes in flow rate. All of the earlier carrier instruments employed carriers at the power-line frequencies of 50 or 60 cps because of the obvious simplicity in obtaining large magnet currents.[7,10,22–26,28–30] James pointed out the advantage of using a very low carrier frequency to reduce the relative amplitude of the transformer component. His instrument utilized a carrier of 11.7 cps but, of course, had a quite limited frequency response.[14,16] Denison used a 30-cycle switching rate in his original square-wave flow meter.[31] To study quite rapid transients, such as the water-hammer effect in rigid hydraulic systems, Arnold utilized a carrier of 5000 cps. This allowed the observation of flow transients as brief as 1-msec duration.[15] While the use of a carrier frequency greater than 60 cycles for blood flowmeters had been suggested,[29,30] it was not until Denison's 240-cycle square-wave flowmeter[32,33] that the use of fast carriers had been described. Recently a group from Kolin's laboratory described an instrument using a carrier of from 400 to 1000 cps.[35]

In order to utilize the magnetic flowmeter on conscious animals, several workers have recently concentrated on reducing the size of the magnet assembly to allow its surgical implantation about the artery for recording after recovery from the surgery. The first implanted flowmeter method was tried by placing the electrode sleeve only about the vessel and bringing the leads out through a small incision. The magnetic field was provided by a large electromagnet with a gap sufficient to enclose that portion of the animal containing the sleeve.[23] While Kolin later showed some methods to correct for the variable orientation of the vessel and electrodes within the field, this technique has not persisted.[26,36] A very small magnet and electrode assembly implanted about the carotid artery of a conscious active dog, using a 60-cycle carrier and later a 400-cycle carrier, has been demonstrated by Kolin's group.[35,38] Cardiac output recording from conscious dogs has been carried out by Spencer and Denison using their square-wave instrument and a miniature magnet assembly implanted on the ascending aorta.[34] A similar technique using a 400-cycle sine-wave carrier has also been employed by Olmstead.[37]

The magnetic flowmeter has been used to record the flow of blood to several vascular areas, most often in the dog. The two vessels most frequently studied are the femoral[21–23,28–31] and the carotid[5–7,26,35,38] arteries. Wetterer's original instrument was used on the ascending aorta of the rabbit.[8,9] Recording from the ascending aorta of the dog was mentioned above.[34,37] Kolin measured the flow in the descending thoracic aorta of the dog by his sleeve-implantation method.[26] Spencer and Denison also studied descending aorta flow with the square-wave method[39] as well as the flow in renal and mesenteric arteries.[31] This is by no means a complete list of all the regions studied by various workers, but only a summary of those mentioned by people active in the development of the instrument.

Design Considerations for a Cardiac Output Flowmeter

With the exception of the fraction supplying the heart musculature by way of the coronary vessels, cardiac output may be determined by recording the rate of flow through the ascending aorta. In the dog, this vessel is from 8 to 20 mm in diameter and can be dissected free for a length of 25 to 30 mm. This provides enough freedom to apply an electrode sleeve and poles of an exciting magnet about the vessel.

A special problem in applying the magnetic method to flow recording from this vessel is the very large cardiac action potential (EKG) existing at the flow pickup electrodes. This potential is of the order of 5 mv peak-to-peak, while that due to typical rates of flow through a magnetic field of 200 to 500 gauss is only in the range of 50 to 500 μv. Attempts to reduce this source of interference by either guard electrodes or differential amplifiers have not met with success.

The solution to the EKG problem presents itself through the use of a carrier-operated system. The spectrum occupied by the EKG, which is quite rich in harmonics, lies between zero and 100 cps. On the other hand, the signal produced by pulsatile blood flow through an alternating field of, say 480 cps, consists of a suppressed 480-cycle carrier plus the upper and lower sidebands. To obtain an over-all net frequency response for the flow signal from zero to 100 cycles, the net pass band of the amplifier through the demodulator need only be from 380 to 580 cps. All frequencies outside this passband can be rejected by filters in the amplifier or through the use of signal-balanced demodulation followed by a low-pass filter. Thus, by selecting the carrier frequency, spectral regions containing large noise voltages can be avoided.

As the carrier frequency is raised, the problem of elimination of the transformer component from the flow signal becomes increasingly difficult, since its magnitude is proportional to the rate of change of the magnetic field. Of the several methods that have been proposed, the square-wave method introduced by Denison and Spencer offers the greatest promise. The important principle of this method is that the flow signal be sampled during a period when the magnetic field is constant; the waveform of the current used to energize the magnet must have a flat top, but need not have a short rise time. In general, it would seem that the ideal waveform would be a triangular wave with the peaks clipped

[38] G. Herrold, A. Kolin, N. S. Assali, and R. Jersey, "Miniature electromagnetic flowmeter for chronic implantation," Abstracts of the Natl. Biophys. Conf., Columbus, Ohio, p. 34; March, 1957.

[39] M. P. Spencer and A. B. Denison, "The aortic flow pulse as related to differential pressure," *Circ. Res.*, vol. 4, pp. 476–484; July, 1956.

off quite flat for a brief period. For a given frequency, the triangular wave has the minimum rate of change and therefore will induce the minimum transformer component. The flow signal sampling period need only be very brief and positioned somewhere near the center of the flat portion of the magnetic field waveform. If the width of the flat portion is made somewhat greater than the period of signal sample, a tolerance is obtained to some drift in the sample phase with no appearance of transformer component in the output.

A uniform magnetic flux density over the entire cross-sectional area of the flowing stream is necessary in order that the calibration of the flowmeter be independent of the nature of the flow profile; a magnet with long pole faces is thus required. Due to fringing of the field at the edges of the pole faces, the faces must be somewhat longer than the diameter of the largest vessel to be studied. This requires a rather large, bulky magnet assembly when using the conventional horseshoe magnet structure. Since our purposes would be served by making recordings from the anesthetized, open-chest preparation, no attempt was made to reduce the magnet size for implantation.

Circuit Description

Block Diagram

The block diagram of the cardiac output flowmeter is shown in Fig. 4 along with the signal modification diagram. The 480-cycle carrier frequency is locked to a multiple of the power-line frequency by deriving the master synchronizing pulses from the 16th harmonic of the power frequency. The master pulse generator supplies this pulse as a trigger to the magnet-current generator, and, after an appropriate delay, to the amplifier gate-pulse generator and to the demodulator carrier generator. The magnet-current waveform has a very flat top, but no attempt is made to obtain a particularly fast rise, since this would only serve to increase the amplitude of the switching transient. The pickup electrodes are held in contact with the surface of the aorta by a plastic sleeve selected to constrict the vessel slightly in order to maintain a constant diameter and to hold the electrodes in firm electrical contact. The sleeve is held rigidly between the poles of the magnet. The signal from the pickup electrodes passes through a series calibrator to the amplifier input where the switching spike is partially balanced out by the split-lead method. The amplifier raises the signal level from a few tens or hundreds of microvolts up to a few volts where it is then sampled by the gate. The gate is open for only a small fraction of each half cycle and adjusted to occur late in the half cycle when the magnet field is most steady. The square wave character of the gated flow signal is then restored by a pulse stretcher. Demodulation is carried out by a carrier- and signal-balanced ring demodulator, and all higher order products are removed by a low-pass filter. The output dc amplifiers have only a small voltage gain and are used primarily as impedance transformers to provide current outputs of plus and minus 20 ma into low-resistance galvanometers. One output channel is used for recording instantaneous flow and the second for recording mean flow. The mean-flow signal is derived by passing the flow signal through an additional resistance-capacitance filter of 3.3-seconds time constant.

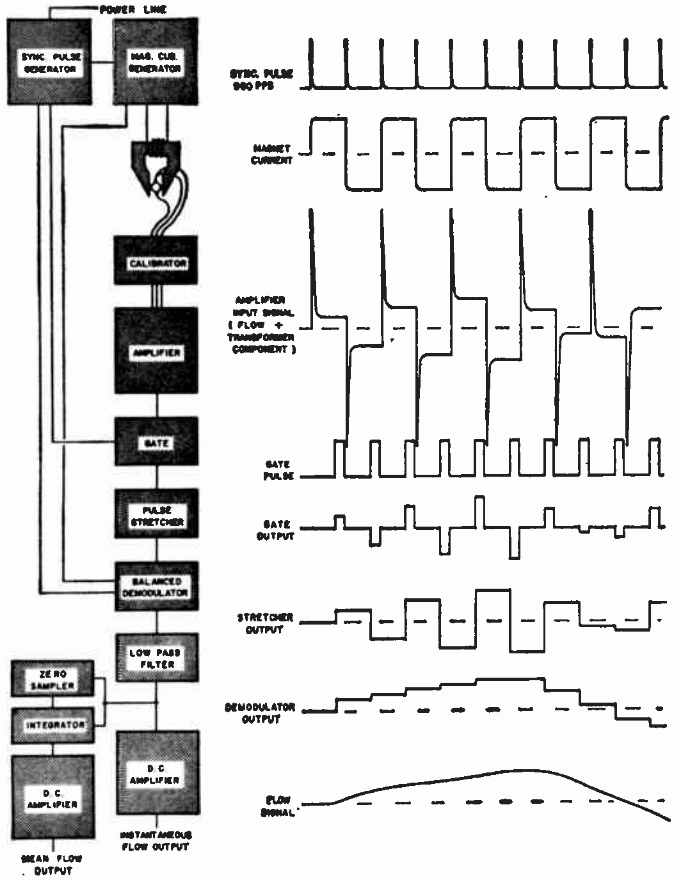


Fig. 4—The cardiac output flowmeter.

Magnet Assembly

The magnet assembly, shown in Fig. 5, consists of two parts: the magnet itself, and a removable sleeve. The sleeve, machined from acrylic plastic, is used to support the silver electrodes against the surface of the aorta and to fix its diameter. Several sleeves are made to cover a range of vessel diameters. On one side of the sleeve there is a removable insert to allow placement of the sleeve about the aorta. The insert is then replaced and the magnet applied by sliding it down from above, thus holding the insert firmly in place. The electrodes terminate in a pair of lugs on the sleeve which in turn slip under knurled nuts on the magnet; these serve both to hold the sleeve in the magnet and to connect the electrodes to the input cable. The lead from one of the lug terminals on the magnet is split and carried to the cable terminals on either side of the magnet core. The other terminal is brought out by a single lead. The input cable consists of three small coaxial cables from the terminals on the side of the magnet to the calibrator. The magnet

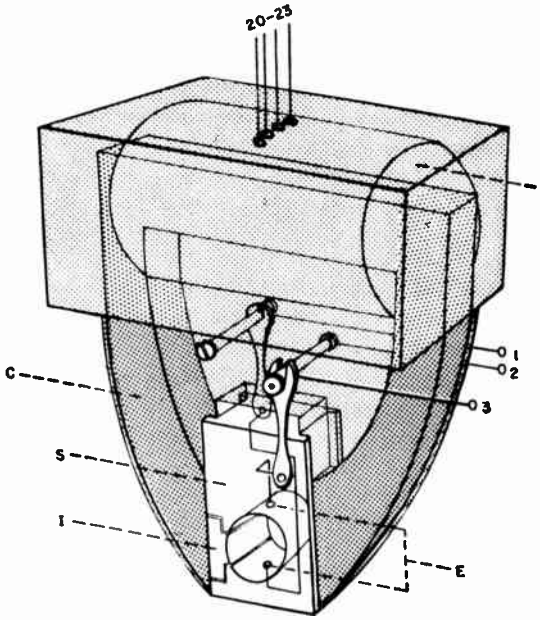


Fig. 5—Magnet and sleeve assembly. *S*, the sleeve with removable insert, *I*, and silver electrodes, *E*; *C*, laminated magnet core; *W*, windings enclosed in an electrostatic shield; 1, terminal of one electrode; 2 and 3, terminals via split lead of the other electrode; 20 to 23, magnet winding terminals.

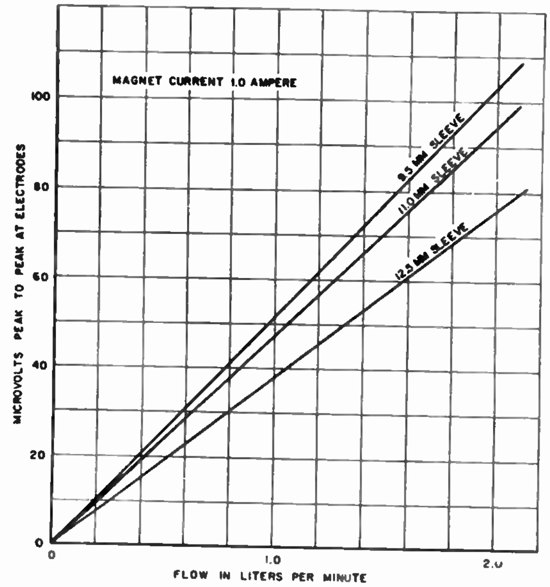


Fig. 6—Sleeve calibration.

core is made up from transformer laminations and has a cross section of 6 by 15 mm. The faces of the poles are 6 by 25 mm and the gap 20 mm wide. The magnet is excited by a pair of windings of 250 turns of number 22 Formvar wire which are electrostatically shielded by copper foil. The magnet is cast in epoxide plastic to provide electrical insulation, prevent corrosion, and to aid in preserving dimensional stability. Each of the sleeves is calibrated in terms of microvolts per flow rate in liters per minute at one ampere magnet current. Sleeve calibration is done by inserting beef arteries in the sleeves and perfusing at a known flow rate. Typical curves, which are quite linear, are shown in Fig. 6.

Input Circuit

The input circuit is shown in Fig. 7. The signal from the magnet assembly is carried through a series calibrator to the amplifier. The calibrator provides a 480-cycle square wave of 20, 50, 100, 200, or 500 μv, peak-to-peak, in the same phase as the flow signal. This calibration voltage is applied across a very low value of resistance (2 to 50 ohms) inserted in the single electrode lead. This method of calibration, long used in electrocardiography, takes into account any loading of the signal source by the input impedance of the amplifier, since the calibration signal appears in series with the source impedance of the flow signal. A double-balanced ring modulator is used to generate the calibration voltage. By operating the modulator at a carrier level of at least ten volts peak, the output amplitude is determined primarily by the 1.34-volt mercury cell and the attenuation produced by the isolation transformer and the resistance dividing network, and only to a small extent

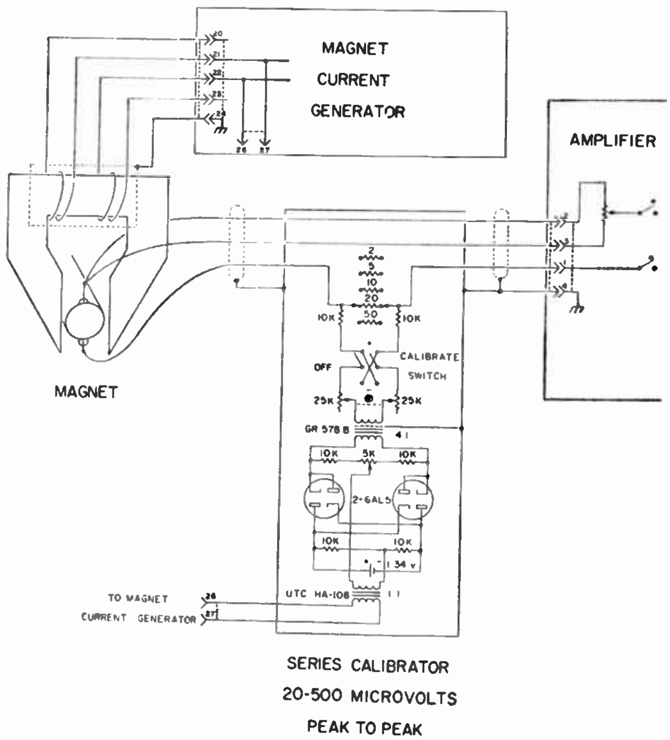


Fig. 7—Input circuit.

by the magnitude of the carrier and the diode characteristics. Isolation from ground is quite essential, and is provided by the shielded isolation transformer. The calibrator output level is adjusted by switching in the proper series resistance in the signal lead. Either phase of the calibration can be selected by the momentary, center-off, double-pole, double-throw switch.

The Amplifier

The amplifier consists of four push-pull triode stages plus an output cathode follower, and is shown in Fig. 8.

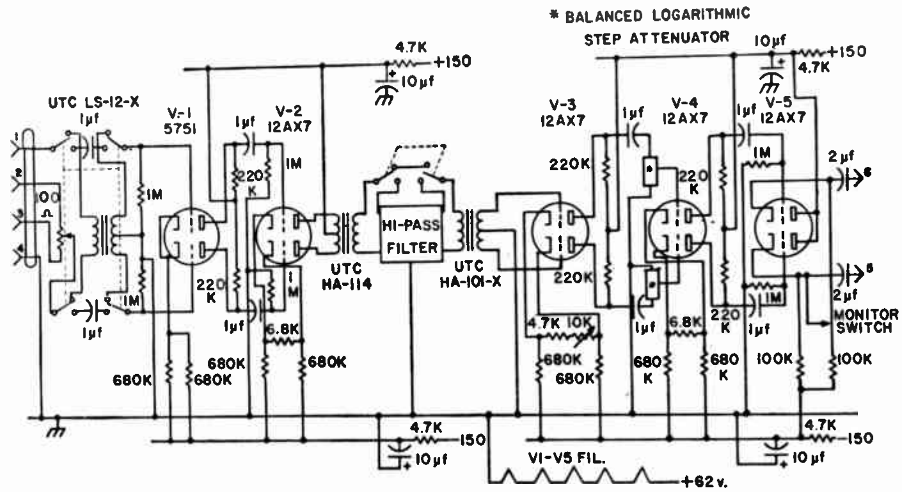


Fig. 8—Carrier amplifier. 1–4, signal input from magnet and calibrator; 5 and 6, to gate.

The switching transient balancing is done with a 100-ohm multiturn potentiometer in the split lead from the magnet assembly. This adjustment is necessary primarily to reduce the otherwise extremely large spikes which would swamp later circuits. A switch allows a choice of either low-impedance transformer input or high-impedance capacitance input. The low-impedance input has proved convenient for improving the signal-to-noise ratio where additional gain is needed in recording from smaller vessels. The high-impedance input is used for larger flows that provide signals well out of the noise. To aid stability and reduce hum and noise to a minimum, the input stage is shock mounted, and amplifier heaters are supplied from a regulated dc source. All dc supply voltages in the entire instrument are provided by electronically regulated power supplies. The coupling time constants of each stage are quite long in order to minimize phase-shift distortion of the signal square wave. The first two stages are operated at maximum gain, while the gain of the third stage is continuously adjustable over a two-to-one range by a potentiometer between the cathodes. A step attenuator between the third and fourth stages allows the over-all gain to be adjusted by a factor of 1, 2, 4, 8, or 16. Cathode degeneration in the third and fourth stages is used to assure adequate dynamic range and linearity to handle the higher signal level. Between the second and third stages provision is made for inserting a 500-ohm high-pass filter. The filter, shown in Fig. 9, consists of two half *m*-sections and two *k*-sections designed for a cutoff frequency of 75 cycles and an infinite rejection frequency of 60 cps. Plate-to-line and line-to-grid transformers are used for proper impedance matching to and from the filter. Use of the filter makes it possible to reduce unusually large interfering potentials originating from cardiac action currents or induced from the power lines. In order to operate the amplifier at a point somewhat remote from the balance of the circuitry, a low-impedance output is provided by a cathode follower stage. In

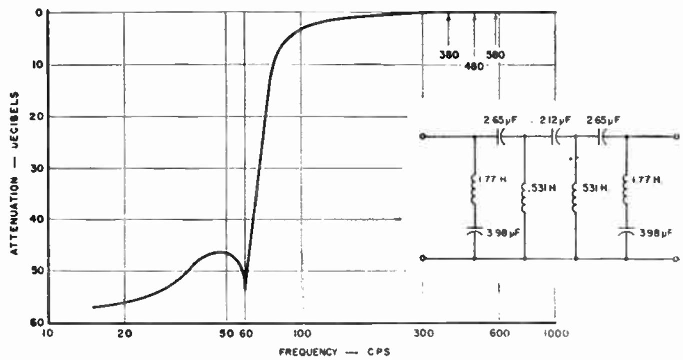


Fig. 9—500-ohm high-pass filter.

models that incorporate the amplifier with the rest of the circuits, this stage is omitted. Since the input signals are rather small and high gain is used, care is necessary in component placement and lead dress. The avoidance of any stray coupling of carrier current into the amplifier is of particular importance.

Gate and Demodulator

Fig. 10 shows the circuit details of the gate and demodulator. V6 through V10 carry signals, while V11 through V15 are concerned with providing the proper switching waveforms. The gate tube, V6, is normally held at cutoff by a high positive bias on the cathodes from V11. The gate pulse from the synchronizing pulse generator chassis is shaped by the Schmitt trigger, V12, and coupled to V6 by way of the dual cathode follower, V11. The gate pulse removes the high cathode bias, opening the gate and allowing the signal to appear at the cathodes of the stretcher diodes, V7. The gate pulse is adjusted in phase to occur toward the end of each half-period of the magnet cycle when the field is most steady. The stretcher diode conducts only during the gate-open phase and allows the signal to charge the 0.0025-μf capacitors. The discharging time constant of these capacitors during the gate-closed period is made suffi-

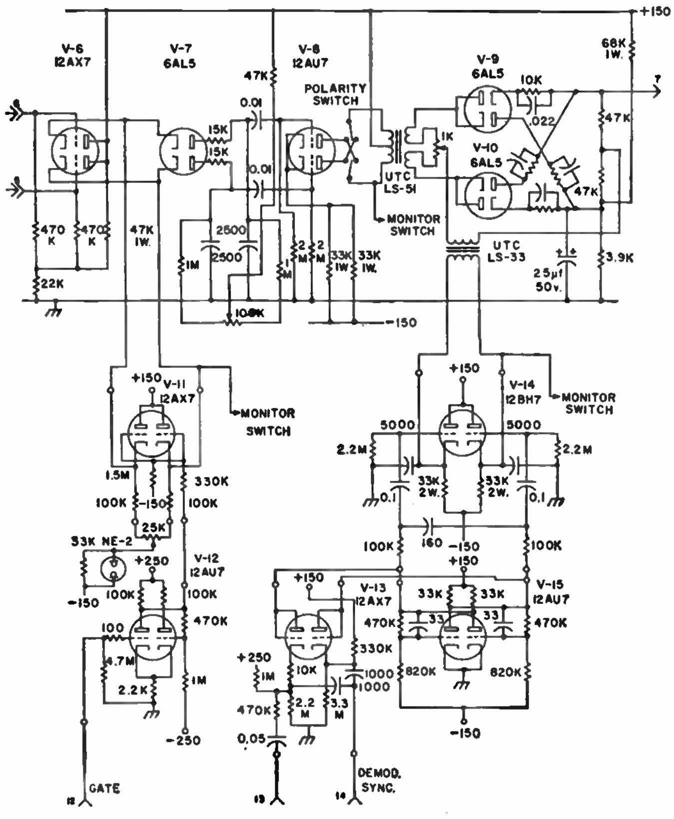


Fig. 10—Gate and demodulator. 5 and 6, from carrier amplifier; 7, to output dc amplifiers; 12, gate-pulse input; 13, demodulator phase-locking pulse input; and 14, demodulator carrier trigger input.

ciently long to stretch the gated signal out to an approximately square wave. V8 serves simply as an amplifier to provide the proper signal level for the demodulator input. A phase reversal switch in the plate circuit is a convenience in allowing selection of the proper direction of output deflection.

The demodulator is a signal- and carrier-balanced ring modulator similar to that used in the calibrator. A plate-to-line transformer is used to provide the proper impedance transformation from V8. The maximum signal level across the secondary is 10 volts, peak-to-peak. In order to provide linear operation for this signal level, a carrier of 150 volts, peak-to-peak, is used. The carrier for the demodulator is provided by a bistable multivibrator, V15, coupled through the cathode follower, V14, and the one-to-one isolation transformer. The demodulator carrier is synchronized with the gate phase by pulses from the synchronizing pulse generator coupled through the disconnect diodes, V13. In order to avoid a 180° phase ambiguity in the demodulator carrier, a 480-cps square wave from the magnet current generator is added to the 960-cps demodulator carrier trigger pulses in one of the disconnect diodes. This addition eliminates alternate trigger pulses to one half of the multivibrator. Therefore, if it should start in the wrong phase, the next synchronizing pulse is absent and causes it to wait another half cycle, whereupon it resumes operation in the proper phase.

The demodulator output consists of the flow signal, now a dc voltage of amplitude and polarity according to flow, harmonics of the carrier, and sidebands of the 960-cps gate produced by low-frequency interference such as EKG and hum. The demodulator is signal-balanced, and therefore will not pass any lower frequencies that manage to get through the gate directly. In order to remove all the undesired modulation products and interference, a low-pass filter, shown in Fig. 11, is used. The filter consists of two *m*-sections and a *k*-section designed for a cutoff frequency of 192 cps and infinite rejection frequencies of 240 and 480 cps. The impedance is 2000 ohms. The filter provides a minimum of 60-db attenuation for all frequencies above 400 cps, and yet allows absolutely uniform response to flow signals from zero to 100 cps, and is down only 1 db at 150 cps.

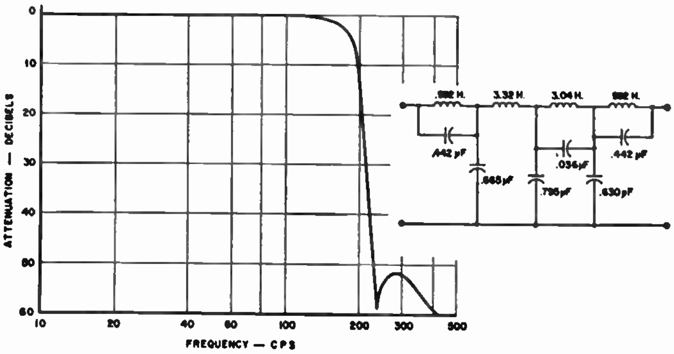


Fig. 11—2000-ohm low-pass filter. This filter curve also represents the net response of the flowmeter.

Output Stages

The output direct-coupled amplifiers are shown in Fig. 12. To provide the proper impedance match from the demodulator to the low-pass filter, a cathode follower, V16, is used. The output of the filter is divided and connected through isolation resistors to two similar direct-coupled amplifiers. One amplifier is used to provide instantaneous flow output, shown as "A," while the other provides mean flow, and is shown as "B." The amplifiers consist of two stages of voltage gain, V17 and V18, followed by the output cathode followers, V19 and V20. Only a small amount of over-all gain is needed, and in order to stabilize gain and provide good linearity, a large amount of voltage feedback is employed. The net gain is determined by the feedback resistors from V17 to V19 and V18 to V20 cathodes. The gain of the mean flow amplifier is made somewhat greater than the instantaneous flow amplifier by using less feedback. The output impedance of the amplifiers is very low due to the large amount of feedback, but is limited to 100 ohms at the output terminals by series resistors. This is done in order to avoid overdamping the particular galvanometers used for recording.

The mean-flow signal is derived from the flow signal by passing it through an RC filter with a 3.3-second time constant. The choice of time constant is a com-

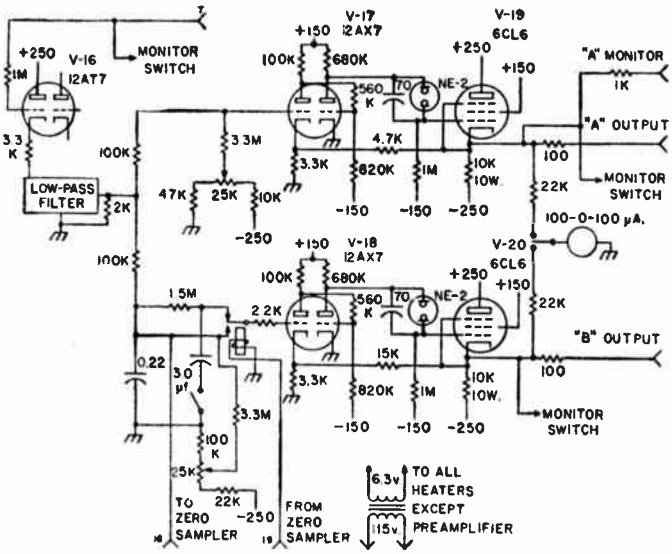


Fig. 12—Output dc amplifiers. 7, from the demodulator; "A" output, to instantaneous flow galvanometer, and "B" output, to mean flow galvanometer.

In the typical cardiac output waveform, the beginning of diastole consists of a sharp negative-going spike. Therefore, by simply differentiating this negative-going portion of the flow signal, a pulse is available for triggering a monostable multivibrator. V21 and V22 are an amplifier and clipper of the flow waveform. This signal is then differentiated and applied as a trigger through a disconnect triode to the delay multivibrator, V23 and V24. The trailing edge of the delay multivibrator output triggers a second monostable multivibrator, and the length of the zero-flow sample is set by the second or duration multivibrator. A cathode follower, V26, provides the relay pulse at the proper current level. Continuous monitoring of the time of the zero-flow sample is done by applying the signal on the grid of V26 to the *Z* or intensity axis of the oscilloscope used to monitor the flow waveform. When recording under conditions of widely varying pulse rate, it is necessary to "ride" the delay control to maintain the sampling during the zero-flow phase.

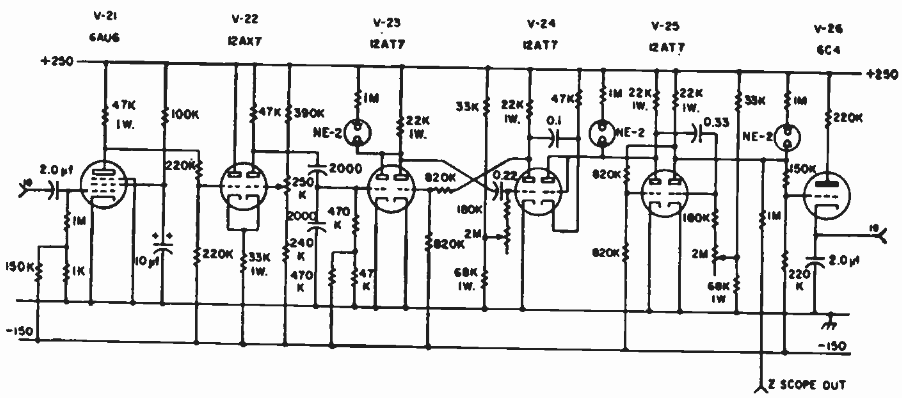


Fig. 13—Base line sampler for mean-flow channel. 18, from output dc amplifiers; 19, to sampling relay.

promise between the degree of smoothing or "integrating" desired and the desired speed of response to changes in the mean flow. Since the mean flow is only about one-fifth the magnitude of the peak flow, higher gain is used, as mentioned above. Due to slight shifts in the base line brought about by respiratory movements, in conjunction with the fact that the magnetic field is not perfectly steady during the gate phase because of a small remaining tilt in the current waveform, it was found desirable in some cases to continuously record the zero-flow level. The opportunity to do this presents itself in the ascending aorta, since during most of the latter part of diastole the net flow is zero. The base line recording is done by momentarily disconnecting the capacitor of the "integrating" RC network during an appropriate period late in diastole when the flow is zero. The resulting record is a dashed line representing mean flow, beneath which is a dotted line representing zero flow. The "zero-taker" circuit is shown in Fig. 13. It is essentially a device for operating the "integrating" capacitor disconnect relay at an appropriate time in the cardiac cycle.

Throughout the signal channel, careful attention was paid to signal level to assure linear operation. In order to avoid the possibility of operating the flowmeter in the region of nonlinearity, the bias voltages in the dc amplifier were adjusted so that complete limiting occurs above a certain signal level and is readily apparent as a sharp clipping of the flow signal waveform. Up to the clipping level, the over-all gain is perfectly constant. This has the additional advantage of preventing accidental burnout of expensive galvanometers when large voltages are introduced in the input due to manipulation in the course of an experiment.

Master Synchronizing Pulse Generator

To prevent any low-frequency beats, the flowmeter carrier is synchronized with the power line frequency by deriving the master synchronizing pulses from the 16th harmonic of the line frequency. Referring to Fig. 14, it is seen that generation of the higher order harmonics is done by a pair of full-wave rectifiers fed at 90° phase difference. A two-stage amplifier, V29, tuned to

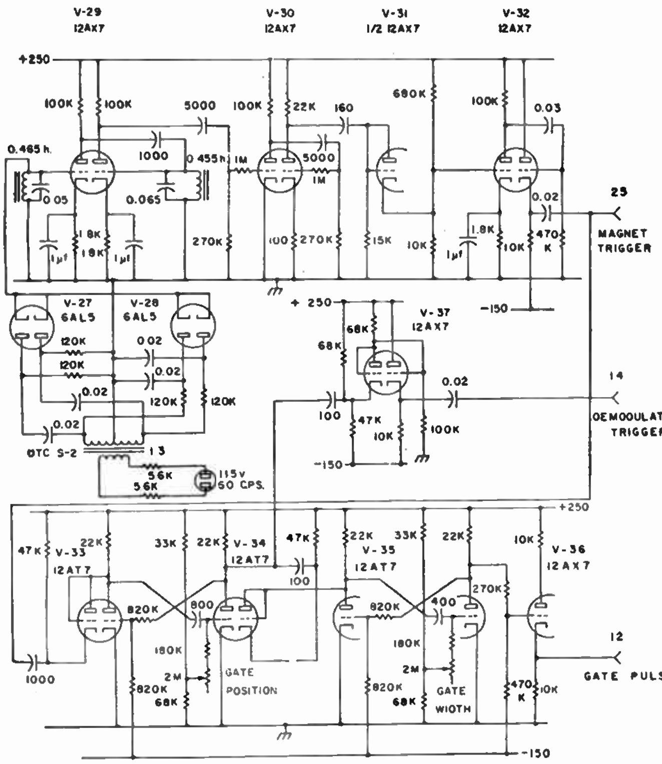


Fig. 14—Master synchronizing pulse generator. 25, trigger pulses to the magnet-current generator; 12, gate pulses to the gate circuit; 14, demodulator carrier trigger pulses.

960 cps with the aid of reasonably high-Q toroidal inductors, provides a clean sine wave. Attempts at utilizing a locked oscillator were unsuccessful, since considerable frequency modulation was always present. The 960-cps sine wave is clipped by V30 and the resulting square wave differentiated. Half of V31 is diode-connected and used to select only the positive-going sharp spike. This master synchronizing pulse, amplified by one half of V32 and passed through the other half connected as a cathode follower, is cabled to the magnet current generator chassis to trigger its multivibrator. The gate pulse is positioned by the delay monostable multivibrator V33 and V34. This is triggered by the master pulse through a disconnect diode, half of V33. The amount of delay, that is, the position of the gate-opening pulse in the magnet half cycle, is set by the length of the quasi-stable phase of the multivibrator output waveform. The trailing edge of this waveform is differentiated and used to trigger a second monostable multivibrator, V35, which generates the gate pulse waveform. The length of time that the flow signal is sampled, *i.e.*, the gate width, is determined by the width of the waveform. In order to cable the gate pulse to the amplifier chassis without distortion, a cathode follower, V36, is used to reduce the output impedance. Since the demodulator carrier must be in phase with the gate, the same delayed trigger pulse that operates the gate waveform multivibrator is used to trigger the demodulator carrier multivibrator, V15. To insure that only the trailing edge is used, and to provide the proper impedance level for cabling, V37 is used as a diode clipper and cathode follower.

The Magnet Current Generator

The magnet current generator is shown in Fig. 15. A 480-cps square wave is generated by a bistable multivibrator, V39, triggered by the 960-cps master synchronizing pulses via a pair of disconnect diodes, V38. The diodes, V40 and V41, acting as shunt limiters, determine the magnitude of the square-wave voltage applied to the driver amplifier, V42. Four dual triodes, V43 to V46, connected in push-pull parallel, serve as a class B power amplifier to provide the current output of up to 1 ampere peak. The magnet is directly coupled in the plate circuit of the output stage. Current feedback, in the form of cathode degeneration, is employed in the output stage to aid in achieving a flat-topped current waveform. The magnitude of the output current is adjusted by setting the clipping level of the shunt limiters, V40 and V41. This is done by adjusting the bias on the control tube, V48. Since the magnitude of the flow signal for a given rate of flow is directly proportional to the strength of the magnetic field, it is important that the magnet current be maintained constant at its set level. To accomplish this, the average magnet current is sampled in its ground return and applied to the grid of control tube, V48. Any change in the magnet current will then produce a correction by way of the feedback loop formed. An ammeter in the cathode circuit serves to monitor the average magnet current. V47 is used as a cathode follower to supply the magnet square wave at low impedance to the amplifier chassis for phase control of the demodulator carrier.

Performance

The usable sensitivity of the flowmeter is determined by the instrument noise rather than its gain, for galvanometers of almost any sensitivity may be connected to the output terminals. The noise generated by the input tube, V1, is by far the greatest noise source, and amounts to approximately 15 μv peak-to-peak at a bandwidth of 180 cps. Referring to the calibration graph, this noise is equivalent to a flow of 400 ml/minute for a 12.5-mm sleeve and 340 ml/minute for a 9.5-mm sleeve. The input transformer, when used, provides an additional gain of 5 ahead of the noise generated by the input tube and therefore reduces the equivalent input noise to about 3 μv or 70 to 80 ml/minute. The systolic peak flow in the ascending aorta of a 12-kg dog is approximately 10 l/minute, and therefore the peak-to-peak noise amounts to about 4 per cent of peak flow when capacitance input is used or less than 1 per cent when the transformer input is used. The input tube noise is very nearly random and therefore proportional to the square-root of bandwidth of the system. The 3.3-second time constant in the mean-flow output channel

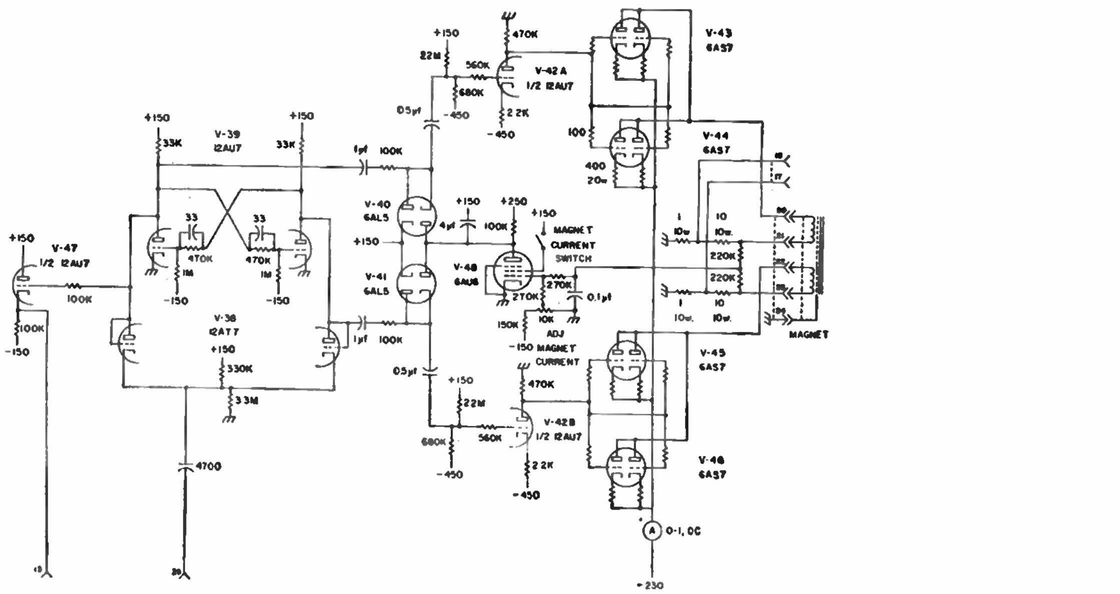


Fig. 15—Magnet current generator. 25, trigger pulse input; 13 demodulator carrier phase-locking pulse output.

reduces the bandwidth by about 3600 times, thereby reducing the equivalent noise in this channel to approximately 6 to 8 ml/minute.

The maximum transconductance of the flowmeter circuit described is 2000 mhos or 2-ma output current per microvolt peak-to-peak flow signal, capacitor input. The transformer input provides an additional gain of 5 times.

The net frequency response is determined by the low-pass filter in the demodulator output and is shown in Fig. 11. This shows the response to be substantially flat to 150 cps.

In order to take advantage of this fast response, galvanometers of high natural frequency (500 cps) are used to record the flow signal.

A typical record of instantaneous cardiac output and simultaneous carotid arterial pressure of a 12-kg dog under nembutal anesthesia is shown in Fig. 16. This record was made on a Hathaway S-14 E oscillograph using 500-cps galvanometers. The timing lines are spaced at one-tenth-second intervals. It should be noted that despite the very large EKG signal existing at the electrode site, there is virtually no trace of it present in the flow record. The brief downward deflection from the zero-flow line shows the rather high-velocity back flow at the time of closure of the aortic valve. The balance of the diastolic period shows no flow. The area under the curve represents the stroke volume. The pressure tracing was made with the aid of a Technitrol capacitance manometer.

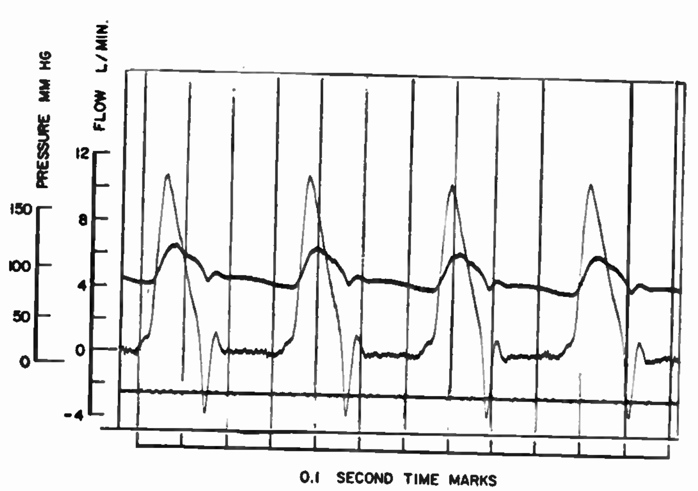


Fig. 16—Instantaneous flow in the ascending aorta and pressure in the carotid artery of a 12-kg dog.

The Use of an Analog Computer for Analysis of Control Mechanisms in the Circulation*

HOMER R. WARNER†, ASSOCIATE MEMBER, IRE

Summary—**Two approaches are presented to the study of regulation in the circulatory system. One consists of programming on an analog computer equations to represent part of the system and then, using suitable transducer, substituting the computer for the biological component. An example is presented in which a part of the mechanism which regulates arterial pressure (the carotid sinus) is simulated. The other approach involves simultaneous solution of equations derived to represent each system component. Simulation of a transient disturbance in blood distribution (Valsalva maneuver) is presented to illustrate the use of this approach in predicting the role of each component in determining over-all system behavior.**

Introduction

The human heart and circulation are a complex closed-loop system consisting of distensible reservoirs, variable flow pumps and branched transmission lines. Not only is the performance of each component determined by the behavior of its immediately adjacent component, but through the medium of the nervous system, the characteristics of each component may vary as a function of events taking place in remote areas of the circulatory system and even outside the system itself. Because of these complexities, it is no wonder that to date a satisfactory analysis of the circulation as a self-regulated system has not been undertaken. The purpose of this paper is to present two approaches to this problem, each of which involves the use of an analog computer.

Part I

One approach consists of using the analog computer to simulate part of a control mechanism. In the experiment presented here, the organ being simulated by the computer is the carotid sinus, a small organ made up of stretch-sensitive nerve endings in the wall of a large artery in the neck. The variable controlled by this organ is arterial pressure.

From work done by others[1,2] using an isolated carotid sinus preparation it is known that the frequency with which action potentials move along the carotid sinus nerve from the carotid sinus toward the brain is directly related to the pressure in the carotid artery (as long as the pressure exceeds a certain minimum value) and is also a direct function of the rate of change of arterial pressure. In the present study the transfer function used to represent the carotid sinus is

$$\frac{n}{p - p_0} = k_1 s + k_2, \tag{1}$$

where n is the frequency of impulses on the carotid sinus nerve, p is the pressure in the carotid artery, p_0 is the minimum static pressure capable of eliciting impulses on the carotid sinus nerve, s is the Laplace operator, and k_1 and k_2 are constants.

The operation of the system may be explained by referring to Fig. 1. A rise in arterial pressure results in a rise in e_i, the input voltage to the computer. The output voltage, e_0, will also rise according to the transfer function shown and result in an increased rate of stimulation of the carotid sinus nerve. The increased frequency of impulses on the carotid sinus nerve produces relaxation of the smooth muscles of small arteries and results in a fall of arterial pressure back toward the control level. By adjusting a bias voltage in the computer, the arterial pressure can be adjusted and maintained over quite a wide range.

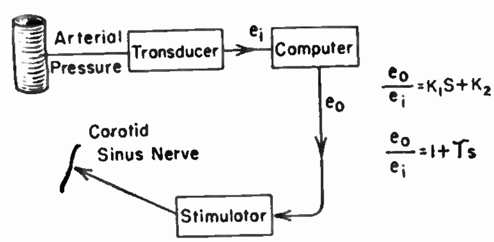


Fig. 1—Diagram of experimental arrangement employed to simulate part of the blood pressure regulating system. The rate of stimulation of the carotid sinus nerve is proportional to the output voltage from the computer e_0.

Experiments were carried out to evaluate the dynamic behavior of this system. The artificial system, when attached in parallel with the dog's own regulating mechanism, acts as an amplifier of carotid sinus function. Variations in pressure were induced by producing variation in flow. The variations in flow were achieved by stimulating a vagus nerve. The time-course of variation in the frequency of vagus nerve stimulation results in a pattern of heart rate and pressure variation of the same frequency but opposite phase. This can be seen in Fig. 2. Although the extremes of variations in amplitude of the forcing function (rate of stimulation of vagus nerve) remained constant, the resulting amplitude of the variation in pressure was frequency dependent. In the control records, the largest variations in pressure occurred when

* Original manuscript received by the IRE, June 8, 1959. Supported by Grant H-3607 U. S. Public Health Services.
† Cardiovascular Lab. of the Latterday Saints Hospital, Salt Lake City, Utah.

[1] D. W. Bronk and L. K. Ferguson, "Impulses in cardiac sympathetic nerves," *Proc. Soc. Exper. Biol. Med.*, vol. 30, pp. 339–341; December, 1932.

[2] H. W. Ead, J. H. Green, and E. Neil, "A comparison of the effects of pulsatile and non-pulsatile blood flow through the carotid sinus on the reflexogenic activity of the sinus baroceptors in the cat," *J. Physiol.*, vol. 118, pp. 509–519; December 30, 1952.

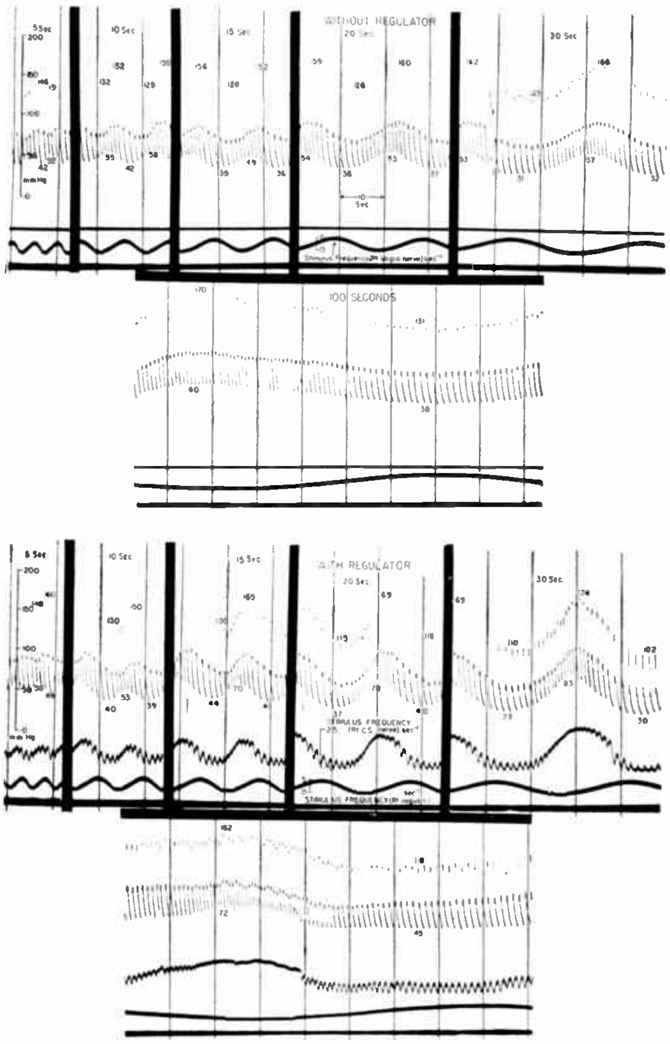


Fig. 2—The effect of amplification of carotid sinus function. Variations in carotid artery pressure are produced by variations in the rate of stimulus of the vagus nerve shown at the bottom of each tracing. Note that although the amplitude of the variations in stimulus frequency of the right vagus nerve remains constant, the amplitude of the pressure variations which result is dependent upon the period of these oscillations. Note that in the bottom set of recordings where carotid sinus function is being amplified by the electrical regulator, the oscillations are even larger.

the pressure was varied with a period of 30 seconds. That this phenomenon is the result of the activity of the dog's pressure regulating system is evident from the fact that amplification of carotid sinus function (employing the electrical regulator in parallel with the dog's own system) increases the amplitude of the pressure excursions with the forcing function unchanged from the control period. (See Fig. 3.)

The explanation for this phenomenon lies in the fact that the response of the arterial smooth muscle to nerve stimulation is very sluggish, having a time constant of approximately 10 seconds.[3] When the period of the pressure variation is 20 to 30 seconds, the phase lag between input and output of the pressure controlling system is 180°, and amplification rather than attenuation occurs.

[3] H. R. Warner, "The frequency-dependent nature of blood pressure regulation by the carotid sinus studied with an electric analog," *Circ. Res.*, vol. 6, pp. 35–40; January, 1958.

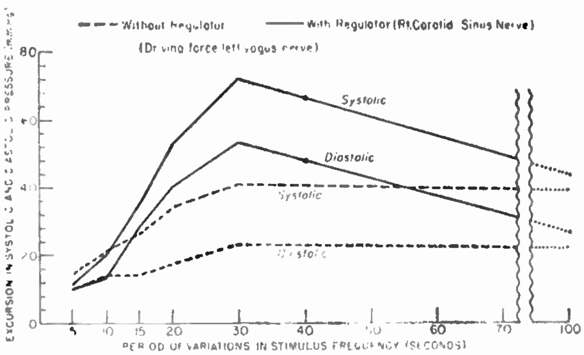


Fig. 3—Excursions in systolic and diastolic (maximum and minimum with each heart cycle) pressure measured from recording shown in Fig. 2 and plotted against the period of the variations in forcing function (stimulus frequency of the left vagus nerve).

Examination of (1) shows that the carotid sinus does have a mechanism for anticipating changes in pressure in that it is sensitive to the rate of change of pressure. This might in part compensate for the lag at the level of the artery smooth muscle were it not for the nature of the arterial pressure wave. Notice in Fig. 2 that with each heart cycle, large excursions in pressure occur. The rate of change of pressure with each heart beat is so rapid that it completely masks the slope of any variations in mean pressure. For this reason the "lead" in the carotid sinus mechanism is ineffective as a device for anticipating changes in mean pressure. That this is true is shown by the fact that k_1 of (1) could be varied from zero to large values without detectable affect on the dynamic response of the pressure regulating system.

Part II

The analog computer may be used in another way to study the regulation of the circulation. Through the simultaneous solution of a set of differential equations derived from current knowledge to represent each of the system's components, behavior of the system as an intact unit may be predicted. Verification of such predictions must then be made by comparing predicted with observed system behavior.

Fig. 4 is a block diagram of the circulation. The system is lumped as follows: the left atrium and pulmonary veins are treated as a single reservoir, the left ventricle is described as a system with two states, systole (contraction and emptying) and diastole (relaxation and filling), and the arterial bed is treated as a transmission line. The system is symmetrical, the equations of the two sides differing only in their coefficients. The analysis will involve the variables, volume, flow, pressure and time.

Fig. 5 shows the equations used to represent each of these circulatory components and a diagram of the way in which each of these equations is represented on the computer. Another set of equations identical to these except for their coefficients is used to represent the right side of the circulation. The first equation expresses the volume of the pulmonary vein-left atrial system as the sum of the initial volume and the integral of inflow

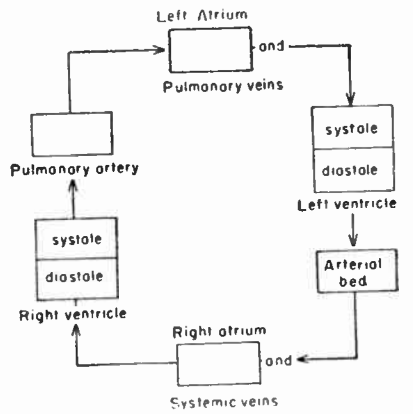


Fig. 4—This is a block diagram of the components into which the circulation is lumped for purposes of the present analysis.

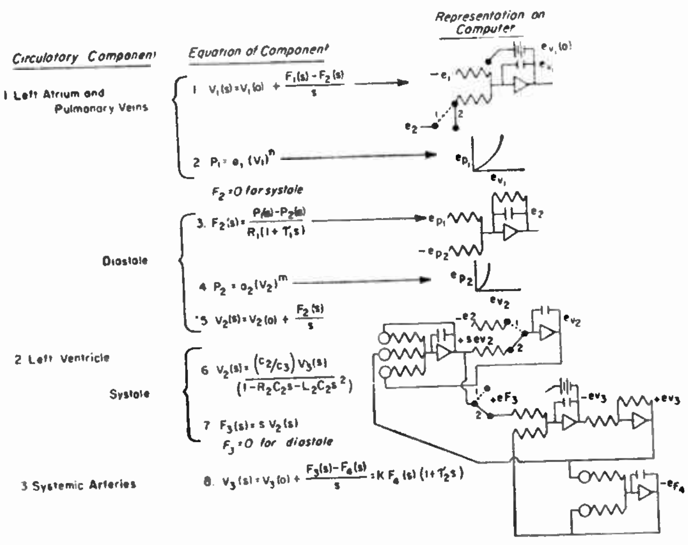


Fig. 5—Equations used to represent the systemic half of the circulation using the variables volume (v), flow (f), pressure (p) and time. Laplace operator notation (s) is used where convenient. The relays which determine the systolic or diastolic state are controlled from a magnetic tape recording of left ventricular pressure, thus synchronizing the computer operation with the biological system.

minus outflow. Laplace operator notation "s" is used where more convenient. This summation and integration is carried out in the computer by feeding the voltages representing flow 1 and minus flow 2 into an integrating amplifier. The output voltage then represents v_1. Initial condition voltage on the amplifier is set to correspond to the initial volume. Eq. (2), in Fig. 5, expresses the relationship between pressure and volume in this component. The function is plotted for the desired values of the coefficients "a" and "n" and set up on a multiple-diode function generator.

Eq. (3) (Fig. 5) represents the flow into the left ventricle during diastole as a function of the pressure gradient across the valve and a time lag (T_1) which depends on the inertia of this element. The pressure-volume curve for the ventricle is similar to that for the atrium but with different coefficients. The left ventricle during diastole will have a volume which is the integral of its inflow plus its volume at the end of the preceding systole. This is shown in (5) in Fig. 5. The volume of the left ventricle during systole is given by (6) in Fig. 5. For the derivation of this, refer to Fig. 6. During systole, the ventricle can be likened to a capacitor which has been charged up during diastole to a certain volume. The rate of discharge of this capacitor will depend upon three characteristics of the myocardium: 1) the static volume—pressure relationship represented here as C_2; 2) the coefficient of frictional forces (R_2), which limit the rate of emptying; and 3) an inertia coefficient (L_2), which limits the rate at which a given rate of emptying can be achieved. In addition, the rate of ventricular emptying will depend upon certain factors on the other side of the aortic valve; namely, the distensibility (C_3), and volume (q_3) of the aorta, the inertia of the column of blood in the aorta (L_3), and the frictional resistance to flow in the arteries (R_3). Eq. (6), in Fig. 5, is an explicit expression for left ventricular volume (v_2) during systole based on this electrical analogy.

Flow into the aorta (F_3) is zero during diastole and is given by the derivative of (6) in Fig. 5 during systole. The volume of the aorta (V_3) at any time is the integral of the difference between inflow and outflow of the arterial bed. Flow out of the arterial bed into the vein (F_4) is related to the aorta volume by a constant and a lag factor with a time constant of T_2.[4] A similar set of equations, but with different coefficients, is used to represent the right side of the circulation. In the computer the systolic and diastolic states are determined by two alterative positions of a set of relays. These relays are synchronized with systole and diastole of the experimental animal being studied. From a recording on magnetic tape of the subject's ventricular pressure pulse a signal is derived which allows the computer to trigger the relays at the proper time.

Fig. 7 shows examples of the type of systematic analysis that can be performed with this technic. Here the effect on the predicted time course of left ventricular volume and aortic inflow of changing certain equation coefficients is shown. Each record shows a control tracing together with the function obtained after increasing one of the system's parameters by a factor of two. A decrease in stroke volume may result from increasing either the frictional coefficient (R_2) or the capacitance (C_2) of the ventricle. Ignoring inertia of the ventricle (L_2) produces an unrealistic initial spike on the aortic inflow curve. Increasing peripheral resistance (R_3) decreases stroke volume, while increasing arterial capacitance increases stroke volume. Increasing diastolic capacity of the ventricle ($1/a_2$) increases both stroke volume and end systolic residual volume. Each of the system's parameters may be systematically studied in this manner. It seems likely that the rapid progress being made in the development of transducers should soon make it possible to test the accuracy of these predictions in experimental animals.

[4] H. R. Warner, "A study of the mechanism of pressure wave distortion by arterial walls using an electrical analog," *Circ. Res.*, vol. 5, pp. 79–84; January, 1957.

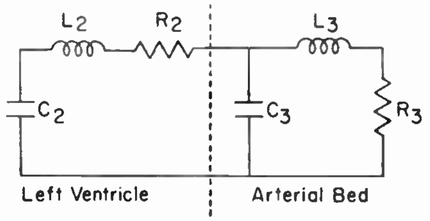


Fig. 6—Electrical analog used to derive the equations describing the performance of the left ventricle and the arterial bed.

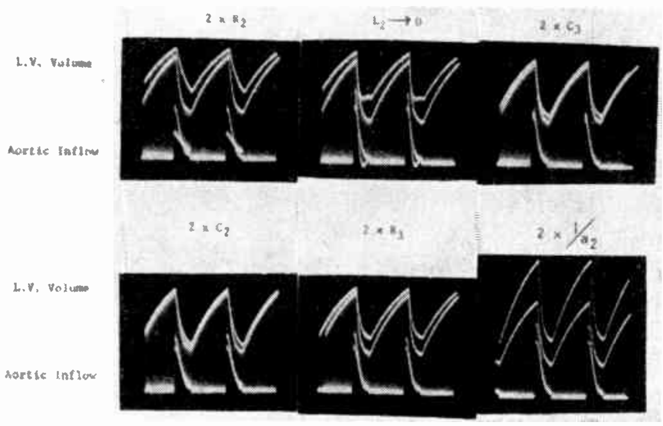


Fig. 7—The effect on the predicted time course of left ventricular volume and aortic inflow of changing certain equation coefficients. Each record shows superimposed a control tracing and the waveform obtained when the system parameter designated is doubled.

Using the equations here presented, a prediction may be made of the response of the whole circulatory system to a transient disturbance. Since the system is a closed loop, such a prediction involves solving all the equations of the system simultaneously. One such solution is shown in Fig. 8. The physiologic situation being simulated is called a Valsalva maneuver and is performed by the subject attempting a forced expiration against a closed glottis. Such a maneuver increases intrathoracic pressure and thus prevents blood from returning to the heart. Because all the blood which leaves the lungs is not replaced, a redistribution of blood between the pulmonary veins and systemic veins occurs. Thus, to simulate the state of affairs at the end of a Valsalva maneuver the initial voltage on the pulmonary vein—left atrial integrator is made lower than its equilibrium value and the voltage on the right atrial-systemic vein integrator higher than its equilibrium value. Upon starting the problem, the predicted response of the system to release of the Valsalva can be observed.

The solution shown on the left in Fig. 8 was obtained with constant coefficients in all of the equations. This predicted time course of aortic volume is similar to the time course of aortic pressure obtained in an animal deprived of reflex activity by prior administration of blocking drugs. The response shown in the tracing on the right was obtained by allowing peripheral resistance (R_3) to vary as a function of aortic volume and a time lag such as was demonstrated in the first part of this paper for the carotid sinus mechanism. An overshoot is

Fig. 8—Time course of aortic volume predicted from simultaneous solution of all sixteen equations when R_3 is constant (*left*) and when R_3 is allowed to vary as a function of arterial volume (*right*).

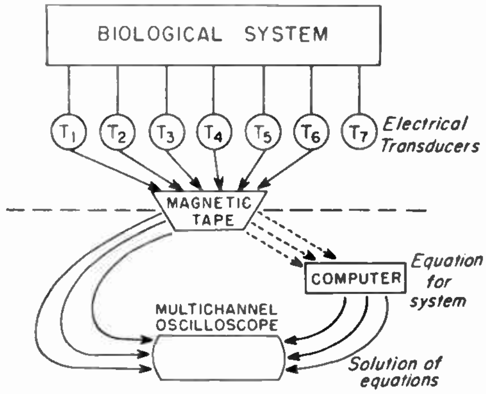


Fig. 9—Block diagram of system currently being employed to evaluate the transfer function of biological systems in the author's laboratory.

now evident such as is seen in the arterial pressure curve of a normal subject with an intact reflex system following release of a Valsalva maneuver.

This set of equations predicts a rapid return to equilibrium for the circulatory system following transient disturbance of several types. This is true in the presence of a wide variety of values for each of the system's parameters with one exception, the exception being the case of very high flow in which some overshoot may be observed even with constant coefficients.

The adequacy of the equations presented is currently being evaluated using the system shown in block diagram in Fig. 9. Up to seven variables may be recorded from the biological system during an experiment. These are recorded on multi-channel magnetic tape using frequency modulation. The tape recorder has two transports. Data are originally recorded using a reel-to-reel transport and then segments are copied onto a tape loop for analysis along with additional programming data. Thus, transient phenomena recorded on the tape may be reproduced over and over again for analysis. Some variables reproduced from the tape are displayed on a multichannel oscilloscope for comparison with values for the corresponding variables being predicted by the computer. The computer coefficients are varied until the best possible fit between predicted and recorded waveforms is obtained. If a solution is not found which accurately predicts the experimental observation, a new equation must be sought. On the other hand, when an equation is found whose solution conforms to the experimental data, the generality of this equation must be judged by its ability to predict system behavior under other experimental conditions.

Some Engineering Aspects of Modern Cardiac Research*

D. BAKER†, STUDENT MEMBER, IRE, R. M. ELLIS‡, D. L. FRANKLIN†, AND R. F. RUSHMER†

Summary—**Classical investigation into the function and control of the heart has been conducted on anesthetized open-chested dogs. Unfortunately, both the anesthetics and the exposure of the heart affect cardiac function. Hence, more realistic information would be obtained if the heart could be studied in intact conscious animals. A system has been developed to make possible continuous analysis of the action of the heart in the healthy unanesthetized dog during its spontaneous activities. This system involves the continuous measurement of the pressure within the chambers of the heart, the size of these chambers, and the flow of blood out of the heart. Heart rate, stroke volume, average blood flow, effective cardiac power and work, and other information are continuously derived from the directly-measured parameters by means of analog computers. Several new instruments were developed to solve the problems unique to measurement in an intact animal. The dimensions of the heart chamber are obtained by measuring the transit time of pulsed sound passing across the chamber. Blood flow is measured by comparing the upstream and downstream transit times of bursts of sound passing through the moving blood. An isothermal flow meter utilizing a tiny thermistor on the tip of a catheter provides an alternate measure of flow. A miniature, differential transformer type of pressure transducer was developed for measuring pressure within a heart chamber. The system provides a means by which hypotheses regarding cardiovascular function and control may be rapidly and accurately evaluated.**

Introduction

ANALYSIS of the heart as a pump poses problems not generally encountered in engineering practice. Since the heart is normally inaccessible, direct measurement of crucial variables is extremely difficult. Measurements made while the heart is exposed may not reveal its normal function. Exposure of the heart greatly alters its environment, so that its activity is distorted. In addition, the anesthetics necessary to the surgical operation greatly depress, or eliminate, the controls normally exerted by the nervous system. Nevertheless, the classic concepts of cardiac function and control have been derived primarily from studies conducted on exposed or isolated hearts of experimental animals. It is obvious that more realistic information can be obtained if the size and function of the heart are measured in intact animals.

The standard method of studying the size and shape of an internal organ is by means of X rays. Although the combination of X-ray techniques with still or motion picture photography has provided evidence of errors in cardiologic theory and is valuable in the diagnosis of heart diseases, application of these methods to the study of the normal changes in the heart and its four chambers is limited. Among the limitations is the need for many serial exposures which must be individually measured. Also, since the subject must remain squarely between the X-ray tube and the photographic film or flourescent screen, responses to common activities such as running cannot be studied.

For all these reasons, a completely new set of electrical and electronic techniques was developed. These permit continuous measurement of the basic parameters of cardiac function in intact conscious dogs while they are eating, sleeping, and exercising; startle reactions and other emotional responses can also be recorded. The use of analog computers eliminates much tedious analysis for derived functions; storage of basic experimental data on tape makes it possible to perform these analyses at leisure.

The requirements for the gauges were unique and very stringent. For example, the transducers to be installed within the chest must be sterilized and must produce only minimal interference with the function of either the heart or the lungs. The wires must extend through the body wall and sustain repetitive flexion by each heart beat and breathing movement without breakage. To assure complete recovery of the animal before measurements are made, the gauges must function for considerable periods while continuously bathed in fluids which are as corrosive as sea water. The instruments must not cause blood clots or be inactivated by formation of scar tissue. It was deemed necessary that the output of the gauges be recorded continuously by direct-writing instruments so that the changes in heart function could be accurately related to specific actions such as running on a treadmill and external events such as the slamming of a door.

The function of a pump can best be described in terms of changes in three basic parameters: dimensions, pressure, and flow. Specially designed gauges which meet the general criteria listed above have been successfully used to measure each of these variables in the enclosed heart. The derived functions obtained by means of the analog computers have included heart rate in beats per minute, instantaneous rate of change of dimensions, a function of "power" developed by the contracting muscle, a function of "work" performed per

* Original manuscript received by the IRE, July 16, 1959. This work was supported by grants from the National Institutes of Health and the American Heart Association.

† Dept of Physiology and Biophysics, University of Washington School of Medicine, Seattle, Wash.

‡ Seattle Development Laboratory, Minneapolis-Honeywell Regulator Co., Seattle, Wash.

stroke, and accumulated "work" per unit time.[1] One device, an ultrasonic flowmeter, registers instantaneous volume flow per unit time. Its output is integrated to determine volume flow per stroke and accumulated volume flow per unit time.[2]

Measurement of Heart Dimensions

The principal dimensions to be measured are the diameter, the length, and the circumference. The gauges used successfully for these measurements have included variable-inductance diameter and length gauges, variable-resistance circumference and length gauges, and mutual-inductance and sonar diameter gauges.

Variable Inductance Gauge

This linear displacement transducer may be used to record the distance between two points when the space between them is unobstructed.[3] The gauge consists of a long thin coil of copper wire wound on a nylon cylinder into which a magnetic stainless-steel core can be slid. The end of the coil is fastened to the inner side of one wall of a heart chamber and the end of the core is fastened to the opposite wall. The inductance of the coil is varied by the position of the core and is therefore a measure of the distance between the two points of attachment.

A precalibrated gauge of this type is reliable and stable. It may be balanced in the bridge circuit of a carrier-wave strain-gauge amplifier, and the signal applied to a direct-writing recorder. Coil transducers with a diameter of 2 mm and a working range of 1.5 cm to 3 cm have been used very successfully to record the inside diameter and length of the left ventricle in dogs' hearts. A typical application is shown in Fig. 1(a). Although gauges of this type have the disadvantage of requiring a difficult surgical operation for placement, several have functioned satisfactorily for as long as 26 days after surgical installation.

Variable Resistance Gauge

A small-bore rubber tube filled with mercury can serve as a simple and effective dimensional gauge[4,5] when electrodes are attached to its ends. As the tube is elongated, the mercury column has greater length and a smaller cross-section area. The electrical resistance of the mercury column is a function of the length of the gauge. Measurement of the resistance change is accomplished by connecting the gauge as one arm of a bridge circuit excited by either alternating or direct currents. This gauge has been employed in many applications, a few of which are illustrated in Fig. 1(b).

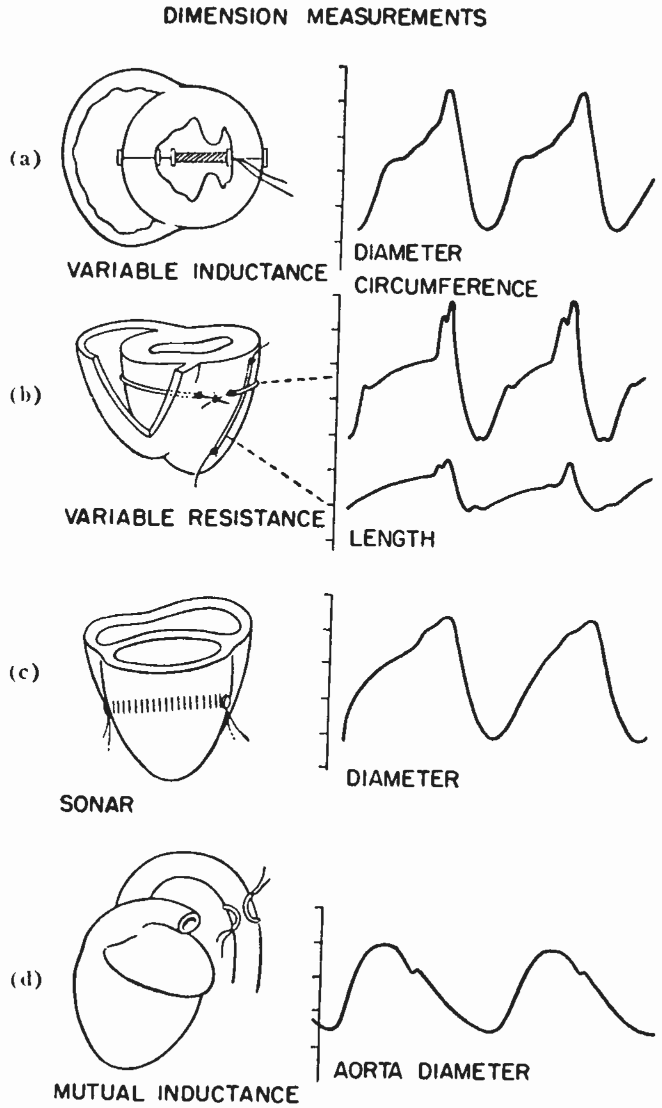


Fig. 1—Diagrams of various techniques used to provide instantaneous and continuous records of heart dimensions and representative tracings. (a) Changes in diameter vary inductance of coil; electrical change is recorded by means of Sanborn carrierwave amplifier. (b) Mercury-filled rubber tubes change their electrical resistance when stretched. When such a tube is attached to heart wall, circumference or length can be measured. (c) Transit time of sound bursts traveling between barium titanate crystals is a measure of distance. (d) Amount of coupling between two adjacent coils is a measure of distance between them.

The variable resistance gauge provides useful information concerning changes in dimensions, but lacks the long-term stability required in typical chronic preparations. The length and circumference of the heart and the circumference of the aorta[6,7] and other vessels have been successfully recorded over periods of one to two weeks.

[1] R. F. Rushmer and T. C. West, "Role of autonomic hormones on left ventricular performance continuously analyzed by electronic computers," *Circ. Res.*, vol. 5, pp. 240–246; May, 1957.

[2] D. L. Franklin and R. M. Ellis, "A pulsed ultrasonic flowmeter," *Federation Proc.*, vol. 17, p. 48; March, 1958.

[3] R. F. Rushmer, R. M. Ellis, and A. A. Nash, "Continuous measurements of left ventricular dimensions in intact, unanesthetized dogs," *Circ. Res.*, vol. 2, pp. 14–21; January, 1954.

[4] R. F. Rushmer, "Length-circumference relations in the left ventricle," *Circ. Res.*, vol. 3, pp. 639–644; November, 1955.

[5] R. F. Rushmer, "Pressure-circumference relations of the left ventricle," *Amer. J. Physiol.*, vol. 186, pp. 115–121; July, 1956.

[6] R. F. Rushmer, "Pressure-circumference relations in the aorta," *Amer. J. Physiol.*, vol. 183, pp. 545–549; November, 1955.

[7] A. P. Greer, H. Irisawa, and R. F. Rushmer, "Applications of electrokymography based on a comparison with records from dimensional gauges in situ," *Amer. Heart J.*, vol. 57, pp. 430–437; March, 1959.

Mutual Inductance Gauge

A device utilizing the effect of mutual inductance is presently being applied to the study of the pressure-volume relationships in the large veins leading to the heart.[8] The technique previously described by Perry and Hawthorne[9] is quite simple and extremely effective [Fig. 1(d)]. It consists of placing two copper coils on the opposite sides of a large blood vessel. One coil is driven at frequencies up to 600 kc; the other, a pick-up coil, has an output proportional to the distance between the coils. The output from the pick-up coil is detected and filtered for additional amplification and presentation on a Sanborn recorder. Through careful shaping of the coils, the output can be rendered quite linear over a small range of diameters (*e.g.*, 0.01 cm and 3 cm). The mutual-inductance gauge has been used on blood vessels of moderate size (*e.g.*, the coronary arteries). Its construction and compact size give it attributes well suited to measurements in intact animals, provided the angular rotation can be controlled.

Sonocardiometer

The principles of sonar and pulse timing circuitry provide a means of measuring the dimensions of internal organs in the intact animal. The first application of pulsed sonar at this laboratory was for measurement of the diameter of the left ventricle of the dog's heart.[10] Preliminary investigations revealed that the velocity of sound in tissue is approximately 1500 mps. The distance to be measured was expected to vary between 25 mm and 70 mm. This range corresponds to transit times varying from 16 μsec to 45 μsec, which can be measured easily.

Transducers made of barium titanate crystals with a resonant frequency of 3 mc (in the thickness mode) were chosen to obtain good resolution. The crystals were mounted on opposite ventricular walls to serve as transmitter and receiver respectively [Fig. 1(c) and Fig. 2].

The transmitter crystal was excited through miniature coaxial cables from the transmitter located in a relay rack. When short bursts of 3 mc sound are passed between the two crystals the transit time is a measure of the distance between them. The repetition rate of these sonic bursts is variable from 1000 to 2500 pps. A bistable multivibrator can be used to generate a pulse whose width is proportional to the transit time of these bursts. This is accomplished by triggering the multivibrator at a time slightly delayed from the instant of transmission, and then having it turn off at the instant the signal is received. The bistable pulse can be integrated and filtered to provide a dc output voltage which can be calibrated in terms of a linear dimension. Since the repetition rate of the transmitted pulses is over 1000 pps, the frequency response of the sonocardiometer is in excess of 50 cps. This response characteristic makes it possible to measure very rapid changes in dimensions. The application of the sonocardiometer is by no means restricted to dimensional measurements of the heart. Two of these instruments have been successfully used to register simultaneously the dimensions of the spleen and liver.[11] In this instance the circuits were synchronized as would be required whenever the sonocardiometer is used with other pulse equipment. In general, this instrument can measure the distance between the transducers in many different biological structures. The present effort to improve this instrument is to extend its usable operating range for measurements down to 2 mm so that small blood vessels can be studied.

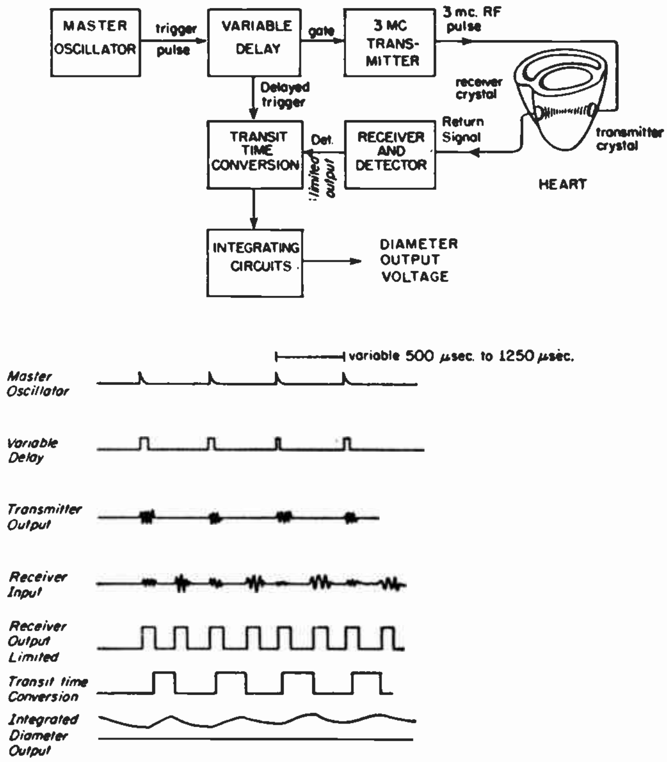


Fig. 2—Sonocardiometer. Top, block diagram of instrument to measure travel time of 3 mc bursts of sound passing between two crystals attached to the heart. A short time after the transmitter fires, a bistable circuit is triggered. This transit time conversion pulse is terminated by the received signal. The width of the resulting pulse is proportional to travel time and when integrated forms a voltage which can be calibrated in units of distance. Bottom, waveforms obtained.

Measurement of Pressure

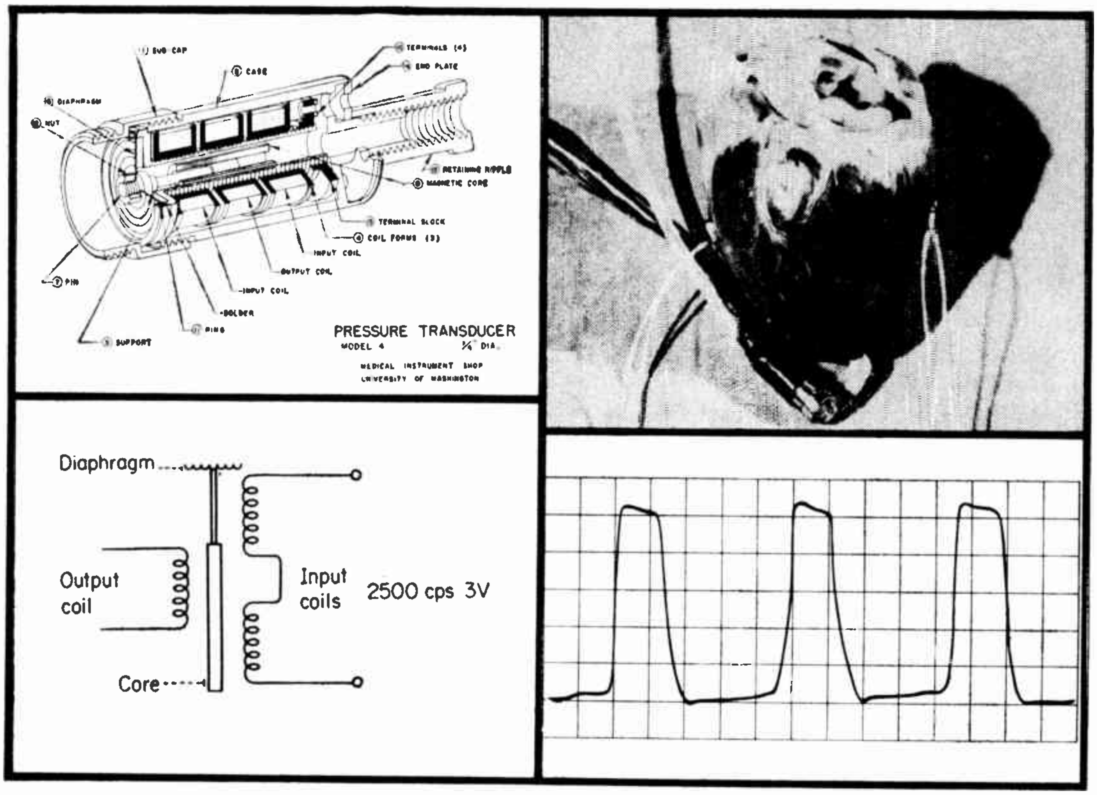
Standard commercial pressure transducers are sufficiently sensitive to record the pressure developed within the chambers of the heart. However, these gauges can-

[8] H. Irisawa, A. P. Greer, and R. F. Rushmer, "Changes in the dimensions of the venae cavae," *Amer. J. Physiol.*, vol. 96, pp. 741–744; April, 1959.

[9] S. L. C. Perry and E. W. Hawthorne, "A new method for measurement of instantaneous dimensional changes of left ventricle, kidneys and other organs in animals," *Federation Proc.*, vol. 17, p. 123; March, 1958.

[10] R. F. Rushmer, D. L. Franklin, and R. M. Ellis, "Left ventricular dimensions recorded by sonocardiometry," *Circ. Res.*, vol. 4, pp. 684–688; November, 1956.

[11] W. G. Guntheroth, "Function of liver and spleen as venous reservoirs," *Federation Proc.*, vol. 17, p. 63; March, 1958.



(a) (b) (c) (d)

Fig. 3—Miniature pressure transducer for internal use in intact animals. (a) Cutaway view; displacement pressure applied to diaphragm (16) is coupled to core (8), which causes electrical unbalance in differential transformer windings. (b) Complete gauge positioned to measure pressure in the left ventricle of the dog heart. (c) Schematic diagram of transformer. (d) Waveform of left ventricle pressure.

not be installed within the chest cage because they are too large, too heavy, and will not withstand the corrosive action of body fluids. To eliminate these objections, a miniature differential-transformer pressure transducer was specifically designed for mounting at or near the heart in dogs to be studied some time later. This instrument was designed by Wayne E. Quinton, Supervisor of the Medical Instrument Shop of the Health Sciences Division, University of Washington, Seattle. A pressure difference across the diaphragm displaces the iron core within the differential transformer to produce an electrical unbalance [Fig. 3(a), 3(c)]. The resulting output is amplified and demodulated by a Sanborn strain-gauge amplifier and presented continuously on recording paper.

The average sensitivity of these gauges is 0.0384 volt, per mm Hg, per volt applied to the input coils. The sensitivity has been checked over a period of several months and found to be very stable. The zero drift has been measured at less than 0.2 mm Hg/°C. The frequency response exceeds 50 cps.

Because of its small size ($\frac{1}{2}$ inch by $\frac{1}{4}$ inch), the gauge can be readily mounted adjacent to the heart, with connecting tubing extending into its pumping chambers [Fig. 3(b)]. Such gauges performed satisfactorily for days or weeks after being installed within the bodies of experimental animals.

Measurement of Flow

Continuous measurements of the instantaneous flow into and out of the heart are essential for a complete understanding of its function. By relating blood flow to other continuously recorded parameters, one can derive many otherwise hidden relationships. A wide variety of techniques are available for recording blood flow into or out of the heart in anesthetized animals. In general these methods correspond to those employed in measuring flow of fluids through pipes and include orifice meters, rotameters, and bubble and bristle flow meters. In each case, the sensing element is in contact with the blood and therefore is not suitable for measurements over days or weeks because blood clots change the calibration. For studies of blood flow in intact animals, flow meters must be applied to the outside of blood vessels, must have long-term stability of base line and calibration, and must be sensitive to levels of flow between zero and several liters per minute.

The electromagnetic flow meter promises to be a very useful instrument for these purposes.[12,13] However, our

[12] A. B. Denison, Jr., M. P. Spencer, and H. D. Green, "A square-wave electromagnetic flowmeter for application to intact blood vessels," *Circ. Res.*, vol. 3, pp. 39–47; January, 1955.

[13] A. Kolin, "An ac induction flowmeter for measurement of blood flow in intact blood vessels," *Proc. Soc. Exper. Biol. Med.*, vol. 46, pp. 235–239; February, 1941.

attention was directed to the development of pulsed ultrasonic flow meters, which were promptly employed for studies in intact animals during various kinds of activity.[2,14] An isothermal flow meter has also been developed and used successfully.[15,16] The effect of nuclear magnetic resonance has also been used in measuring blood flow.[17]

Sonic Flow Meter

Experience gained from the development and application of the sonocardiometer suggested the possibility of building a flow meter utilizing similar techniques. If bursts of 3-mc sonic waves pass diagonally through a stream of flowing blood, the transit time between two points is longer when the waves pass upstream, and shorter when the waves pass downstream. The difference in transit time upstream and downstream is proportional to the mean velocity of blood flow. If the diameter of the segment of vessel between the crystals is held constant, the volume flow per unit time is proportional to the mean velocity. With this technique a flow section was constructed with piezo-electric crystals mounted as shown in Fig. 4.

Under these conditions the transit time of ultrasonic vibrations upstream between the two crystals can be expressed by:

$$tu = \frac{d}{c - V} \cos \theta.$$

Conversely the transit time downstream can be expressed by:

$$td = \frac{d}{c + V} \cos \theta,$$

d = distance between crystal faces,
c = velocity of sound in still medium,
θ = angle between crystal and flow axis.

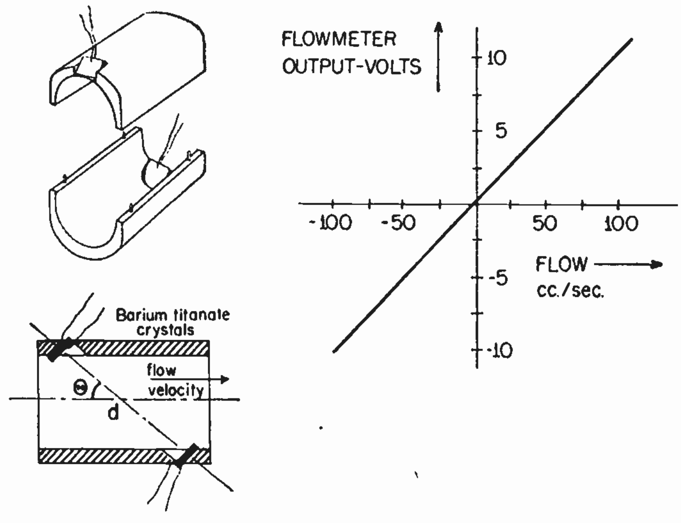
These equations can be arranged to show the difference in transit axis time:

$$\Delta t = td - tu = \frac{2dV}{c^2 - V^2} \cos \theta.$$

Since the stream velocity is less than one thousandth of the velocity of sound in a still medium, the term V^2 can be considered negligible. Since the angle ϕ and the distance between crystals are held constant,

$$\Delta t = \frac{2d}{c^2} \cos \theta V = KV.$$

[14] D. L. Franklin, R. M. Ellis, and R. F. Rushmer, "Aortic blood flow in dogs during treadmill exercise," *J. Appl. Physiol.* (in press).
[15] S. Mellander and R. F. Rushmer, "Venous blood flow recorded with an isothermal flowmeter," *Acta Physiol. Scand.* (in press).
[16] S. Mellander, "Venous blood flow recorded with an isothermal flowmeter," *Federation Proc.*, vol. 17, p. 394; March, 1958.
[17] P. Buchman, "Nuclear Magnetic Resonance Blood Flow-Meter," M.S. thesis, University of Washington, Seattle; 1959.



(a) (b)

Fig. 4—(a) Flow section for ultrasonic flow meter has barium titanate crystals mounted to pass 3-mc sound bursts diagonally through a blood vessel. Blood flow is not restricted when the section is mounted around a vessel. (b) Flow meter output voltage is a linear function of flow.

Although the velocity of sound c in the medium has been found to remain unchanged during the period of the experiment, this potential variable can be monitored at will by using an oscilloscope.

The time of travel of a short burst of sound from one crystal to the other is measured by using a ramp voltage. During the transit time of the sound a capacitor is charged at a constant rate (ramp function). The arrival of the ultrasonic waves at the receiver crystal triggers the discharge of this capacitor. The maximum voltage attained is a measure of the transit time. To measure the difference in transit time a series of sonic bursts are switched alternately upstream and downstream by mechanical choppers.

At zero flow the transit times upstream and downstream are equal. The envelope of the ramp voltages upstream and downstream is a 400-cycle square wave with an amplitude which is proportional to the difference in transit time. This square wave is amplified, synchronously detected, and filtered to provide a dc voltage which can be calibrated in terms of flow.

The primary circuit blocks and the significant waveforms for one series of transmitted pulses either upstream or downstream are shown in the diagrams of Figs. 5 and 6.

The synchronizing circuits generate a series of pulses (6–10) which trigger the transmitter, shown on line 1 in Fig. 6(a). The transmitter responds to each trigger pulse by exciting the transmitter crystal with 0.2-μsec 80-volt pulse. These pulses are coupled through choppers to the transmitting crystal on the flow section. The receiver input appears on line 3 of Fig. 6(a). At the instant the transmitter fires, an electrical coupling occurs at the receiver input. A few microseconds later, according

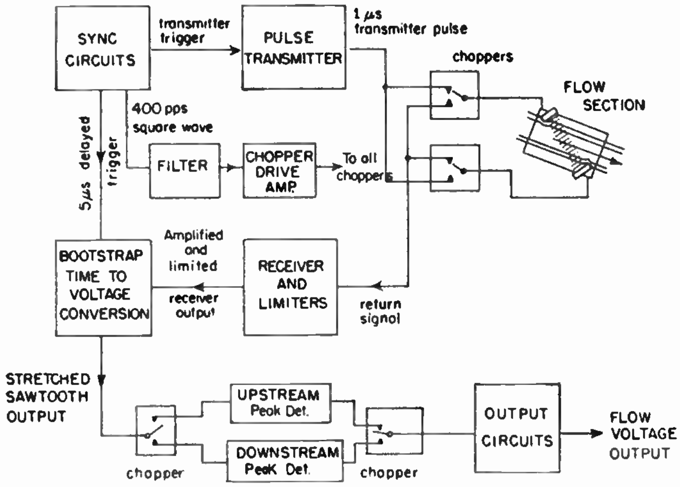


Fig. 5—Operational block diagram of a pulsed ultrasonic flow meter used to measure blood flow in intact unanesthetized dogs. When groups of 3-mc sound bursts are alternately sent upstream and downstream, difference in travel time is a measure of flow velocity. Device uses standard pulse techniques.

6(b). The choppers, Fig. 5, switch the pulse trains into their proper channels, the upstream pulse trains to the upstream peak detectors, and so on. The difference in voltage levels appearing at the output of the upstream and downstream peak detectors is a measure of the difference in transit time. The chopper preceding the output circuits alternately picks these voltages off in synchrony with the upstream and downstream pulse trains. The result is the square wave shown in Fig. 6(b), line 2. The output circuits synchronously detect and filter this wave to provide a voltage output which is proportional to flow.

Synchronism of several flow meters is easily achieved due to the pulse nature of the circuitry. This characteristic makes possible the instantaneous measurement of flow at several points. The sonocardiometer has been successfully synchronized with the flow meter on many occasions to secure simultaneous dimension and flow measurements.

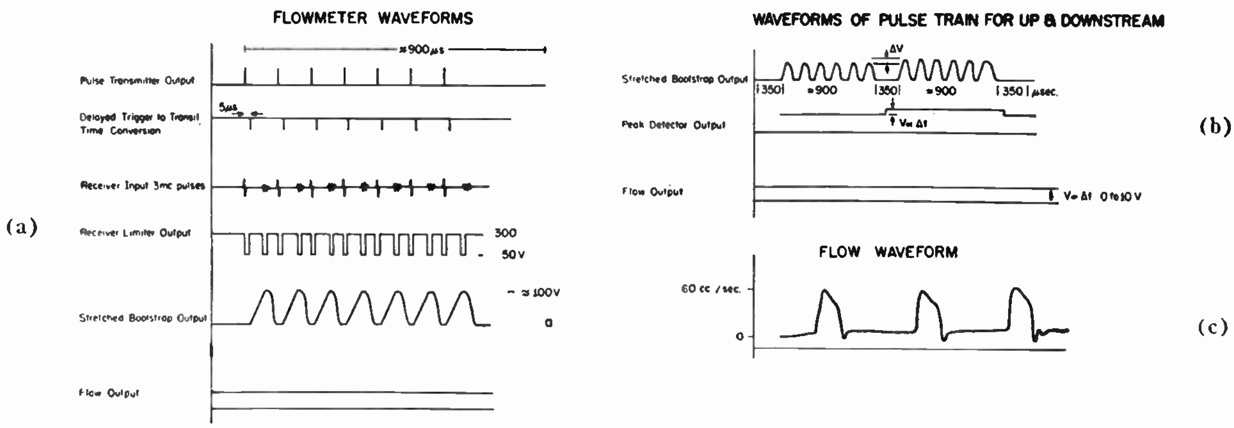


Fig. 6—(a) Waveforms for the ultrasonic flow meter illustrating timing of transmissions in either upstream or downstream direction. Groups of 6–10 pulses are sent alternate directions at a rate of 800 groups per second. Amplitude of linear voltage rise (bootstrap output) is proportional to travel time of transmitted 3 mc sound burst. (b) Difference in amplitude of stretched sawtooths is proportional to difference in transit time between upstream travel and downstream travel. A voltage representing this difference is a measure of flow velocity. (c) Waveform of aortic flow is measured with this device.

to the transit time, the sonic vibrations traveling through the blood arrive at the receiving crystal. The small electrical signals appearing at the receiver input are amplified and limited. This limited output appears on line 4 in Fig. 6(a). From the synchronizing circuits comes another output, a delayed trigger, as indicated in line 2 of Fig. 6(a). It controls a bistable multivibrator which gates the linear ramp function generator (bootstrap circuit), line 5 in Fig. 6(a). The receiver output pulse corresponding to the received signal is differentiated and used to close the gate, interrupting the ramp function. The amplitude attained by the ramp function is then a measure of the transit time. Measurement of the peak amplitude is facilitated by stretching the bootstrap sawtooths by a ringing circuit [line 5, Fig. 6(a)]. An upstream and downstream pulse train to illustrate the difference in peak amplitude, corresponding to a difference in transit time, appears in line 1 of Fig.

Isothermal Flow Meter

When a heated thermistor bead is placed in a flowing stream, the rate at which heat is carried away by the stream is a measure of the mean flow velocity. A flow meter based on this principle has been developed in this laboratory.[15,16] A thermistor bead is attached to a long wire-like probe called a catheter. The thermistor can easily be placed in a blood vessel by threading the catheter down the vessel from some convenient entry point.

The thermistor bead is connected into one arm of a balanced dc bridge circuit so that cooling of the thermistor by increased flow will cause unbalance of the bridge. A change of balance is thus a measure of the temperature of the thermistor. This effect is utilized to provide an error signal to a circuit whose purpose is to heat or allow cooling of the thermistor so that balance of the bridge is maintained. The output of this circuit

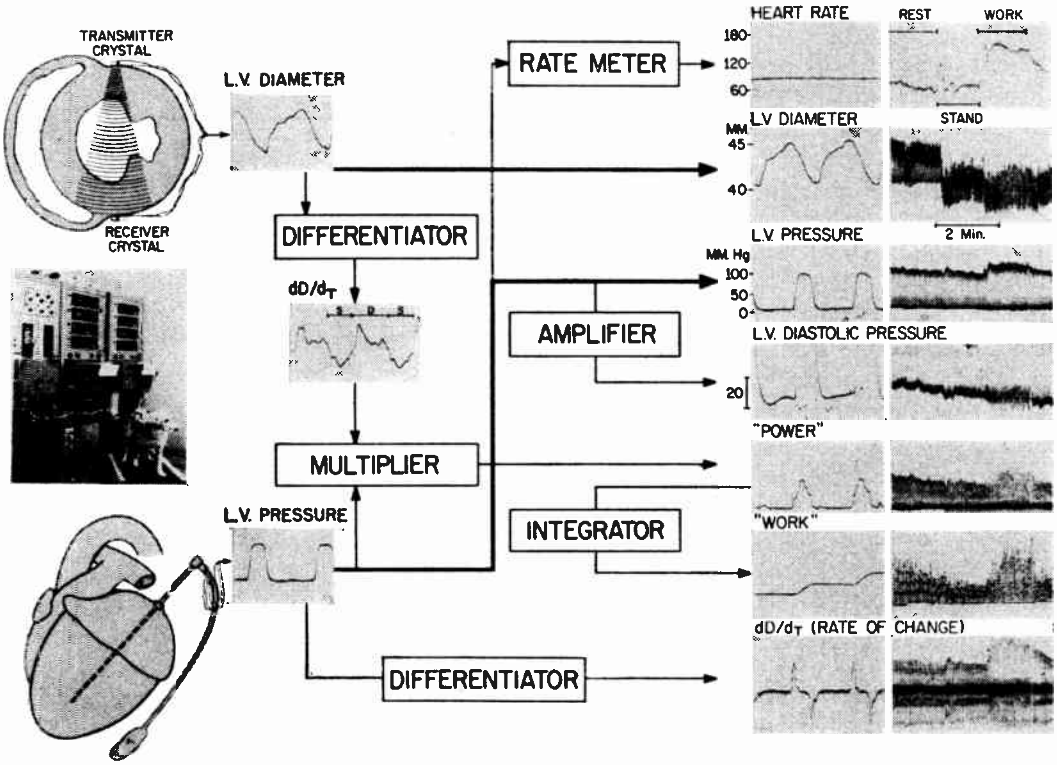


Fig. 7—The analog computer arrangement needed to compute the effective power of the heart wall. Power is defined as the rate of change of ventricle diameter multiplied by the ventricle pressure. Each block consists of an operational amplifier with the proper feedback to perform a given function.

is an ac voltage which is capacitively coupled to the thermistor. If the power fed to the thermistor is continuously monitored, a signal which is proportional to flow is developed.

Nuclear Magnetic Resonance Flow Meter

Nuclear magnetic resonance has been used experimentally to measure blood flow. P. Buchman,[17] devised a technique to accomplish this while working on a master's thesis in electrical engineering. When protons, whose spin axis has been properly aligned, are passed through a varying magnetic field, the amount of energy required to resonate and flip them over is a measure of the number passing per unit time. The energy absorbed is a measure of the volume flow.

The development of this meter was undertaken as a feasibility study. Initial evaluation indicates that it can be used in certain blood flow measurements.

Data Processing and Storage

Analog Computations

Analog computers are a powerful tool in the derivation of new concepts of the heart as a pump. Data representing the fundamental parameters of dimensions, pressure and flow have been reduced to provide such information as heart rate, pressure-volume loops, and functions of power, work, and accumulated flow.[1,14,18]

[18] R. F. Rushmer, O. Smith, and D. Franklin, "Mechanisms of cardiac control in exercise," *Circ. Res.*, vol. 7, pp. 602–627; July, 1959.

The computation of rate, power, and work is shown systematically in Fig. 7. The integrating, differentiating, and multiplying circuits are built up around standard amplifier units. An indication of the effective power generated by the heart walls is derived by multiplying the left ventricular pressure by the rate of change of the left ventricular diameter or circumference. If the diameter or circumference can be considered continuously proportional to the volume we then have an uncalibrated record of the power output of the heart. Work is then the integral of power. These functions are recorded as continuous information along with the primary parameters. The output of blood per stroke and accumulated flow per unit time (*i.e.*, two seconds) can be derived from the instantaneous flow information by integration.

Tape Recorder

An extremely valuable tool used in conjunction with the computer is a four-channel FM tape recorder. Besides providing storage for valuable data, the tape recorder permits the investigator to concentrate on recording the basic parameters.

Then, at a later date, all the derived functions can be computed from the tape. This method of operation reduces the instrumentation required for a particular experiment and permits more extensive data reduction and analysis.

Rate Meter

The measurement of heart rate at frequencies as low as 40 beats per minute posed a problem. Here, again, the problem was one of providing a dc output voltage proportional to the variable measured.

The repetition rate of the heart is equal to the reciprocal of the period between beats. The sharp rising pressure pulses, occurring with each heart beat, trigger a circuit for measuring the period. This trigger controls a ramp function (bootstrap circuit) which is allowed to rise during the interval between pressure pulses. The ramp generator is reset and immediately restarted at the leading edge of each pressure pulse. The peak voltage, which the ramp function attained, is proportional to the length of the period. This peak voltage is then stored in a capacitor, Fig. 8. A phantastron divider circuit develops the reciprocal of this voltage. The output of this circuit is then calibrated in terms of heart rate. Tests have shown that the instrument is linear from 30 ppm to 250 ppm. It responds to changes in cycle length which occur from one beat to the next. This feature is important because of the random cycle length of the heart beats. The rate meter can be used to analyze taped data, either at the time of an experiment or later.

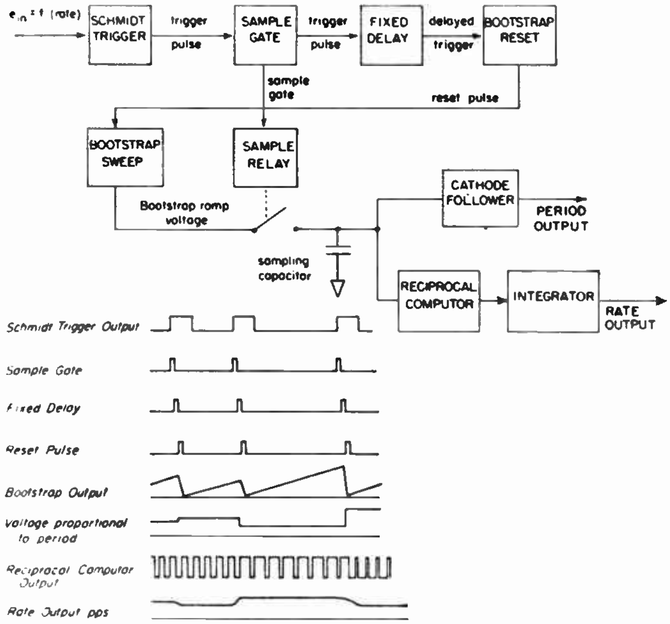


Fig. 8—Top, block diagram of meter to register heart rate by measuring the period between beats. During interval between beats a linear voltage rise is generated by charging a capacitor at a constant value (bootstrap sweep). Rate is computed by taking the reciprocal of the peak voltage attained at end of each linear rise. The reciprocal computer is a circuit whose output pulse width is proportional to the reciprocal of period. When these pulses are integrated, a function of rate is formed. Bottom, waveforms obtained.

A System Concept

Information from a complicated experiment can be obtained by several technicians hovering over a few independent instruments. However, this method of operation can never enjoy the flexibility and potential of a well-planned, fully integrated system. This has been the engineering philosophy in our laboratory. Our ultimate goal has been to develop a centralized instrumentation system which is highly compatible with the physiological environment, and which can be operated by medical investigators with minimal engineering background.

Stability, Oscillations, and Noise in the Human Pupil Servomechanism*

LAWRENCE STARK, M.D.†

Summary—**The pupil reflex to light has been considered as a servomechanism, a self-regulated error-actuated control device. This cybernetic approach, requiring the experimenter to make quantitative measurements in animals with a fully intact central nervous system, was made possible using a pupillometer designed for awake, cooperative human subjects. This instrument provided an electronically controlled light stimulus as well as continuous records of both pupil area and light intensity. Sinusoidal changes in light intensity, small enough for linearization assumptions, were injected in an open loop fashion to determine the transfer function for pupil system behavior. The pupil servo is quite stable and has a low gain with an attenuation slope of 18 db per octave beyond 1.5 cps. One line of investigation using pharmacological agents has suggested the triple lag to be contributed by the physical law representing the viscosity of the iris neuromuscular system. Another experiment used artificially increased gain to produce instability oscillations whose frequency was predictable from the low gain transfer function. Still another investigation has shown the pupil system to contain much noise. This noise is not a result of instability, nor generated by the smooth muscle of the iris, nor by other elements of the pupil servoloop, but is injected into the loop from another part of the brain. Further studies in progress are defining nonlinearities in the pupil and retinal system in order to set up an accurate analog model of the pupil system in the form of a program for a digital computer. The general manner in which pupil behavior is shaped for various operating ranges and frequencies throws light on the power of an adaptive servomechanism to maximize its utility to the organism. The value of the cybernetic approach is demonstrated by both the clarification which these concepts, derived from control and communications engineering, have introduced into the understanding of pupil behavior as well as by the precision of the experimental data obtained using measurement techniques adapted for physiological purposes.**

Introduction

THE PUPIL of the eye acts as a regulator of light impinging upon the retina. The transfer function and noise characteristics of this stable type zero servomechanism (Fig. 1) will be presented in this paper. The normal behavior of the pupil system can be modified by changing the experimental conditions. Then such interesting phenomena as instability and oscillations can be demonstrated.

The pupil was chosen for study from a host of possible examples of biological servosystems for several reasons.[1,2] First, its motor mechanism, the iris, lies exposed behind the transparent cornea for possible measurement without prior dissection. This had previously been exploited for scientific and clinical researches by using high-speed motion picture cameras. Further, by employing invisible infrared photographic techniques, measurements can be made without disturbing the system, because its sensitivity is limited (by definition) to the visible spectrum. Second, the system can be disturbed or driven by changes in intensity of visible light, a form of energy fairly easy to control, and painless in its administration to the subject. The first two advantages lead to still a third: the possibility of performing experiments on awake, unanesthetized animals whose nervous system is fully intact and functional. In fact, all of the experiments to be discussed below have been performed on human subjects. Lastly, the system responds with a movement having only one degree of freedom, a change in pupil area, which simplifies the system equation analysis.

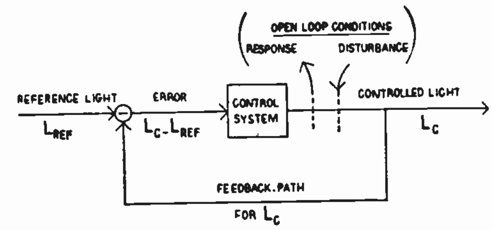


Fig. 1—A simple servosystem. This shows forward and feedback paths in the servoloop and different components therein. Symbols are explained in diagram and text. Dashed lines indicate where loop might be opened and a disturbance be introduced, and response around the loop measured.

The pupil is so widely observed an organ that most persons are already acquainted with certain basic facts of its anatomy and physiology. The pupil is the hole in the center of the iris muscle which enables light to enter the eye and impinge upon the retina, the sensitive layer of the back of the eye. The retina is comprised of primary sense cells containing photosensitive pigments which trap photons and subsequently stimulate nerve cells. The retina is part of the central nervous system and possesses a complex multineural integrative (*i.e.*, information transforming) apparatus. The optic nerve leads mainly to the visual cortex of the cerebral hemispheres via a relaying station, the lateral geniculate body. However, some fibers, called the pupillomotor fibers, go directly to the brain stem and relay in the pretectal area and thence to the Edinger-Westphal nucleus. This nucleus contains the nerve cells, part of the parasympathetic system, whose fibers (after an external relay in the ciliary ganglion) control the powerful sphincter

* Original manuscript received by the IRE, July 23, 1959. Aided by grants from the Natl. Science Foundation and the Natl. Inst. of Neurological Diseases and Blindness. In addition to colleagues and students who are co-authors of papers referenced here, the author wishes to thank Profs. G. H. Glaser and P. N. Schultheiss for their advice and encouragement.
† Asst. Prof. of Neurology, Yale University School of Medicine, New Haven, Conn.
[1] W. S. McCulloch, "Cerebral Mechanisms in Behavior—The Hixon Symposium," John Wiley and Sons, Inc., New York, N. Y.; 1951. See "Why the mind is in the head," pp. 42–57.
[2] N. Weiner, "Cybernetics," John Wiley and Sons, New York, N. Y., 149 pp.; 1948.

muscle of the iris. Fiber tracts also go to the sympathetic system in the spinal cord. Here, nerve cells send fibers back to the orbit, after relaying in the superior cervical ganglion. The dilator of the pupil is controlled by these sympathetic fibers and is responsible for the wide dilatation of the pupils after the administration of adrenalin.

Excitation of the Edinger-Westphal nucleus produces constriction of the pupil and it is also probable that inhibition, *i.e.*, decrease in the operating level, of this nucleus is also the most important mechanism for dilating the pupil.

Any further relevant and necessary details of neuroanatomy will be discussed in the main body of the paper. The reader is assumed to possess a knowledge of linear servomechanism as presented in a college text such as Schultheiss and Bower.[3]

Experimental Methods

In order to obtain careful, quantitative data from the human pupil system under a variety of experimental conditions we felt it would be impossible to use the older infrared photographic techniques,[4,5] and developed a simpler modification of this technique. In Fig. 2 the essential nature of the experimental arrangement of our pupillometer is shown.

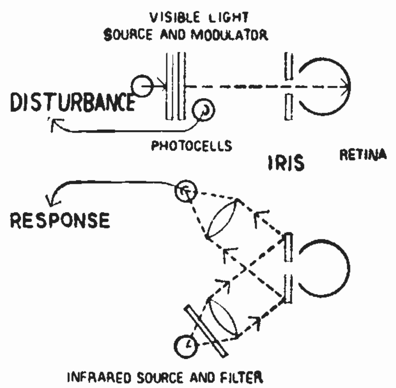


Fig. 2—Experimental arrangement. Optical portions of visible light stimulus path are not shown. Modulator consisted of polaroid filters in rotary oscillation with respect to each other. A fixation point was provided, as well as a biteboard.

The pupil area was measured continuously by reflecting infrared light from the iris onto a photocell. The pupil is ordinarily black because most of the light passing through the pupil into the eye is absorbed by pigment layers behind the retina. Thus, when the pupil is large (and the iris small) less light is reflected from the front of the eye onto the photocell. When the pupil is small (and the iris large) more infrared light is reflected onto the photocell. In this manner we obtained a convenient and continuous measurement of the system response. At first an attempt was made to use specular reflection from the iris in order to improve linearity of the response, but retinal specular reflection was also obtained and this was not negligible. Therefore the use of scattered light reflection from the iris onto the infrared sensitive photocell was an important part of the experimental instrument design. Another way of increasing the signal-to- noise ratio was by the use of a relatively small area of infrared illumination of the iris. This was arranged to be only slightly larger in diameter than the largest diameter of the pupil. The use of a narrow spectral band of infrared light shaped to the infrared spectral sensitivity of the photocell also aided in this experimental approach. The elimination of the long wavelength infrared meant that most of the heat and discomfort to the subject was removed. The output of the photocell, a vacuum-type no. 917, was fed directly into the high impedance input of the recording amplifier. There was a small capacitance across it in order to remove high-frequency noise, and the tube and leads were shielded. The photocell housing could be shifted for studying the right or left eye. The infrared sensitive photocell was shielded with an infrared Wratten filter to eliminate the effects of stray visible light. Whenever beam splitters were placed in the path of the infrared light these were made dichroic to minimize attenuation of infrared light. The infrared light source was a 35-mm slide projector with a built-in fan for cooling, and it was supplied from a constant voltage transformer to obtain stability of light intensity.

In order to translate the photocell currents into pupil area measurement it was necessary to calibrate the instrument. Such a calibration is shown in Fig. 3. This was obtained by taking flash photographs of the pupil and at the same time noting the amount of photocell current. The flash of light naturally produced a pupillary response in the subject but the photographic measurements were over and completed before the pupil had a chance to respond. There has been found a fairly proportional relationship between pupil area and photocell current, but in the figure one can see that there is a diversion from linearity. However, this is small and not significant within the range of most of the experiments, and was not corrected for. The calibration camera was a permanent part of the apparatus and was a single lens reflex camera with a built-in viewfinder. A dichroic beam splitter was placed so that most of the visible light reflecting from the iris was transmitted to the camera while interfering relatively little with the measuring infrared light. The electronic photoflash light was heavily filtered to permit only blue light to illuminate the iris, a practice which markedly reduced intensity and discomfort for the subject, as well as the effect on the infrared measuring photocell. An event marker automatically indicated on the recording graph the instant the photograph was taken. The camera view-

[3] P. W. Schultheiss and J. L. Bower, "An Introduction to the Design of Servomechanisms," John Wiley and Sons, Inc., New York, N. Y., 500 pp.; 1958.

[4] O. L. Lowenstein, "Pupillary reflex shapes and topical clinical diagnosis," *Neurology*, vol. 5, pp. 631–644; 1955.

[5] S. A. Talbot. "Pupillography and Pupillary Transient," Ph.D. dissertation, Dept. of Physics, Harvard University, Cambridge, Mass.; January, 1938.

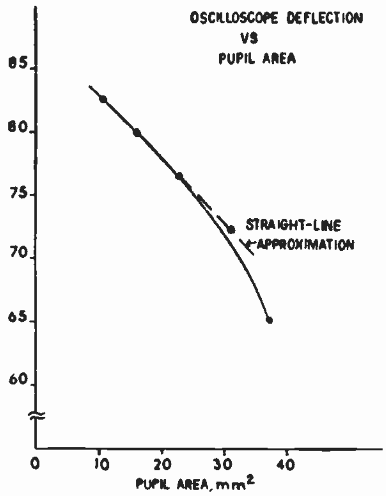


Fig. 3—Response calibration. Response photocell current is shown in arbitrary units. Pupil area was determined from diameter measurements, from enlarged photographs. Each point is average of two or three flash photographs taken at same light intensity.

finder had excellent split-image focusing so that the subject could be accurately determined to be in the focal plane of the camera. Further details concerning the positioning of the subject's head and eyes will be given in a later portion of this section. Each individual that was studied in these series of experiments was separately calibrated because of the differences in infrared reflectivity of various subjects' irises. It was noted that subjects with brown eyes generally gave more infrared reflection at a stated pupil diameter than subjects with blue eyes.

The stimulus light had to be controlled in a quantitative manner. Originally, rotating polaroid films were used to obtain sinusoidal intensity changes. However, since the original work was reported, an electronically controlled light source has been developed to provide a flexible stimulus system in which light intensity can be varied sinusoidally or in a stepwise fashion, the dc level may be adjusted, and other experimental arrangements which will be discussed later are made possible. The central features of the stimulating light system are shown in Figs. 4 and 5. Current passing through a glow modulator lamp (Sylvania 1131C) changes the light intensity. In order to linearize the power amplifier and glow modulator lamp a vacuum photocell (no. 929 with a wide linear range) was inserted into the system to sample a portion of the visible light stimulus. The output of this monitoring photocell was fed back into the first stage of the high-gain amplifier. Thus, the high-gain amplifier and the power amplifier, controlling current through the glow modulator, were made to follow the output of the monitoring photocell. Other aspects of this circuit included a filter to reduce high-frequency response and keep the amplifier stable and the use of limiters to protect the glow modulator tube. Various inputs were used to drive the light source. They included a low-frequency sinusoidal oscillator, step generator, and dc level changes.

It was also possible, since the measuring system provided a voltage proportional to the pupil area, to feed back this pupil area voltage to control light intensity. For example, if the pupil became larger we could use the pupil area signal to decrease the light intensity. In this way we could oppose the effect of the pupil system and so open the loop. Again, the polarity of the signal representing the pupil area could be changed and then fed back to the light control amplifier. In this way a higher gain system could be obtained. For example, if the pupil area became smaller the light intensity would be turned down. In this way the pupil area change would have a much greater effect than it would have from consideration of the optical geometry. It was possible to control quantitatively the amount of pupil area feedback to the light control system. Increasing the gain, as shown later, makes the entire system unstable and oscillations of the pupil system are obtained.

The optical arrangement of the stimulus light was most important. Several arrangements of the stimulating light are shown in Fig. 6. Under normal closed-loop operating conditions light was diffuse so that changes in the pupil area could affect the amount of retinal flux. Provision was made, however, for careful focusing of the stimulus light so that the entire light beam entered the pupil in the form of an image of a small disk whose diameter was smaller than the smallest diameter of the pupil. Changes in pupil size could then have no effect on flux on the retina. This arrangement of stimulus light, shown in Fig. 6(b), produced the open-loop operating condition by removing the influence of system response over stimulus. The pupillometer can be arranged for a further experimental operating condition shown in Fig. 6(c). Here the small focused disk of light used in the open-loop operating condition can be situated on the edge of the pupil. In this experiment it is found that a narrow bar of light from an ophthalmological slit lamp was also able to be used and provision was made for supplying this type of stimulus also. Under either of these conditions very small movements of the iris muscle will cause a very large change in retinal illumination. Thus, the gain of the pupil system has been greatly increased and the pupil develops oscillations.

The pupillometer measures continuously the amount of light impinging upon the eye of the subject. In this manner, a continuous pen recording of both stimulus and response is provided. The pen recorder is able to be modified so that four channels are in use. The extra two channels exhibit the response of the system in a modified way. For example, if the gain of the amplifying system is turned up, then small fluctuations of the pupil can be seen. This high-gain measuring system, however, may be compared with the lower gain continuous record of the regular response recording channel. An additional channel might display the response after passing through a narrow band-pass filter, so that the fundamental response is emphasized.

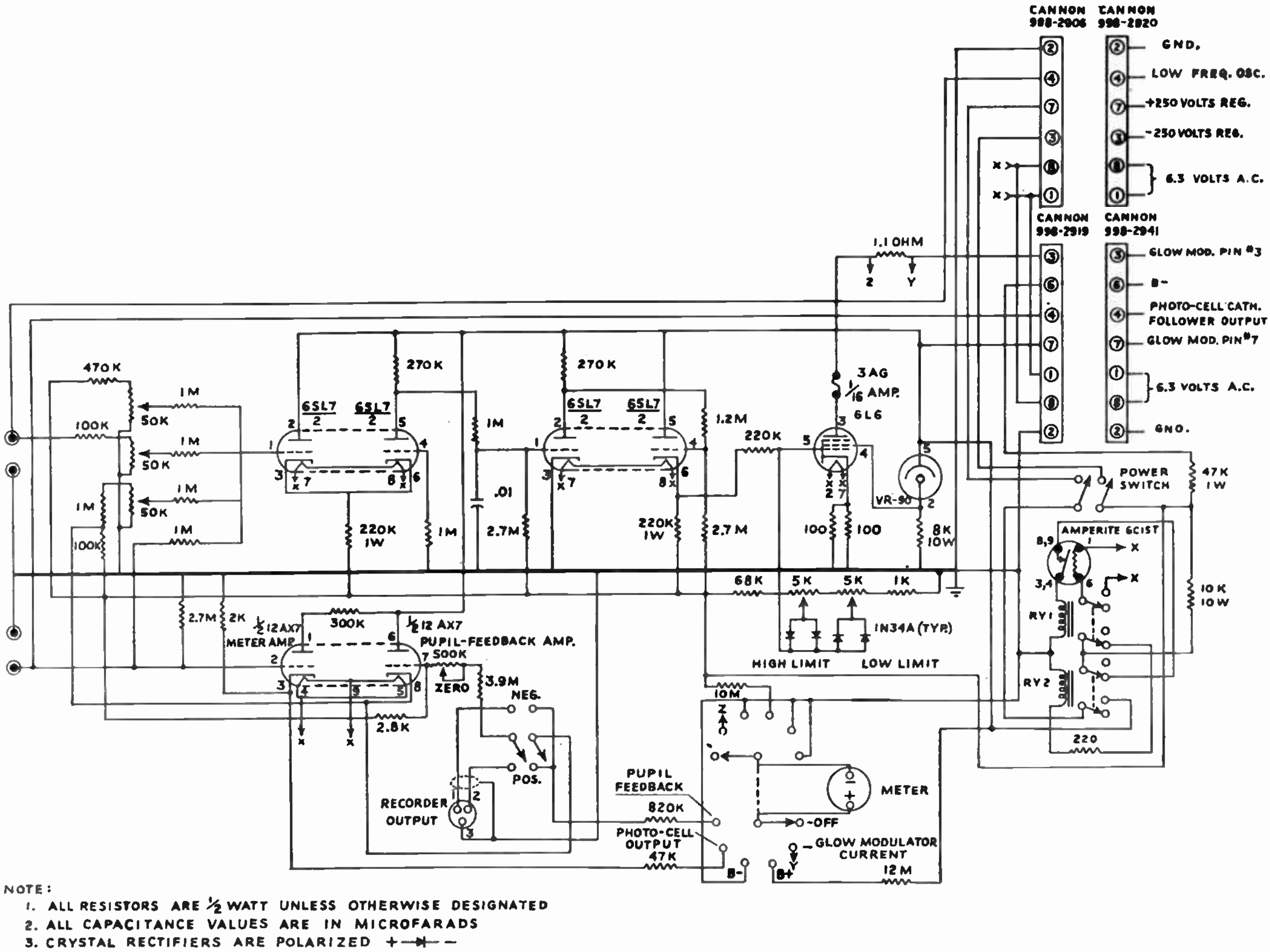


Fig. 4—Circuit diagram of light control unit, a servoamplifier with feedback from output of monitoring photocell which linearizes power amplifier and glow modulator tube. Sinusoidal, dc, and pupil-area-feedback voltage signals are mixed in the first stage of amplifier.

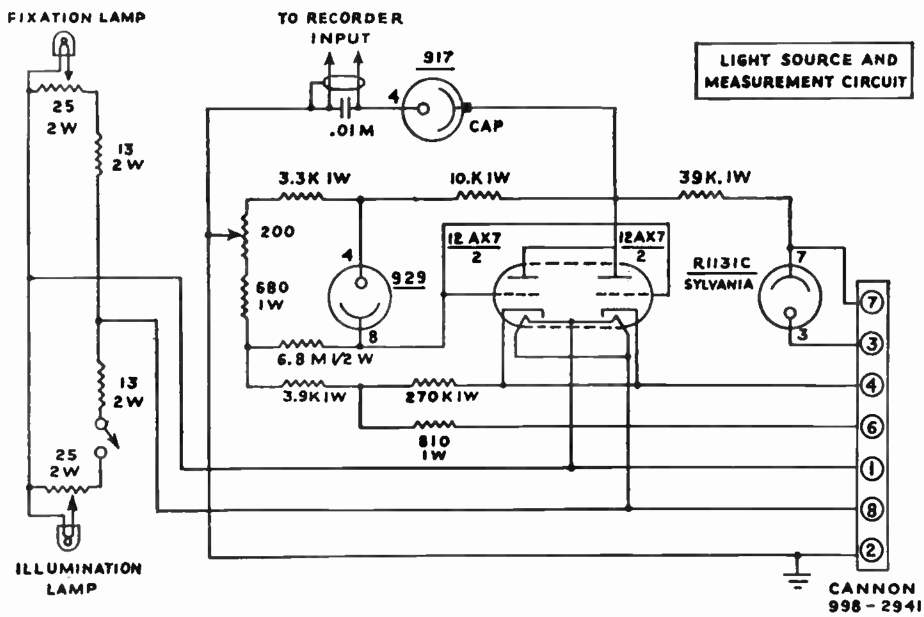


Fig. 5—Circuit diagram of glow modulator monitoring photocell and cathode follower unit as well as accessory illumination light and fixation light supplies.

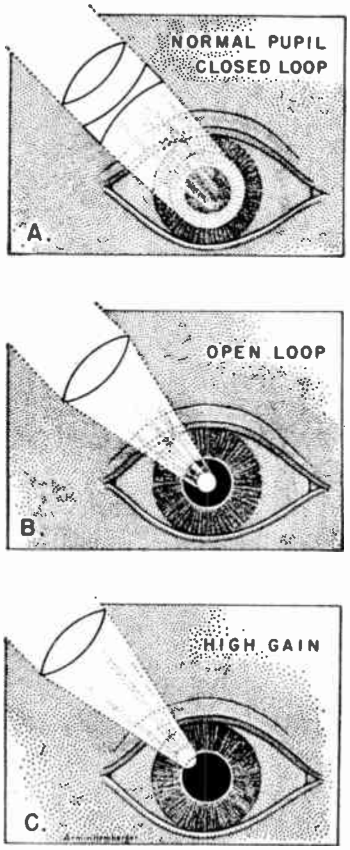


Fig. 6—Three techniques used for stimulation. (a) Normal condition in which movement of iris changes intensity at retina. (b) Technique generally used for obtaining amplitude-phase frequency data. Light entering center of pupil is unaffected by movement of iris. Maxwellian View. (c) Light is here focussed on borders of iris of pupil. Small movements of iris result in large changes in light intensity at retina.

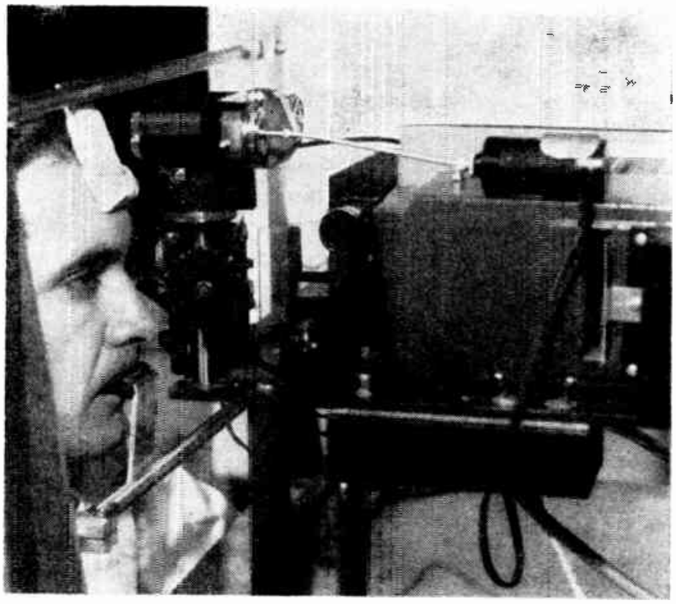


Fig. 7—Physical arrangement of pupillometer, subject, and bite-board. Subject's head is supported and with a proper fixation point the eye is held quite still. Further details of stimulation and response system are described in text.

In order to obtain accurate measurements of a human subject's pupillary area it is of course necessary that the subject remain absolutely still. In order to enable a cooperative subject to do this two varieties of bite boards were prepared. For most of our careful quantitative experiments on normal subjects, individual tooth impressions made of dental wax were placed on a metal angle and clamped to the pupillometer, as shown in Fig. 7. Another method used was to split and glue a rubberized bite ring onto the metal bite board. In this way the subject did not need to have a personal bite board made and this method was widely used for studying neurological patients. Both of these arrangements were found to be quite satisfactory in preventing head movements as the weight of the subject's head rests on his upper teeth, and there is a relative fixity in contact between the subject's head and the apparatus. There is no pressure against the soft parts of the head such as obtains in a chin rest, and contraction of the jaw musculature does not shift the head. In order to keep the subject's eye still, a fixation point was provided with a special light and placed at an infinite distance (optically) from the subject. In this way the subject was able to fixate comfortably on a small spot and relatively little eye movement appeared. The amount of residual eye and head movement could be determined by observing changes in recorded area under experimental arrangements wherein no changes in pupil area would be expected. Under these conditions the base line remains extremely stable. Trained subjects were able to remain in the apparatus for several hours during which time repeatable measurements could be obtained. Untrained subjects such as patients from the Neurology Clinic could also be adapted to the instrument since there was little or no discomfort in the procedure. Removing the heat-radiating long infrared spectrum from the infrared lamp was an important part of the experimental design in this respect. Another detail found to be helpful was the arrangement for quickly changing the configuration used as the fixated point. In this way the subject's attention and his eye could be again brought to focus on the fixation point.

The viewfinder of the calibration camera was used together with a special illumination light which lit up the subject's eye, in order to position the subject carefully and to check on the subject's condition again and again throughout the course of the experiment. Millimeter lines were ruled on the viewfinder and a good calibration could be obtained by estimating the size of the subject's pupil and comparing this with photocell output at that instant. As the apparatus has developed toward more stable operation we have gradually shifted from continuous to stepwise controls. In this way a trained technician can run a careful experiment in a short time on a clinic patient.

Transfer Function Analysis

Initial experiments were performed using open-loop injection of small sinusoidal signal stimuli to obtain a linear transfer function of the pupil system.[6] Studying a system under open-loop operating conditions is an important dissection technique widely used by the servo-engineer. This method is indicated in Figs. 1 and 8 by dashed lines representing a break in the servo loop.

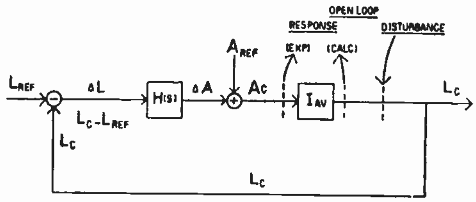


Fig. 8—Linearized approximation to pupil servosystem. Actual pupil system is more complicated than this figure indicates. However, we linearized by using small variations about a fixed operating point. Thus the necessary calculations are simplified and linear servoanalytic methods may be applied. A_{REF} is reference area. A_C is controlled area. ΔA is a change in area generated by control system whose transfer function is $H(s)$, I_{AV} is average intensity value used to multiply controlled area to yield controlled light flux. This approximation is explained in text.

A disturbance is injected, transmitted around the loop, and measured at the point of break. Thus, system response has no influence over the disturbance, which remains entirely in the control of the experimenter. The input-output relation, that is, the transfer function, is simplified and for this reason an essential part of this experimental approach to the pupil reflex was the development of the method for open-loop operating condition discussed in the preceding section describing the pupillometer. There were three reasons for our use of sinusoidal stimuli: experimental techniques for obtaining given level of accuracy are simple, the mathematical analysis is well understood and relatively easy to manipulate, and system design and performance are evaluated readily. As an example of the first point, once the retina had adapted to mean light intensity, we were able to vary sinusoidal modulations over the entire frequency range while the pupil system remained in a steady state. The small signal approach enables us to have a linearized approximation to the pupil servosystem. The actual pupil system is much more complicated than Fig. 8 indicates and this will be discussed in a later section. However, in attempting to obtain a value for absolute gain it is necessary to utilize this small signal approach. The pupil system is also very sensitive to operating levels. Thus, by using a small signal the operating level remains relatively constant, and the response of the pupil system can be made quite reproduceable. Furthermore, there are several nonlinear operators in the pupil system; for example, the log transducer which represents the Weber-Fechner Law, and the actual geometrical multiplication of pupil area by light intensity to obtain retinal flux. By using small signals we were able to ignore the logarithmic operator, and also to linearize and so substitute a subtraction operator for the multiplication operator.

At this point clarification of the concept of gain and the operational definition given it in our experiments is necessary. Open-loop gain is the ratio of the magnitude of the response of the servomechanism in open-loop operating condition to a signal injected into it, providing that the signal has completely traversed the loop up to the point of injection. It is of course a dimensionless quantity since the injected signal and the response measured at the same point must be in dimensionally similar terms. It is irrelevant where the break in the loop is, as a fundamental theorem in linear servoanalysis proves the identity of all over-all loop transfer functions.

The break point in the pupil loop is just before the injection of the disturbance (see Fig. 22). The disturbance is a change in retinal flux caused by a change in external intensity. The change of light intensity must be multiplied by the area of the open-loop disk of light (provided that the intensity is measured at this point).

$$F_e = A_{AV} \cdot \Delta I. \quad (1)$$

The response is the change of flux due to the pupil area change which would have occurred if the pupil area had been allowed to operate on the light entering through the pupil. Thus the change in area is multiplied by the average intensity considered as if it were evenly distributed over the entire pupil area.

$$F_i = I_{AV} \cdot \Delta A. \quad (2)$$

This assumes the retina cannot distinguish between various light intensity distributions at the plane of the pupil; reasonably valid for a Maxwellian view. Reference to Fig. 6, showing the open-loop light distribution, and to Figs. 1, 8, 21, and 22, showing the open-loop break position, will help to clarify this argument. The small signal approximation is needed for the gain calculation

$$G(s) = \frac{F_i}{F_e} = \frac{I_{AV} \cdot \Delta A}{A_{AV} \cdot \Delta I} = \frac{\Delta A / A_{AV}}{\Delta I / I_{AV}}. \quad (3)$$

The numerator (response of the system) and the denominator (injected signal) are both in dimensions of flux; thus a dimensionless gain is obtained.

A sample experiment is shown in Fig. 9. The two oscillatory traces represent oscillations in light intensity of the stimulus and oscillations in the pupil area of the response. It will be noted that the response curve contains noise, and a discussion of this will form a later section of this paper. There is also some harmonic distortion noticeable and this again is evidence of the non-

[6] L. Stark and P. M. Sherman, "A servoanalytic study of the consensual pupil reflex to light," *J. Neurophysiol.*, vol. 20, pp. 17–26; 1957.

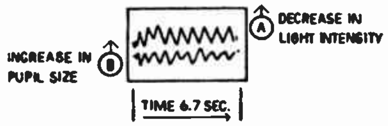


Fig. 9—Typical experimental data. (a) Stimulus intensity varying sinusoidally at a frequency of 1.3 cps. Upward deflection indicates decrease in light intensity. (b) Pupil area changing at the same frequency but approximately 180° behind stimulus. Upward deflection indicates increase in pupil size. Note variability in response and harmonic distortion. Time proceeds toward right.

linearity of the pupil system and introduces an error of up to 10 per cent in our phase shift measurement.[7]

The basic data obtained in the analysis of this simple experiment are the relative amplitude of input and output and the phase shift between input and output. These two values are obtained for each of a number of frequencies of intensity modulation. In summary, the experimental procedure was designed to adapt servo-analytic methods to study the pupil reflex. Small sinusoidal light stimuli of varying frequency were applied. The sinusoidal response was measured and its amplitude and phase relationship to the stimulus was determined. Data were obtained in open-loop and closed-loop operating condition. A number of frequencies could be studied quickly using our pupillometer on a subject in one sitting.

From these amplitude and phase data several displays of system behavior can be obtained. The first is the Bode plot in Fig. 10. From this graph one can readily see certain qualitative features of the pupil system: low gain, steep attenuation of gain at high frequencies, and a large phase shift. Another method of displaying system behavior used the Nyquist diagram shown in Fig. 11. This vector plot of gain and phase angle is often employed to show clearly the desired characteristics of a servomechanism: stability of operation with adequate speed and accuracy of response. Stability means that the system must not oscillate excessively as it attempts to correct for error. Other characteristics of interest are degree of stability and range of frequency over which the system will respond. Enclosure of the critical point at 180° phase shift and unity amplitude by the Nyquist curve indicates the system instability and predicts sustained or divergent oscillation in response to any disturbance. The fact that the curve in Fig. 11 does not so enclose the critical point indicates that the system is stable. Furthermore, the degree of stability of the system can be determined by the distance of the curve to the critical point. Since the gain is 0.12 at 180° phase shift the pupil system is clearly very stable.

In addition to the above qualitative and semi-quantitative discussion of the pupil system it is also desirable to have a full but concise mathematical de-

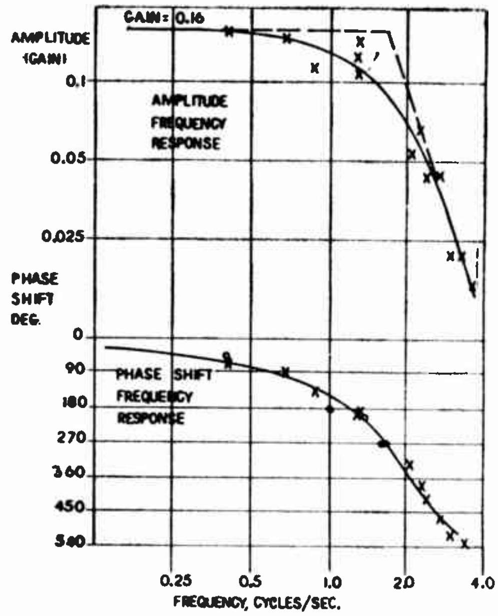


Fig. 10—Open loop frequency response. Amplitude is plotted on log-log scale while phase shift (lag) is on log-linear scale. Points are experimental and continuous lines are fitted. Dashed lines are asymptotes.

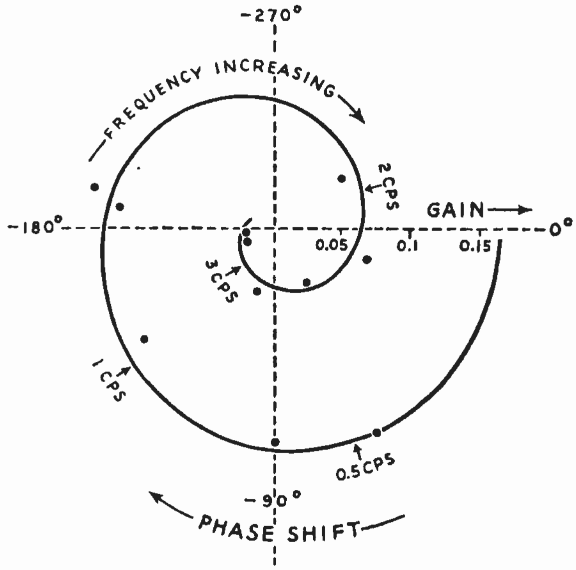


Fig. 11—Nyquist diagram of pupil response, a vector plot of gain and phase shift. Scale of modulus and a few frequencies are indicated. Curve is derived from fitted lines from gain and phase frequency-response graphs, while points are experimental.

scription of system behavior. Such a canonical expression is the transfer function, for convenience written as a function of the "complex frequency" operator s. In Fig. 8 loop elements of the transfer function, $G(s)$, are shown. Included are $H(s)$, the transfer function relating area change output to light change input, and the linearized intensity multiplier, I_{AV}. $G(s)$ is independent of the actual break point, provided only that the response has traveled completely around the loop to the point of injected disturbance. The data displayed in the Bode diagram will now be used to derive the open-

[7] J. Stegeman, "On the influence of sinusoidal variations in brightness upon pupillary size," *Pflügers Arch. ges. Physiol.*, vol. 264, pp. 113–122; 1957

loop transfer function, $G(s)$. Low frequency gain is 0.16. The attenuation curve appears to have an asymptotic slope of 18 db per octave beyond the break frequency. This slope can be represented by three time lag factors. The actual values of the time constants are hard to determine exactly from the experimental data, but as a rough approximation we set each equal to 0.1 second. These time lags account for a portion of the phase shift at higher frequencies. By referring to the phase shift of the Bode diagram it can be seen that the actual phase shift is 540° at 4 cps. The minimum phase shift discussed above accounts for only 270° of this. The remaining 270° can be attributed to a nonminimum phase shift, equivalent to a time delay of 0.18 second, expressed as $\exp(-0.18s)$.

These calculations now may enable us to write the transfer function.

$$G(s) = \frac{0.16e^{-0.18s}}{(1 + 0.1s)^3} \cdot \tag{4}$$

Several parameters in the open-loop transfer function can be compared with data from closed-loop and step experiments. The closed-loop transfer function is related to the open-loop transfer function by

$$F(s) = \frac{G(s)}{1 + G(s)} \cdot \tag{5}$$

Low-frequency closed-loop gain is calculated to be 0.14. Closed-loop gain was experimentally determined at several operating points as 0.15. Time delay was measured in step function experiments and found to be 0.18 second. The agreement shown here is satisfactory, especially considering our difficulty in achieving experimentally a good small signal approximation (10 per cent modulation in open-loop experiments). Further experiments have shown that both the consensual and direct pupil response have a similar transfer function.

High-Gain Oscillations and Stability

The feedback pathway which characterizes servomechanisms necessitates a system analysis in order to elucidate its physiology. This means that there are certain properties such as stability and oscillations which cannot be attributed to any individual component of this system but are properties of the entire system. In the particular example to which this section is devoted, the nature of "induced pupillary hippus" becomes comprehensible when considered as sustained oscillations in the pupil servosystem experimentally made unstable by greatly increasing the gain of the loop. Further, the nature of this servo approach permits the design of quantitative experiments to test predictions based on these concepts. The experiments to be described below are of this type. Stability is a fundamental property of a servomechanism which is closely related to other important characteristics such as gain and phase lag. Indeed, servo-engineers devote much effort toward eliminating instability in the design of useful servosystems, even paying for stability by reducing the size and speed of the response. The characteristic fault or failure of an unstable servomechanism is exhibited as sustained or increasing oscillations. The pupil servoloop, as mentioned in the preceding section, is stable because of its low gain despite large phase lag. However, it is possible to increase the gain in the pupil system experimentally. When this is done the instability of the modified pupil loop shows itself as sustained oscillation with a frequency in accordance with the prediction of linear servoanalytic theory.[8,9] In this section these instability oscillations are studied and their general nature elucidated. Further, the pupil system is changed by means of drugs and it is shown how the oscillations are affected in a manner predicted by alterations in the normal low-gain pupil system.[10] The accuracy of these predictions is further evidence showing the validity of the application of linear servoanalytic theory to this neurological servomechanism. The methods used in the experiment to be described are illustrated in Fig. 6 which shows the high-gain operating condition. The operation of shining light on the edge of the pupil must have been performed many times by ophthalmologists studying the anterior chamber of the eye by means of a slit lamp, and several published descriptions and studies are available.[11–13] The gain of the pupil system was increased and the resultant oscillations were termed by these earlier studies "induced pupillary hippus." Fig. 12 shows an example of spontaneous oscillation in pupil area obtained under these conditions. It should be carefully noted that light intensity was kept constant throughout this experimental condition as shown in the figure. Thus this oscillation is very different from the driven oscillations whereby the frequency response data are obtained. It was generally always possible to produce oscillations in normal individuals although in some cases an adjustment of light intensity had to be made. The frequency of the high-gain oscillations were determined by a straightforward method of averaging several waves, and the vertical line in Fig. 13 marks the frequency of this oscillation on the phase frequency diagram. It can be noted that this is about the frequency at which 180° of phase lag was found in the normal low-gain pupil system.

[8] L. Stark, "Oscillations of a neurological servomechanism predicted by the Nyquist stability criterion," in "Selected Papers in Biophysics," Yale University Press, New Haven Conn.; 1959 (in press).

[9] L. Stark and T. N. Cornsweet, "Testing a servoanalytic hypothesis for pupil oscillations," *Science*, vol. 127, p. 588; 1958.

[10] L. Stark and F. Baker, "Stability and oscillations in a neurological servomechanism," *J. Neurophysiol.*, vol. 22, pp. 156–164; 1959.

[11] H. J. Stern, "Simple method for early diagnosis of abnormalities of pupillary reaction," *Brit. J. Ophthal.*, vol. 28, pp. 278–276; 1944.

[12] F. W. Campbell and T. C. D. Whiteside, "Induced pupillary oscillations," *Brit. J. Ophthal.*, vol. 34, pp. 180–189; 1950.

[13] K. C. Wybar, "Ocular manifestations of disseminated sclerosis," *Proc. Roy. Soc. Med.*, vol. 45, pp. 315–320; May, 1952.

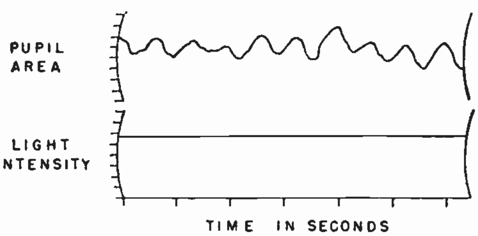


Fig. 12—Example of spontaneous oscillations in pupil area obtained with constant light stimulus using high gain operating condition illustrated in Fig. 6(c). The frequency in cycles per second of several such runs determined and averaged to obtain high-gain oscillation data displayed in Fig. 13.

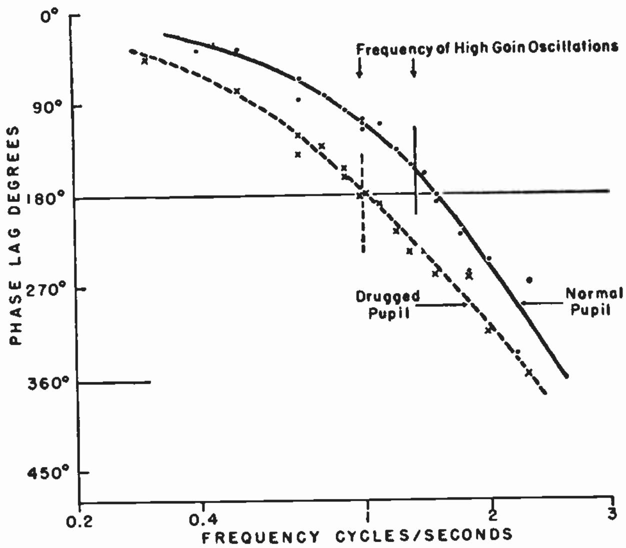


Fig. 13—Data from experiment F.B.200 using drugs to shift system parameters. Solid curve represents phase-frequency data from normal pupil low-gain operating condition. Solid vertical line marks frequency of spontaneous oscillations in high-gain operating condition which corresponds to 180° phase crossover frequency. Dashed curve represents phase-frequency data from drugged pupil in low-gain operating condition. Dashed vertical line marks frequency of spontaneous oscillations of drugged pupil in high-gain operating conditions again showing correspondence with the new 180° phase crossover frequency.

Although the association of these two frequencies is quite close it was determined that an effort should be made to alter the pupil system and measure the two frequencies in the altered pupil system. It was felt that this task of being able to shift the two frequencies in a parallel manner would be a more stringent set of conditions which should be met if one were to have confidence in the linear servoanalytic approach. Therefore several topical autonomic drugs were applied to the eye in the following way: a 1 per cent solution of hydroxyamphetamine hydrobromide was applied as five successive 1-drop doses spaced by 2-minute intervals. In approximately 5 to 10 minutes the pupil was widely dilated. Then a 1-drop dose of a 0.1 per cent solution of eserine was given. This had the result of reducing pupil diameter again to about normal, that is, 3 to 4 mm. Thus many parameters of the pupil system such as dc operating level (average pupil area) remain constant. The pupil seemed to respond fairly well to both direct and consensual stimulation but, as will be brought out shortly, a quite definite change had occurred. These pharmacological studies have been used mainly as a tool to illustrate clearly the correlation of the frequency of the high-gain oscillation with the 180° phase crossover frequency. However, it is possible to use this sensitive system to define drug action. Fig. 13, which is the phase curve of the Bode diagram, plots the experimental results from these drug experiments. The solid line represents the pupil and its high-gain oscillation in an experiment before drugs, and the dashed line represents the pupil of the same eye 30 to 60 minutes later, after the application of the drug. It will be noted that the high-gain oscillation frequency and the 180° phase crossover frequency shifted in a parallel fashion (the undrugged pupil remained unchanged). Several experiments of this type have been performed and a consistent correlation was noted.

It has already been noted that stability is a fundamental property of a servomechanism. The characteristic fault or failure of an unstable servomechanism is exhibited as sustained oscillations or the process of building up to violent oscillations limited in magnitude only by the limit of linearity of the servomechanism. Among the methods that servo-engineers have developed for dealing with this problem, the Nyquist stability criterion stands highest because of its graphic representation of the physical state.

The Nyquist criterion has a fundamental mathematical meaning based upon the analysis of the characteristic equation of the system and the distribution of the roots of this equation. If the roots of the characteristic equation do not have a positive real component, then the response of the system to a bounded input will approach zero at some finite time: thus the system is stable. When the roots of the characteristic equation have positive real values, then instabilities exist and the system will respond to a bounded input by giving an unbounded response with increasing time. Although the Nyquist criterion is based upon this mathematical consideration, it allows one to analyze system stability without solving the equation, by using a graphic display of the open-loop response determined experimentally. Furthermore, unlike the analysis of root distributions themselves, this criterion possesses a visual indication of the degree of stability, and indeed suggests methods to correct for possible instability in the design of a control system.

The Nyquist plot as shown in Fig. 11 and again in Fig. 14 plots the open-loop response of the system as a vector quantity with gain as modulus and phase shift as angle. The critical point representing a gain of one at 180° phase lag is not encircled by the curve of a stable system. This is illustrated by the inner curve in Fig. 14 and the Nyquist curve in Fig. 11. In an unstable system, in particular in the pupil system which has been modified as in the high-gain experiment described above, the Nyquist curve encircles the critical point. This is

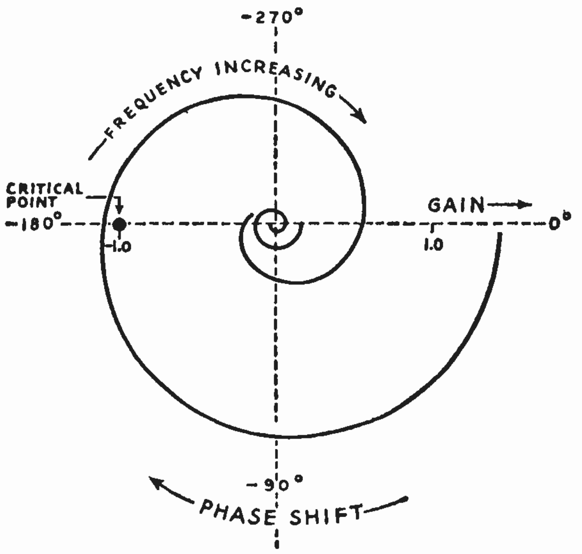


Fig. 14—Illustrated Nyquist diagram comparing plot of unstable (outer curve) with stable (inner curve) servosystem. The curve of stable system lies between Nyquist critical point and origin, indicating stability. Conversely, curve of unstable system encloses critical point indicating instability. As noted in text, critical point is that point in graphical plane of open-loop transfer function which represents gain of one at phase lag 180°. Since no frequency dependent parameters have been changed, the 180° phase cross-over frequency of the high-gain and low-gain systems are identical.

shown in the outer curve of Fig. 14 where the gain of the system is greater than one at a frequency where 180° phase lag is obtained. Any error is added to and enlarged each time the signal goes around the loop, because, at 180° phase lag, the control device is operating in the opposite direction to that required to cancel out the error. This regenerative or positive feedback produces system instability and oscillations. Thus it is possible to determine the degree of stability from the graphic display of open-loop system response data in the Nyquist diagram. In our experiments the gain of the system is increased by changing the light distribution at the plane of the pupil as shown in Fig. 6. This changes no other system characteristic but gain. Thus, the Nyquist plot of the system response will have only the scale of the modulus changed. We are, of course, assuming that no gross frequency dependent nonlinearity developed with these changes in operating conditions of the pupil. Fig. 14 shows both the Nyquist diagram of a low-gain stable system and the Nyquist diagram of a high-gain unstable system. Frequency is a monotonically increasing function around the curve, and since to a first-order approximation only the scale factor has been altered in this transformation, we would expect the high-gain unstable oscillations to develop at the frequency of 180° phase crossover of the low-gain system, and indeed this has been our experimental finding. Such is the Nyquist stability criterion and its application in the pupil servosystem.

The Nyquist stability criterion as well as the ability to use frequency response experiments depends on the assumed linearity of the servosystem under study. This means that if twice the magnitude of a signal is injected then twice the response will be observed; in other words, the theory of superposition holds for the system. It is realized, of course, that almost biological and indeed physical systems must be nonlinear. However, it is important to attempt to approach these fundamentally nonlinear systems by means of a linear approximation, because only in a linear domain are we able to transform and manipulate relationship in an easy manner. Therefore in our experiments we used a small signal approximation to aid in this linearization approach. This avoided large signal responses, permitted us to make approximations for the multiplication function in the control mechanism, and enabled us to neglect the harmonic distortion and other frequency dependent nonlinearities. The accuracy of the prediction obtained from this straightforward linear analysis is a good estimate of the accuracy of the approximation of the linear transfer function to the real pupil servosystem. Thus, the evidence presented here strongly justifies our application of linear analysis to the pupil servomechanism.

Pupil Unrest: An Example of Noise in a Biological Servomechanism

The pupil of the human eye continuously undergoes small fluctuations in area even in steady illumination.[14] These movements may be readily seen on close inspection. We considered it worthwhile to investigate this pupil unrest as an example of noise in a biological system. We have therefore measured and described the phenomenon, performed experiments to decide between possible mechanisms for its production, and finally considered its possible function.

Description of the Pupil Unrest

One-minute samples of pupil area fluctuations were recorded. A part of one such typical time function record is shown in Fig. 15 together with the autocorrelation function (Fig. 16) and power spectrum (Fig. 17) of this record. The power spectra are unreliable below 0.05 cps due to the short length of the time functions, and above 2 cps because the sampling rate limits the significance of the data. The autocorrelation function and the power spectra both show that the major component of the pupil unrest is random noise in the frequency range from 0.05 to 0.3 cps. In the region from 0.75 to 2.5 cps, all three power spectra we have obtained show a general slope of 36 db per octave with a corner frequency near 0.5 cps. One power spectrum obtained from a time function of pupil diameter fluctuations showed the same corner frequency and an 18 db per octave slope. Power spectra of diameter fluctuation should have half the slope of area fluctuations.

[14] L. Stark, F. W. Campbell, and J. Atwood, "Pupil unrest; an example of noise in a biological servomechanism," *Nature*, vol. 182, pp. 857–858; September 27, 1958.

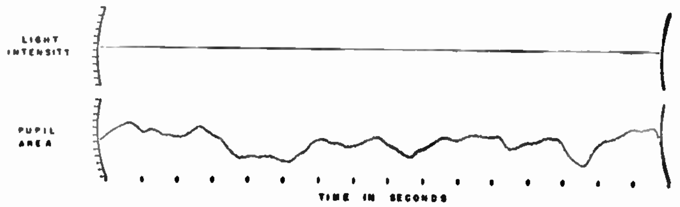


Fig. 15—Segment of record of pupil unrest from experiment no. 5 on subject J.F. Area and light intensity in arbitrary units.

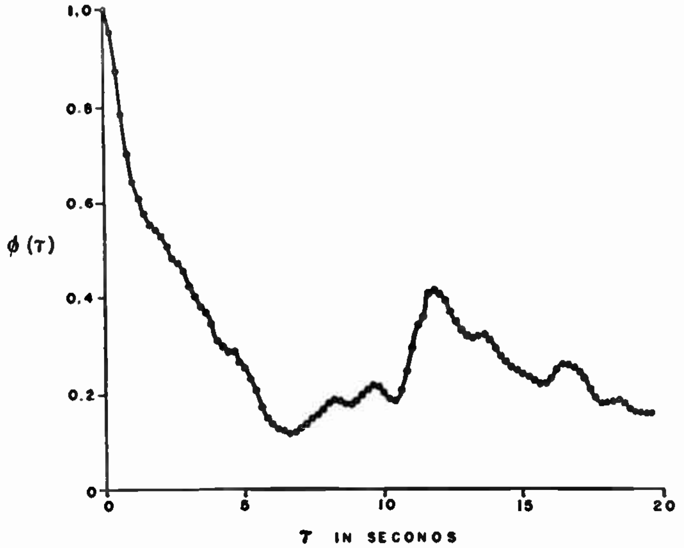


Fig. 16—Normalized autocorrelation function $\phi(\tau)$ computed from 71 seconds of recorded unrest of experiment no. 5.

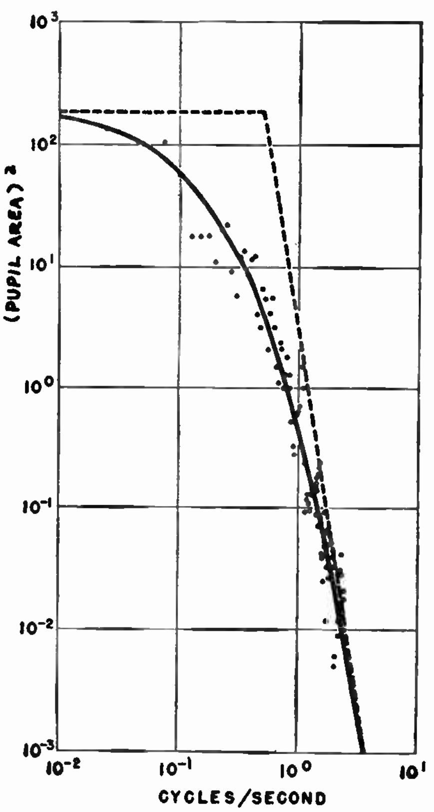


Fig. 17—Power spectrum computed from autocorrelation function of Fig. 16. Broken lines indicate approximate asymptotes. Area squared in arbitrary units.

The power spectra in the frequency range from 0.5 to 2 cps show an irregular distribution of power with frequency, which can be seen in the log-log plot of Fig. 17. Both the time functions and the autocorrelation functions have features which suggest the existence of coherent signals in this frequency range. However, study of the amplitude of these features in the autocorrelation functions and the power spectra show that if such signals exist, they represent less than 5 per cent of rms value of the time function.

The Origin of Pupil Unrest

It is possible to make some deductions about the origin of the pupil unrest from information already known about the pupil light reflex, from the new data presented and from the results of certain critical experiments described below.

The unrest cannot be a manifestation of instability in the pupil light reflex considered as an error actuated servomechanism. Such systems may oscillate, if the feedback path which completes the loop carries a response delayed by time lags and is sufficiently amplified to produce regenerative action. Considerable lags exist in the pupil system but the necessary amplification is not present because the gain is equal to 0.16. Eq. (4) is the transfer function $G(s)$. Thus the pupil will not continue to respond to a stimulus that does not itself continue, and the pupil light reflex servoloop is therefore stable. The absence of a large coherent frequency component at the frequency at which such instability oscillations would be expected is additional evidence against this possibility. Furthermore, as no regenerative instability is possible when the loop is open, pupil unrest should continue and appear unchanged under open-loop conditions. This is found to be true experimentally (Fig. 18). Thus, quite independent of the quantitative arguments above, we can disregard the hypothesis that an unstable pupil loop produces the pupil unrest.

Other linear properties of servosystems such as "backlash" or "dead space" also cause oscillations. However, the pupil loop does not show the change in gain which would be found if substantial amounts of backlash or dead space were present. Again, these nonlinearities require closed-loop operation in order to produce oscillations. Thus the experimental results shown in Fig. 18 also exclude these nonlinearities as a possible explanation for pupil unrest.

The above discussion eliminates causes of unrest associated with the closed-loop properties of the light reflex. Therefore, the unrest must originate in or be injected into the loop. Another observation suggests that this point of origin or injection must be into that portion of the loop which is common to both iris muscles. Only central nervous system elements, including the retina, comprise this portion of the loop. Fig. 19 shows the unrest of one iris to be highly correlated with the simultaneously recorded unrest of the contralateral iris. This excludes as an otherwise likely source of pupil unrest,

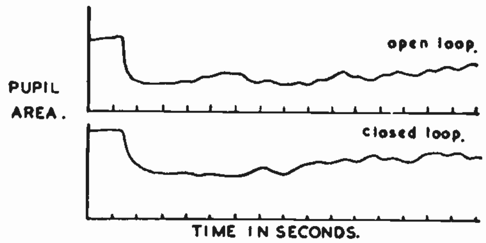


Fig. 18—Record of pupil unrest under (a) open loop and (b) closed loop operating conditions, showing general similarity. Pupil area in arbitrary units. Response to initial step change of light is also shown.

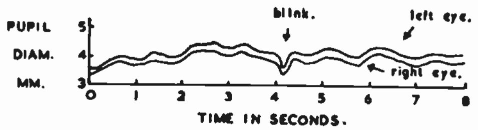


Fig. 19—Record of pupil unrest from right and left eye simultaneously, showing the high degree of correlation.

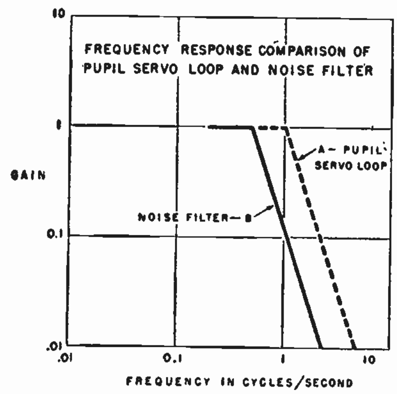


Fig. 20—Approximate asymptotes for (a) pupil servosystem and (b) lumped parameter network which would attenuate white noise to produce a power spectrum similar to Fig. 17. Both are normalized to unity gain in midfrequency range.

spontaneous movements of the iris muscle, of the type seen in a wide variety of smooth muscle when isolated and free from nervous influence.

The distinction between origin within and injection into the pupil reflex system can be made by comparing the frequency response of the open-loop system with the power spectrum of the unrest. Approximate asymptotes are plotted in Fig. 20(a) for the pupil servosystem and in Fig. 20(b) for a lumped parameter network which would attenuate white noise to produce a power spectrum similar to Fig. 17. It is evident that the frequency distribution of the power spectrum of the pupil unrest cannot have been caused by a filtering action of either the open-loop or closed-loop pupil system on noise with uniform distribution of power with frequency. Conversely, the filter which has shaped the power spectrum of this noise cannot be in series with the flow of information in the pupil system loop, because the measured frequency response of the pupil system does not have the 18 db per octave asymptote from 0.5 to 1.0 cps. It must be clearly understood that this bandwidth argument implies nothing about the actual anatomical location of the noise generator. For example, the noise may be leakage from signals external to the pupil system (crosstalk) or it may arise from part of the neurological apparatus which comprises the pupil system, as, for example, the retina. In such cases, although the noise generator and the pupil loop are anatomically united, it must still be true that they are functionally and physiologically separate; transmission of light information relevant to the operation of the pupil system cannot pass in series through the noise generator filter.

However, other experiments could localize the noise generator. For example, careful quantitative comparison under open-loop and closed-loop operating conditions could determine whether the noise is injected at the loop input or elsewhere. Because of the low loop gain and the low magnitude of the fluctuations (10 per cent change in area) this experiment is technically difficult and has not been satisfactorily performed. To identify the input which injects noise into the pupil system it is necessary to establish some measure of cross correlation between the presumptive input and the pupil unrest.

The Role of Pupil Unrest

Pupil unrest might play a specific role in the visual process, or it might represent the tolerance of the pupil servomechanism to uncorrected error. It is unlikely that pupil unrest prevents the disappearance of a stabilized image or produces information of value to the accommodation mechanism since the use of a fixed artificial pupil does not affect the characteristics of these functions. On the other hand, since visual acuity is not appreciably influenced by a 15 per cent change in pupil area, it seems reasonable that pupil area need not be controlled within more precise limits. Perhaps this is an example of economy in construction in the sense that unnecessary requirements are not placed on the noise level in the pupil apparatus.

Block Diagram Representation

The pupil system acts to regulate light impinging on the retina. To appreciate the pupil system fully, its interaction with the retina must be considered. This includes the processes of automatic gain control intrinsic to the retina, which are called dark and light adaptation by the physiologist. First, the sensing element for the controlled output by the pupil loop is the retina. Second, the hypothetical reference light flux, and the comparator of the block diagram in Figs. 1 and 6 can also be assumed to be retinal functions. The reference flux must be modifiable since, after a lag, the pupil system accepts any dc light operating level. This secondary loop feedback path may be presented as feeding into the reference flux with a block lying on the feedback path and describable as a simple lag with a gain of one. However, there is evidence from studies of even primitive retinas[15] that the relation-

[15] H. K. Hartline, "A quantitative and descriptive study of the electrical response to illumination of the arthropod eye," *Amer. J. Physiol.*, vol. 83, pp. 466-483; 1928.

ship between retinal flux and error signal into the pupil loop is similar to the relationship between retinal flux and optic nerve impulse frequency. This supports our original conjecture concerning the retinal location of the elements just discussed: sensor, comparator, reference flux, secondary loop feedback path, and log operator. It further suggests a transformation of the block diagram to one similar to that shown in Fig. 21. Here one block, marked "retinal adaptation," subserves all the previously scattered operators and groups them in a physiologically understandable manner with a retinal flux input and an optic nerve impulse rate output. The $H(s)$ feedback block within the larger retinal block is the simple lag which produces a lead-lag operating characteristic for the larger block, as will be shown in the later section on the simulation equations.

The optic nerve impulses $n(t)$ are delayed through the transport delay and represent the error signal into the pupil system. The next block represents the pupil lags. These pupil lags provide the characteristic slowly rising response of the pupil to a step input noted in many early experiments. Since the mechanism summarized in this pupil block is bilateral, the noise of the pupil system must be injected, fully band-limited ahead of this point in the block diagram (as proven in an earlier section of this paper).

The pupil area multiplied by external light intensity yields retinal flux. This is a simple physical relationship, but it assumes equal luminance values for light entering all parts of the pupil. Because this situation is known to be somewhat incorrect due to the Stiles-Crawford effect, the stimulus used in most of our experiments was designed to illuminate over a 30° retinal field. Thus, this effect is minimized and permits the simple physical law to dominate.

The log operator block is also part of the retina and represents the Weber-Fechner law operating on large signals. When the experimental system is linearized using small signals, this is eliminated. The negative block is inserted to make clear the change of sign of signal necessary if pupil area is the next parameter to be observed. This is especially important when considering the operation of the pupil electronic control system block which represents an electronic feedback control of external light intensity. Because of the low pupil gain this system works well both as open- and closed-loop. With the positive multiplier that relates pupil area to light intensity in the pupil system (remembering the negative block) oscillations occur at the 180° phase cross-over frequency. Of course, the electronic gain factor must be sufficient (the inverse of pupil gain) so that unity gain prevails in the combined loop comprising the pupil and the electronic feedback control system. It should further be noted that phase changes are possible within this electronic control system and thus instability oscillations of many frequencies are possible with various settings of phase and providing always that sufficient gain exists.

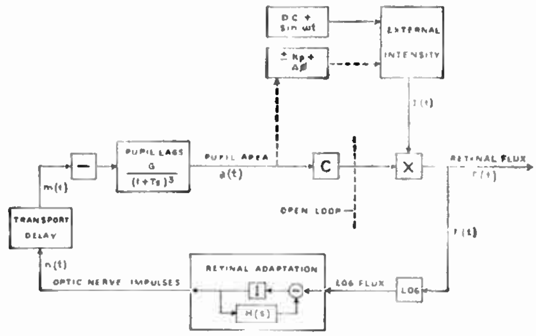


Fig. 21—Block diagram of the nonlinear pupil system with retinal adaptation units included as well as logging operator, multiplication operator, and the electronic feedback system.

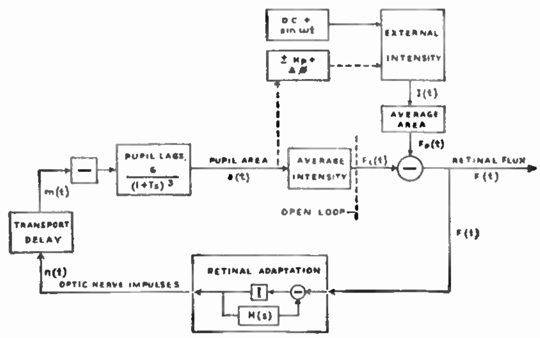


Fig. 22—Block diagram of the expanded pupil similar to Fig. 21 but linearized with approximations permitted by using small signal disturbances. In this way the multiplication operator is replaced by addition, and the logarithmic operator is excluded. Dashed lines represent the point of open-loop break.

The block just above the electronic feedback control system consists of the standard dc and sinusoidal generators for producing driven oscillations. The block marked C is the point at which light on the edge of the iris amplifies the effect of pupil area change sufficiently to produce instability oscillation as discussed in an earlier section of this paper. The dashed line indicates the break point of the system when open-loop operating conditions exist. In this experimental situation, pupil area has no effect on light flux hitting the retina and external intensity is converted into flux via the small area path in the center of the pupil, entirely under control of the experimenter.

The block diagram in Fig. 22 represents the pupil system operating under the small signal condition. Pupil area change multiplied by average intensity is subtracted (actually added if the negative multiplier is considered) from external intensity multiplied by average pupil area to yield retinal flux. The multiplication and logging operated blocks are removed. At frequencies high with respect to the retinal operator $H(s)$, the pupil system degenerates to the one indicated in the block diagram in Figs. 1 and 8. These simplified block diagrams describe many of our experiments quite accurately enough.

An important consideration in using these block diagrams is that they are designed to aid in our understanding of physical mechanism and have no sacrosanct viability of their own. The same is true of arbitrary definitions of servomechanism and regulator. When unorthodox biological systems, in the present-day engineering practice at least, are considered, it sometimes is necessary to concentrate on the actual functioning of the system under study and utilize eclectically all available conceptual tools.

Simulation Equations

As our knowledge of various parts of the pupil servomechanism has accumulated, we have attempted both to be more specific, by writing defining equations for various subentities, and to be more general in appreciating the over-all interaction of the pupil system with other systems, particularly with the retina. Using a digital computer program, an analog of the pupil system is being formed, and some of the defining equations so utilized will form the basis of our final discussion of the various portions of the pupil mechanism.

The relationship between the pupil area (a) and the nerve signal (m) expressed as a continuous function[16] is

$$T^3 \frac{d^3}{dt^3} a + 3T^2 \frac{d^2}{dt^2} a + 3T \frac{d}{dt} a = Gm \qquad (6)$$

where G is the gain of the pupil and T is the time constant. In general G is a function both of the direction of movement, and of the operating level of the pupil. The time constant of constriction is shorter that that of dilatation which gives constriction an apparent lower gain setting at lower frequencies. Retinal light adaptation also acts to decrease constriction gain, thus decreasing its gain even more at lower frequencies. This means, for example, that when the frequency of a steady-state oscillation is decreased from 1.5 to 0.5 cps, the dilatation gain becomes relatively greater than constriction gain. The pupil area has a net increase with each cycle until the dilatation gain (decreasing with increasing pupil area) lowers to a gain equal to constriction gain. At this time a new steady state is reached.

The pupil lags can be evaluated both from a careful open-loop experiment and, even more effectively, by changing their values as in the drug experiment shown in Fig. 13. Here a noticeable flattening of the convexity of the phase curve of the drugged pupil indicates that the shifted lag time constants are producing their effects relatively free of the nonminimum phase change. This gives an opportunity to verify the 270° of minimum phase change and the triple lag form of the operator. If we imagine a linear set of three sheets of low mass in a viscous medium connected by springs with a force pulling the first spring, the displacement of the edge of the last sheet would be related to the forcing function by an equation similar to the one given above. The ability to change these time constants with local application of drugs adds evidence in favor of such a view of the pupil lags, although other relationships such as a neuromuscular transmission lag or an apparent internal muscle viscosity lag (due to limitations in chemical energy transfer rate) are not excluded.

The time delay operator can be expressed either as [$\exp(-0.18s)$] or as

$$m(t) = n(t - l) \qquad (7)$$

where n is the nerve message from the retina, m the delayed nerve message, and l is the amount of the delay. However, although the form of this operator and its constant have been accurately determined, it does not seem likely that it is due to simple nerve conduction time as a 10-cm distance would delay a nerve impulse traveling 5 meters per second only 20 msec. These time estimates are quite conservative. The delay probably lies somewhere in the initial retinal integration process, but no definite evidence about this is available. These system equations have focused attention on the existence of physiological mechanisms which, although active in shaping pupil behavior, obviously have not been satisfactorily identified or studied.

The enormous input light range over which the human eye is able to perform some visual role must now be considered. It is of the order of 8 log units of intensity. The retinal system has at least four internal automatic gain control mechanisms to adjust its sensitivity inversely to retinal flux. One is the interesting mutual inhibition of neighboring units, probably interacting fields or groups of single receptors. This mechanism has been shown to have most important space contrast amplifying properties. It would also act as a sensitivity control more diffusely. The marked effect of cones acting to inhibit rods, as demonstrated by the work of Elenius, would belong to this class of interactions. Secondly, the number of sensory units, rods especially, which converge to form a receptive field, is an inverse function of retinal flux, and has the important effect of reducing sensitivity to light flux. This is inverse to the effect on visual acuity since resolution increases with increased intensity. Evidently the product of sensitivity and acuity is kept constant, thus probably permitting a high constant information handling capacity to exist in the retina at all levels of luminance. The bit per quantum ratio has been estimated to be about 0.25 at threshold intensities and possibly remains fairly close to this value over the available light input range. A third possible, but not proven, mechanism is a result of photochemical bleaching of the photosensitive pigment in the individual sensory cell. Somewhat parallel changes in concentration and sensitivity have been shown under certain experimental conditions and have been adduced as evidence for this third mechanism. The centrifugal fibers in the optic system demonstrated by Jacobson, Marg, and Granit might be a possible fourth adaptation mechanism. The neuro-

[16] It is difficult to handle with present-day analytic methods the pulse rate code of the nerve impulse. In our experimental studies we have usually integrated these nerve impulses with an appropriate filter to obtain a continuous voltage whose amplitude is a function of the pulse rate. The nerve message in the pupil analog is handled similarly.

logical reorganization of the retina to compensate for changes in external brightness seems to dominate in the mammalian eye. However, it is well known that subjective estimates of change in brightness and objective measures of pupillary responses closely parallel one another. Thus, it is quite reasonable to put the retinal adaptation operator with constants from light and dark adaptation studies into the pupil loop. The logarithmic factor and constants of the retinal lead-lag as adaptation operator can also be justified by adducing studies of electrical response of animal retinas in isolation, which show similar characteristics and with values of constants agreeing in the main with parameters utilized in our studies.

Since the retina does have a high internal gain compensation, we must consider the pupil in this context. The response of the pupil is limited in the first instance by the possible extreme values of its output range. These are a smallest area approximately equal to 2 square mm and a largest area approximately equal to 40 square mm.[17] Now, given this limited output range, the response of the pupil system can be expanded in a profligate manner to compensate for light changes over a limited part of the input range, or in an evenly distributed fashion over the entire input range. The former solution is apparent in some nocturnal animals, as the owl, which is almost blind in the daytime. (It should be remarked that a wide variety of pupillary shapes exist in the animal world; some, such as the slit pupil of the cat, may provide a larger pupil output range, possibly without decreasing visual acuity.) The latter solution is evident in the human pupil system. Here only a small percentage of a dc light change is compensated for by the iris mechanism; the remainder of the adjustment is handled by the aforementioned internal retinal mechanism. However, the pupil system does handle a 1 to 2 log compensation range, and after a discussion of frequency dependent factors in the retinal interaction with the pupil system, we shall consider this topic further.

The transfer function of the retinal system relating the retinal flux f with the nerve message n is

$$\frac{n(s)}{[\log f](s)} = k \frac{\left(1 + \frac{1}{kr} s\right)}{\left(1 + \frac{1}{r} s\right)} \tag{8}$$

[17] The smallest area is apparently limited by possible diffraction errors occurring with small pupil which reduces human visual acuity in experiments with artificial pupils (Liebowitz). The large end of the output range is limited by dioptic errors due to various lens aberrations which again have been shown to limit visual acuity in human psychophysical experiments.

or alternatively

$$\frac{1}{r}\frac{d}{dt}n + n = \frac{1}{f \cdot r}\frac{d}{dt}f + k \log f \tag{9}$$

where r is the time constant of retinal adaptation, and k is the dc level of retinal response. Here we see expressed the immediate open circuit of the retina in the pupil loop and then the gradual attenuation of the signal to a low dc level. This k, which represents the ratio of the dc response of the pupil to the dc change of light intensity, also represents the ratio of output range to input range if response is evenly distributed over the ranges.

There are, however, important asymmetries in the retinal system even if we neglect operating level and signal amplitude nonlinearities. The time constant of light adaptation is about 12 seconds and is only one-tenth the time constant of dark adaptation. This means that at frequencies below the active midfrequency range of the pupil, the retinal adaptation acts almost entirely as a factor decreasing the gain of pupillary constriction. This was discussed above when considering the frequency effect interacting with pupil lag asymmetries. In any case, the retinal effect can be noted in a first-order fashion as an attenuation of pupil gain and a phase advance, both effects becoming noticeable below 0.3 cps under certain experimental conditions.

We must carefully note that the output range to input range ratio, which is small and limiting for the dc response of the pupil, is not so completely constraining in the case of the ac midfrequency response of the pupil. In this case the pupil may respond quite fully to a sudden change of light at the expense of the ability to respond in the same direction to an additional stimulus. For example, the logarithmic operator acts to expand the response for a small signal. It is evident that certain probabilities of amplitudes and frequencies of light disturbance in the animal's environment would be optimally compensated for by a system responding as does the human pupil-retinal complex. We are in the process of determining these frequency and amplitude distributions in an attempt to understand more deeply the principles of design of the pupil servomechanism (evolutionary selections of random mutation being, of course, the design mechanism). In any case it is evident from our dynamic analysis that the pupil can respond to compensate for light changes more efficiently over a short period than one would suspect from a study of the dc relationships only. In fact, it can compensate for $\frac{1}{6}$ to $\frac{1}{2}$ of light intensity changes occurring in the midfrequency range of $\frac{1}{4}$ to 1 cps.

What the Frog's Eye Tells the Frog's Brain*

J. Y. LETTVIN†, H. R. MATURANA‡, W. S. McCULLOCH||, SENIOR MEMBER, IRE, AND W. H. PITTS||

Summary—**In this paper, we analyze the activity of single fibers in the optic nerve of a frog. Our method is to find what sort of stimulus causes the largest activity in one nerve fiber and then what is the exciting aspect of that stimulus such that variations in everything else cause little change in the response. It has been known for the past 20 years that each fiber is connected not to a few rods and cones in the retina but to very many over a fair area. Our results show that for the most part within that area, it is not the light intensity itself but rather the pattern of local variation of intensity that is the exciting factor. There are four types of fibers, each type concerned with a different sort of pattern. Each type is uniformly distributed over the whole retina of the frog. Thus, there are four distinct parallel distributed channels whereby the frog's eye informs his brain about the visual image in terms of local pattern independent of average illumination. We describe the patterns and show the functional and anatomical separation of the channels. This work has been done on the frog, and our interpretation applies only to the frog.**

Introduction

Behavior of a Frog

A FROG hunts on land by vision. He escapes enemies mainly by seeing them. His eyes do not move, as do ours, to follow prey, attend suspicious events, or search for things of interest. If his body changes its position with respect to gravity or the whole visual world is rotated about him, then he shows compensatory eye movements. These movements enter his hunting and evading habits only, *e.g.*, as he sits on a rocking lily pad. Thus his eyes are actively stabilized. He has no fovea, or region of greatest acuity in vision, upon which he must center a part of the image. He also has only a single visual system, retina to colliculus, not a double one such as ours where the retina sends fibers not only to colliculus but to the lateral geniculate body which relays to cerebral cortex. Thus, we chose to work on the frog because of the uniformity of his retina, the normal lack of eye and head movements except for those which stabilize the retinal image, and the relative simplicity of the connection of his eye to his brain.

The frog does not seem to see or, at any rate, is not concerned with the detail of stationary parts of the world around him. He will starve to death surrounded by food if it is not moving. His choice of food is determined only by size and movement. He will leap to capture any object the size of an insect or worm, providing it moves like one. He can be fooled easily not only by a bit of dangled meat but by any moving small object. His sex life is conducted by sound and touch. His choice of paths in escaping enemies does not seem to be governed by anything more devious than leaping to where it is darker. Since he is equally at home in water and on land, why should it matter where he lights after jumping or what particular direction he takes? He does remember a moving thing providing it stays within his field of vision and he is not distracted.

Anatomy of Frog Visual Apparatus

The retina of a frog is shown in Fig. 1(a). Between the rods and cones of the retina and the ganglion cells, whose axons form the optic nerve, lies a layer of connecting neurons (bipolars, horizontals, and amacrines). In the frog there are about 1 million receptors, $2\frac{1}{2}$ to $3\frac{1}{2}$ million connecting neurons, and half a million ganglion cells [1]. The connections are such that there is a synaptic path from a rod or cone to a great many ganglion cells, and a ganglion cell receives paths from a great many thousand receptors. Clearly, such an arrangement would not allow for good resolution were the retina meant to map an image in terms of light intensity point by point into a distribution of excitement in the optic nerve.

There is only one layer of ganglion cells in the frog. These cells are half a million in number (as against one million rods and cones). The neurons are packed together tightly in a sheet at the level of the cell bodies. Their dendrites, which may extend laterally from 50 μ to 500 μ, interlace widely into what is called the inner plexiform layer, which is a close-packed neuropil containing the terminal arbors of those neurons that lie between receptors and ganglion cells. Thus, the amount of overlap of adjacent ganglion cells is enormous in respect to what they see. Morphologically, there are several types of these cells that are as distinct in their dendritic patterns as different species of trees, from which we infer that they work in different ways. The anatomy shown in the figures is that found in standard references. Further discussion of anatomical questions and additional original work on them will appear in a later publication.

Physiology as Known up to This Study

Hartline [2] first used the term *receptive field* for the region of retina within which a local change of brightness would cause the ganglion cell he was observing to discharge. Such a region is sometimes surrounded by an annulus, within which changes of brightness affect the cell's response to what is occurring in the receptive field,

* Original manuscript received by the IRE, September 3, 1959. This work was supported in part by the U. S. Army (Signal Corps), the U. S. Air Force (Office of Sci. Res., Air Res. and Dev. Command), and the U. S. Navy (Office of Naval Res.); and in part by Bell Telephone Labs., Inc.

† Res. Lab. of Electronics and Dept. of Biology, Mass. Inst. Tech., Cambridge, Mass.
‡ Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Mass., on leave from the University of Chile, Santiago, Chile.
|| Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Mass.

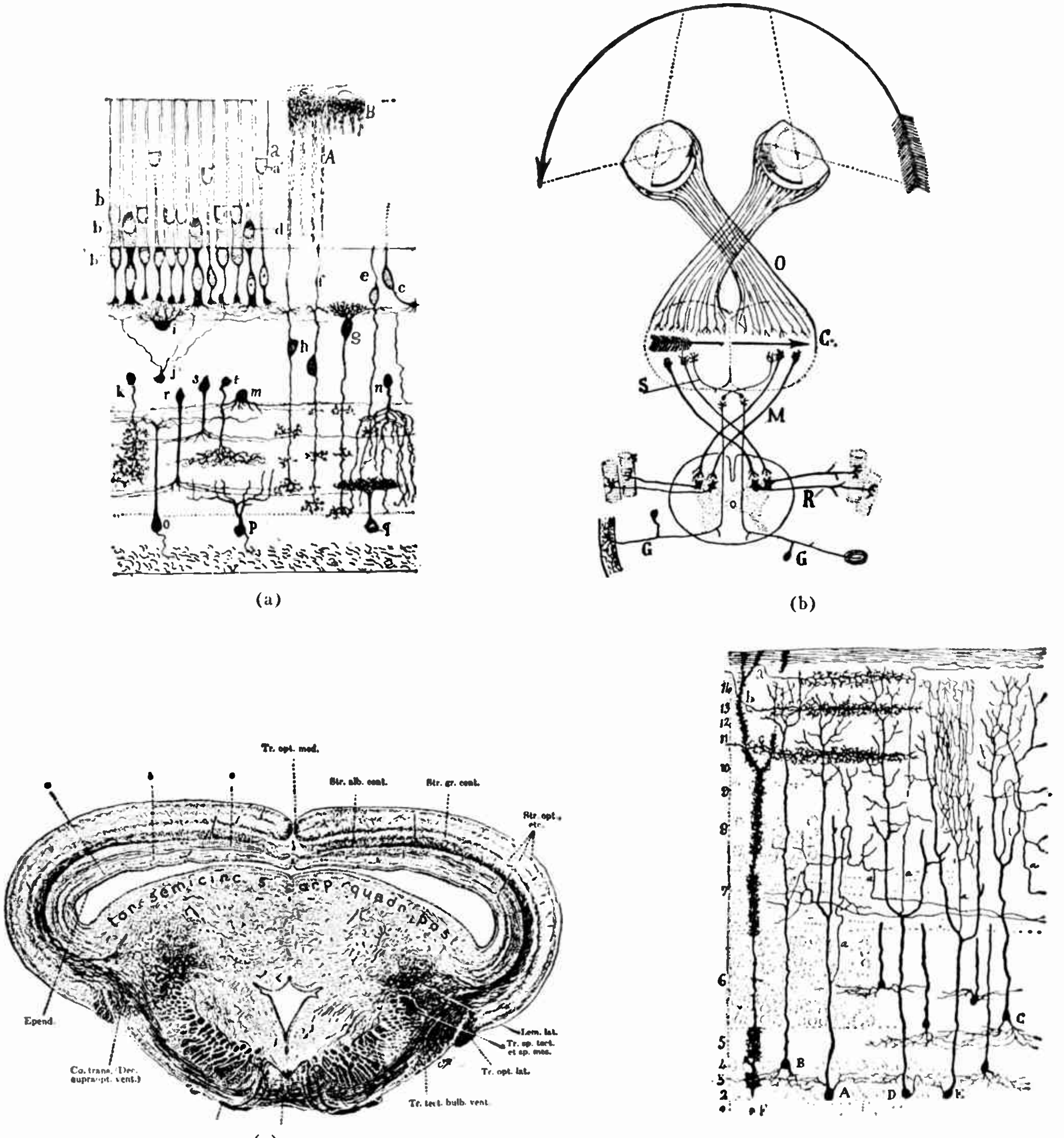


Fig. 1—(a) This is a diagram of the frog retina done by Ramon y Cajal over 50 years ago [9]. The rods and cones are the group of elements in the upper left quarter of the picture. To their bushy bottom ends are connected the bipolar cells of the intermediate layer, for example, *f, g,* and *h*. Lateral connecting neurons, called *horizontal* and *amacrine* cells, also occur in this layer, for example, *i, j* and *m*. The bipolars send their axons down to arborize in the inner plexiform layer, roughly the region bounded by cell *m* above and the bodies of the ganglion cells, *o, p* and *q*, below. In this sketch, Ramon has the axons of the bipolar cells emitting bushes at all levels in the plexiform layer; in fact, many of them branch at only one or two levels.

Compare the dendrites of the different ganglion cells. Not only do they spread out at different levels in the plexiform layer, but the patterns of branching are different. Other ganglion cells, not shown here, have multiple arbors spreading out like a plane tree at two or three levels. If the terminals of the bipolar cells are systematically arranged in depth, making a laminar operational map of the rods and cones in terms of very local contrast, color, ON, OFF, etc., then the different shapes of the ganglion cells would correspond to different combinations of the local operations done by the bipolars. Thus would arise the more complex operations of the ganglion cells as described in the text. (b) This is Ramon y Cajal's diagram of the total decussation or crossing of the optic nerve fibers in the frog [9]. He made this picture to explain the value of the crossing as preserving continuity in the map of the visual world. *O* is the optic nerve and *C* is the *superior colliculus* or *optic tectum* (the names are synonymous). (c) This is Ariens-Kapper's picture of the cross section of the brain of a frog through the colliculus, which is the upper or dorsal part above the enclosed space. (d) This is Pedro Ramon Cajal's diagram of the nervous organization of the tectum of a frog. The terminal bushes of the optic nerve fibers are labelled *a, b,* and *c. A, B, C, D* and *E* are tectal cells receiving from the optic nerve fibers. Note that the axons of these cells come off the dendrites in stratum 7, which we call the *palisade* layer. The endings discussed in this paper lie between the surface and that stratum.

although the cell does not discharge to any event occurring in the annulus alone. Like Kuffler [4], we consider the receptive field and its interacting annulus as a single entity, with apologies to Dr. Hartline for the slight change in meaning. Hartline found three sorts of receptive field in the frog: ON, ON-OFF, and OFF. If a small spot of light suddenly appears in the receptive field of an ON-cell, the discharge soon begins, increases in rate to some limit determined by the intensity and area of the spot, and thereafter slowly declines. Turning off the spot abolishes the discharge.

If the small spot of light suddenly appears or disappears within the field of an ON-OFF cell, the discharge is short and occurs in both cases.

If the spot of light disappears from the field of an OFF cell, the discharge begins immediately, decreases slowly in frequency, and lasts a long time. It can be abolished promptly by turning the spot of light on again.

For all three sorts of field, sensitivity is greatest at the center of each field and least at the periphery.

Barlow [3] extended Hartline's observations. He observed that the OFF cells have an adding receptive field, *i.e.*, the response occurs always to OFF at both center and periphery of that field, and that the effect of removing light from the periphery adds to the effect of a reduction of light at the center, with a weight decreasing with distance.

The ON-OFF cells, however, have differencing receptive fields. A discharge of several spikes to the appearance of light in the center is much diminished if a light is turned on in the extreme periphery. The same interaction occurs when these lights are removed. Thus, an ON-OFF cell seems to be measuring inequality of illumination within its receptive field. (Kuffler [4] at the same time showed a similar mutual antagonism between center and periphery in each receptive field of ganglion cells in the eye of a cat, and later Barlow, Kuffler and Fitzhugh [5] showed that the size of the cat's receptive fields varied with general illumination.) Barlow saw that ON-OFF cells were profoundly sensitive to movement within the receptive field. The ON cells have not been characterized by similar methods.

These findings of Hartline and Barlow establish that optic nerve fibers (the axons of the ganglion cells) do not transmit information only about light intensity at single points in the retina. Rather, each fiber measures a certain feature of the whole distribution of light in an area of the receptive field. There are three sorts of function, or areal operation, according to these authors, so that the optic nerve looks at the image on the retina through three distributed channels. In any one channel, the overlap of individual receptive fields is very great. Thus one is led to the notion that what comes to the brain of a frog is this: for any visual event, the OFF channel tells how much dimming of light has occurred and where; the ON-OFF channel tells where the boundaries of lighted areas are moving, or where local inequalities of illumination are forming; the ON channel shows (with a delay) where brightening has occurred. To an unchanging visual pattern, the nerve ought to become fairly silent after a while.

Consider the retinal image as it appears in each of the three distributed channels. For both the OFF and ON channels, we can treat the operation on the image by supposing that every point on the retina gives rise to a blur about the size of a receptive field. Then the OFF channel tells, with a long decay time, where the blurred image is darkened, and the ON channel tells with a delay and long decay where it is brightened. The third channel, ON-OFF, principally registers moving edges. Having the mental picture of an image as it appears through the three kinds of channel, we are still faced with the question of how the animal abstracts what is useful to him from his surroundings. At this point, a safe position would be that a fair amount of data reduction has in fact been accomplished by the retina and that the interpretation is the work of the brain, a yet-to-be unravelled mystery. Yet the nagging worries remain: why are there two complementary projections of equally poor resolution? Why is the mosaic of receptors so uselessly fine?

Initial Argument

The assumption has always been that the eye mainly senses light, whose local distribution is transmitted to the brain in a kind of copy by a mosaic of impulses. Suppose we held otherwise, that the nervous apparatus in the eye is itself devoted to detecting certain patterns of light and their changes, corresponding to particular relations in the visible world. If this should be the case, the laws found by using small spots of light on the retina may be true and yet, in a sense, be misleading. Consider, for example, a bright spot appearing in a receptive field. Its actual and sensible properties include not only intensity, but the shape of its edge, its size, curvature, contrast, etc.

We decided then how we ought to work. First, we should find a way of recording from single myelinated and unmyelinated fibers in the intact optic nerve. Second, we should present the frog with as wide a range of visible stimuli as we could, not only spots of light but things he would be disposed to eat, other things from which he would flee, sundry geometrical figures, stationary and moving about, etc. From the variety of stimuli we should then try to discover what common features were abstracted by whatever groups of fibers we could find in the optic nerve. Third, we should seek the anatomical basis for the grouping.[1]

[1] This program had started once before in our laboratory with A. Andrew [6], [7] of Glasgow who unfortunately had to return to Scotland before the work got well under way. However, he had reported in 1957 that he found elements in the colliculus of the frog that were sensitive to movement of a spot of light (a dot on an oscilloscope screen) even when the intensity of the spot was so low that turning it on and off produced no response. In particular, the elements he observed showed firing upon movement away from the centers of their receptive fields, but not to centripetal movements. As will appear later, this sort of response is a natural property of OFF fibers.

(Actual) Methods

Using a variant of Dowben and Rose's platinum black-tipped electrode described in another paper of this issue, we then began a systematic study of fibers in the optic nerve. One of the authors (H. R. M.) had completed the electron microscopy of optic nerve in frogs [8], and with his findings we were able to understand quickly why certain kinds of record occurred. He had found that the optic nerve of a frog contains about half a million fibers (ten times the earlier estimates by light microscopy). There are 30 times as many unmyelinated axons as myelinated, and both kinds are uniformly distributed throughout the nerve. The axons lie in small densely packed bundles of five to 100 fibers with about 100 A between axons, each bundle surrounded by one or more glial cells [8]. But along the nerve no bundle maintains its identity long, for the component fibers so braid between bundles that no two fibers stay adjacent. Thus the place a fiber crosses one section of the nerve bears little relation to its origin in the retina and little relation to where it crosses another section some distance away.

Fibers are so densely packed that one might suppose such braiding necessary to prevent serious interactions. On the other hand, the density makes the recording easier. A glial wall surrounds groups rather than single fibers, and penetration of the wall brings the tip among really bare axons each surrounded by neighbors whose effect is to increase the external impedance to its action currents, augmenting the external potential in proportion. Thus, while we prefer to use platinum black tips to improve the ratio of signal to noise, we recorded much the same population with ordinary sharp microelectrodes of bright Pt or Ag. The method records equally well from unmyelinated and myelinated fibers.

We used *Rana pipiens* in these experiments. We opened a small flap of bone either just behind the eye to expose the optic nerve, or over the brain to expose the superior colliculus. No further surgery was done except to open the membranes of connective tissue overlying the nervous structure. The frog was held in extension to a cork platform and covered with moist cloth. An animal in such a position, having most of his body surface in physical contact with something, goes into a still reaction—*i.e.*, he will not even attempt to move save to pain, and except for the quick small incision of the skin at the start of the operation our procedure seems to be painless to him. With the animal mounted, we confront his eye with an aluminum hemisphere, 20 mils thick and 14 inches in diameter, silvered to a matte grey finish on the inner surface and held concentric to the eye. On the inner surface of this hemisphere, various objects attached to small magnets can be moved about by a large magnet moved by hand on the outer surface. On our hemisphere, 1° is slightly less than an eighth of an inch long. In the tests illustrated, we use as stimulating objects a dull black disk almost 1° in diameter and a rectangle 30° long and 12° wide. However, in the textual report, we use a variety of other objects. As an indicator for the stimulus, we first used a phototube looking at an image of the hemisphere taken through a camera lens and focussed on the plane of a diaphragm. (Later we used a photomultiplier, so connected as to give us a logarithmic response over about 4 decades.) Thus we could vary how much of the hemisphere was seen by the stimulus detector and match that area in position and size against the receptive field of the fiber we were studying. The output of this arrangement is the stimulus line in the figures.

Findings

There are four separate operations on the image in the frog's eye. Each has its result transmitted by a particular group of fibers, uniformly distributed across the retina, and they are all nearly independent of the general illumination. The operations are: *1) sustained contrast detection; 2) net convexity detection; 3) moving edge detection; and 4) net dimming detection.* The first two are reported by unmyelinated fibers, the last two by myelinated fibers. Since we are now dealing with events rather than point excitations as stimuli, receptive fields can only be defined approximately, and one cannot necessarily distinguish concentric subdivisions. The fibers reporting the different operations differ systematically not only in fiber diameter (or conduction velocity) but also in rough size of receptive field, which ranges from about 2° diameter for the first operation, to about 15° for the last. The following description of these groups is definite.

1) Sustained Contrast Detectors

An unmyelinated axon of this group does not respond when the general illumination is turned on or off. If the sharp edge of an object either lighter or darker than the background moves into its field and stops, it discharges promptly and continues discharging, no matter what the shape of the edge or whether the object is smaller or larger than the receptive field. The sustained discharge can be interrupted (or greatly reduced) in these axons by switching all light off. When the light is then restored, the sustained discharge begins again after a pause. Indeed the response to turning on a distribution of light furnished with sharp contrast within the field is exactly that reported by Hartline for his ON fibers. In some fibers of this group, a contrast previously within the field is "remembered" after the light is turned off, for they will keep up a low mutter of activity that is not present if no contrast was there before. That this is not an extraordinary sensitivity of such an element in almost complete darkness can be shown by importing a contrast into its receptive field after darkening in the absence of contrast. No mutter occurs then. This memory lasts for at least a minute of darkness in some units.

In Fig. 2 we see the response of such a fiber in the optic nerve. We compare these responses with full illumi-

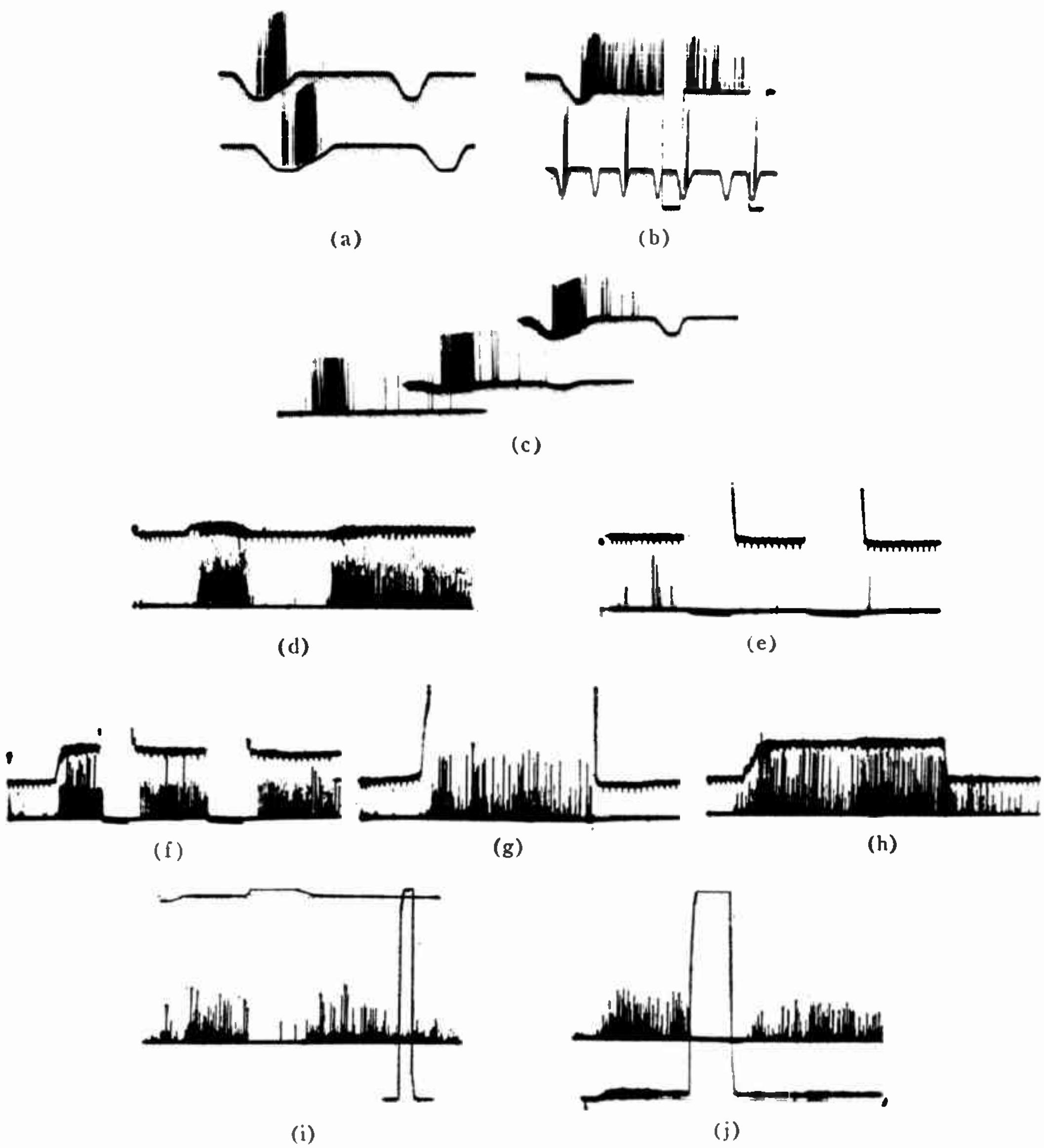


Fig. 2—Operation 1)—contrast detectors. The records were all taken directly with a Polaroid camera. The spikes are clipped at the lower end just above the noise and brightened on the screen. Occasional spikes have been intensified by hand for purposes of reproduction. The resolution is not good but we think that the responses are not ambiguous. Our alternate recording method is by means of a device which displays the logarithm of pulse interval of signals through a pulse height pick-off. However, such records would take too much explanation and would not add much to the substance of the present paper. (a) This record is from a single fiber in the optic nerve. The base line is the output of a photocell watching a somewhat larger area than the receptive field of the fiber. Darkening is given by downward deflection. The response is seen with the noise clipped off. The fiber discharge to movement of the edge of a 3° black disk passed in one direction but not to the reverse movement. (Time marks, 20 per second.) (b) The same fiber shown here giving a continued response when the edge stops in the field. The response disappears if the illumination is turned off and reappears when it is turned on. Below is shown again the asymmetry of the response to a faster movement. (Time marks, 20 per second.) (c) The same fiber is stimulated here to show asymmetrical response to the 3° black object moved in one direction, then the reverse and the stimuli are repeated under a little less than a 3-decade change of illumination in two steps. The bottom record is in extremely dim light, the top in very bright light. (Time marks, 20 per second.) (d) In the bottom line, a group of endings from such fibers is shown recorded from the first layer in the tectum. A black disk 1° in diameter is moved first through the field and then into the field and stopped. In the top line, the receptive field is watched by a photomultiplier (see text) and darkening is given by upward deflection. (Time marks, 5 per second for all tectal records.) (e) OFF and ON of general illumination has no effect on these fibers. (f) A 3° black disk is moved into the field and stopped. The response continues until the lights are turned OFF but reappears when the lights are turned ON. These fibers are nonerasable. (g) A very large black square is moved into the field and stopped. The response to the edge continues so long as the edge is in the field. (h) The 3° disk is again moved into the field and stopped. When it leaves, there is a slight after-discharge. (i) A 1° object is moved into the field, stopped, the light is then turned off, then on, and the response comes back. The light is, however, a little less than 300× dimmer than in the next frame. Full ON and OFF are given in the rectangular calibration on the right. (j) The same procedure as in Fig. 2(i) is done under very bright light. The return of response after reintroducing the light seems more prolonged—but this is due only to the fact that, in Fig. 2(i), the edge was not stopped in optimal position.

nation (a 60-watt bulb and reflector mounted a foot away from the plane of the opening of the hemisphere) to those with less than 1/300 as much light (we put a variable resistance in series with the bulb so that the color changed also). We are struck by the smallness of the resulting change. In very dim light where we can barely see the stimulating object ourselves, we still get very much the same response.

2) Net Convexity Detectors

These fibers form the other subdivision of the unmyelinated population, and require a number of conditions to specify when they will respond. To our minds, this group contains the most remarkable elements in the optic nerve.

Such a fiber does not respond to change in general illumination. It does respond to a small object (3° or less) passed through the field; the response does not outlast the passage. It continues responding for a long time if the object is imported and left in the field, but the discharge is permanently turned off (erased) by a transient general darkness lasting 1/10 second or longer. We have not tried shorter obscurations.

The fiber will not respond to the straight edge of a dark object moving through its receptive field or brought there and stopped. If the field is about 7° in diameter, then, if we move a dark square 8° on the side through it with the edge in advance there is no response, but if the corner is in advance then there is a good one. Usually a fiber will respond indefinitely only to objects which have moved into the field and then lie wholly or almost wholly interior to the receptive field. The discharge is greater the greater the convexity, or positive curvature, of the boundary of the dark object until the object becomes as small as about ½ the width of the receptive field. At this point, we get the largest response on moving across that field, and a good, sustained response on entering it and stopping. As one uses smaller and smaller objects, the response to moving across the field begins to diminish at a size of about 1°, although the sustained response to coming in and stopping remains. In this way we find the smallest object to which these fibers respond is less than 3 minutes of arc. A smooth motion across the receptive field has less effect than a jerky one, if the jerks recur at intervals longer than ½ second. A displacement barely visible to the experimenter produces a marked increase in response which dies down slowly.

Any checked or dotted pattern (in the latter case, with dots no further apart than half the width of the receptive field) moved as a whole across the receptive field produces little if any response. However, if any dot within the receptive field moves differentially with respect to the background pattern, the response is to that dot as if it were moving alone. A group of two or three distinct spots enclosed within the receptive field and moved as a whole produce less direct response to movement and much less sustained response on stopping than if the spots are coalesced to a single larger spot.

A delightful exhibit uses a large color photograph of the natural habitat of a frog from a frog's eye view, flowers and grass. We can move this photograph through the receptive field of such a fiber, waving it around at a 7-inch distance: there is no response. If we perch with a magnet a fly-sized object 1° large on the part of the picture seen by the receptive field and move only the object we get an excellent response. If the object is fixed to the picture in about the same place and the whole moved about, then there is none.

Finally, the response does not depend on how much darker the object is than its background, so long as it is distinguishably so and has a clear-cut edge. If a disk has a very dark center and merges gradually into the grey of the background at the boundary, the response to it is very much less than to a uniform grey disk only slightly darker than the background. Objects lighter than the background produce almost no response unless they have enough relief to cast a slight shadow at the edge.

All the responses we have mentioned are relatively independent of illumination, and Fig. 3 taken as described in the caption shows the reactions to a 3° object and the large rectangle under some of the conditions described.

General Comments on Groups 1) and 2)

The two sorts of detectors mentioned seem to include all the unmyelinated fibers, with conduction velocities of 20 to 50 cm. The two groups are not entirely distinct. There are transition cases. On one hand, some convexity detectors respond well to very slightly curved edges, even so far as to show an occasional sustained response if that edge is left in the field. They may also not be completely erasable (though very markedly affected by an interruption of light) for small objects. On the other hand, others of the same group will be difficult to set into an indefinitely sustained response with any object, but only show a fairly long discharge, acting thereby more as detectors of edges although never reacting to straight edges. Nevertheless the distribution of the unmyelinated axons into two groups is very marked. Any fiber of either group may show a directional response—*i.e.*, there will be a direction of movement that may fail to excite the cell. For the contrast fibers, this will also appear as a nonexciting angle of the boundary with respect to the axis of the frog. Such null directions and angles cancel out in the aggregate.

3) Moving-Edge Detectors

These fibers are myelinated and conduct at a velocity in the neighborhood of 2 meters per second. They are the same as Hartline's and Barlow's ON-OFF units. The receptive field is about 12° wide. Such a fiber re-

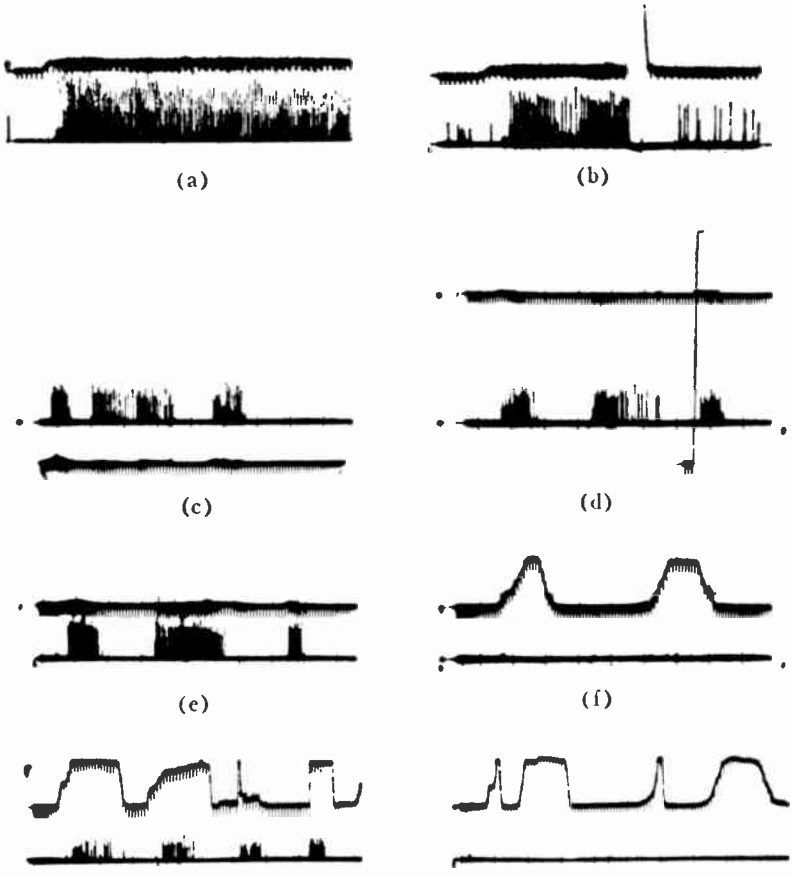


Fig. 3—Operation 2)—convexity detectors. The photomultiplier is used, and darkening is an upward deflection. (a) These records are all from the second layer of endings in the tectum. In the first picture, a 1° black disk is imported into the receptive field and left there. (b) The same event occurs as in Fig. 3(a), but now the light is turned off then on again. The response is much diminished and in the longer record vanishes. These fibers are erasable. (c) The 1° disk, is passed through the field first somewhat rapidly, then slowly, then rapidly. The light is very bright. (d) The same procedure occurs as in Fig. 3(c), but now the light has been dimmed about 300×. The vertical line shows the range of the photomultiplier which has been adjusted for about $3\frac{1}{2}$ decades of logarithmic response. (e) A 1° black disk is passed through the field at three speeds. (f) A 15° black strip is passed through at two speeds edge leading. (g) A 15° black strip is passed through in various ways with corner leading. (h) The same strip as in Fig. 3(g) is passed through, edge leading.

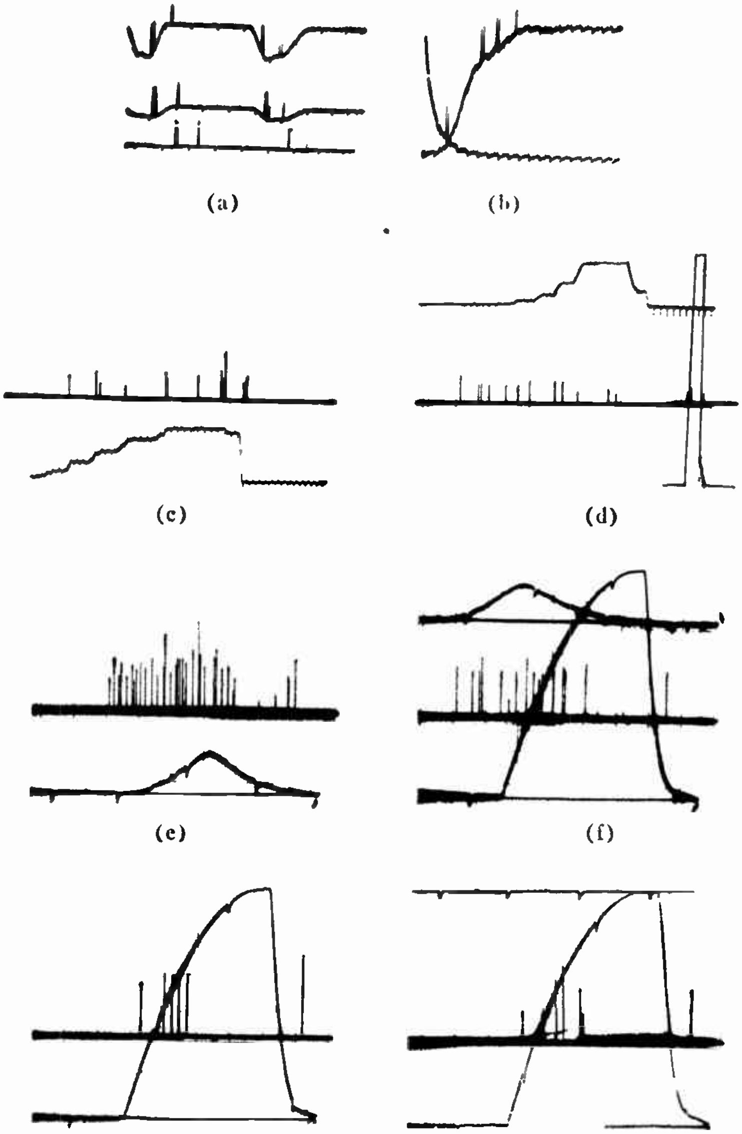
sponds to any distinguishable edge moving through its receptive field, whether black against white or the other way around. Changing the extent of the edge makes little difference over a wide range, changing its speed a great one. It responds to an edge only if that edge moves, not otherwise. If an object wider than about 5° moves smoothly across the field, there are separate responses to the leading and trailing edges, as would be expected from Barlow's formulation. These fibers generally show two or three regularly spaced spikes, always synchronous among different fibers to turning the light on or off or both. The response to moving objects is much greater than to changes in total illumination and varies only slightly with general illumination over a range of 1/300. The frequency of the discharge increases with the velocity of the object within certain limits (see Fig. 4).

4) Net Dimming Detectors

These are Hartline's and Barlow's OFF fibers. But they have some properties not observed before. They are myelinated and the fastest conducting afferents, clocked at 10 meters per second.[2] One such fiber responds to sudden reduction of illumination by a prolonged and regular discharge. Indeed, the rhythm is so much the same from fiber to fiber that in recording from several at once after sudden darkening, the impulses assemble in groups, which break up only after many seconds. Even then the activity from widely separated retinal areas seems to be related. We observe that the surface potential of the colliculus shows a violent and prolonged oscillation when the light is turned off. This oscillation, beginning at about 18 per second and breaking into 3 to 5 per second after several seconds, seems to arise from these fibers from the retina; the same record is seen when the optic nerve is severed and the recording electrode placed on the retinal stump. See Fig. 5.

The receptive field is rather large—about 15°—and works as Barlow describes. Darkening of a spot produces less response when it is in the periphery of the field than when it is at the center. The effect of a mov-

[2] The even faster fibers, with velocities up to 20 meters per second, we presently believe to be the efferents to the retina, but although there is some evidence for this, we are not yet quite certain.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 4—Operation 3)—moving-edge detectors. The first two pictures are taken from a single fiber in the optic nerve. (a) Shows a 7° black disk moving through the receptive field (the photocell was not in registration with the field). There is a response to the front and back of the disk independent of illumination. There is about a 300/1 shift of illumination between top and bottom of the record. Darkening is a downward deflection with the photocell record. (Time marks, 5 per second.) (b) OFF and ON of general lighting. (Time marks, 50 per second.) Note double responses and spacing. (c) This and succeeding records are in the third layer of endings in the tectum. Several endings are recorded but not resolved. Darkening is an upward deflection of the photomultiplier record. The response is shown to the edge of a 15° square moved into and out of the field by jerks in bright light. (d) The same procedure occurs as in Fig. 4(c), but in dim light. Calibration figure is at the right. (e) The response is shown to a 7° black disk passed through the receptive fields under bright light. The sweep is faster, but the time marks are the same. (f) The same procedure as for Fig. 4(e), but under dim light. The OFF and ON of the photomultiplier record was superimposed for calibration. (g) OFF and ON response with about half a second between ON and OFF. (h) Same as Fig. 4(g), but with 2 seconds between OFF and ON.

ing object is directly related to its size and relative darkness. The response is prolonged if a large dark object stops within the field. It is almost independent of illumination, actually increasing as the light gets dimmer. There is a kind of erasure that is complementary to that of group 2). If the general lighting is sharply dimmed, but not turned off entirely, the consequent prolonged response is diminished or abolished after a dark object passes through the receptive field. In this case, the reasons for erasure are apparent. Suppose one turns off the light and sets up a prolonged response. Then the amount of light which must be restored to interrupt the response gets less and less the longer one waits. That is, the sensitivity of the OFF discharge to the ON of light increases with time. If we darken the general lighting only by a factor of 100, we also get a prolonged discharge. However, if we turn off the light completely a few seconds after the 100/1 dimming and then turn it back on to the same dim level, the discharge is increased by the second dimming and is completely or almost completely abolished by the relighting. The effect of moving a dark object through the field after dimming is to impose a second dimming pulse followed by brightening as the object passes.

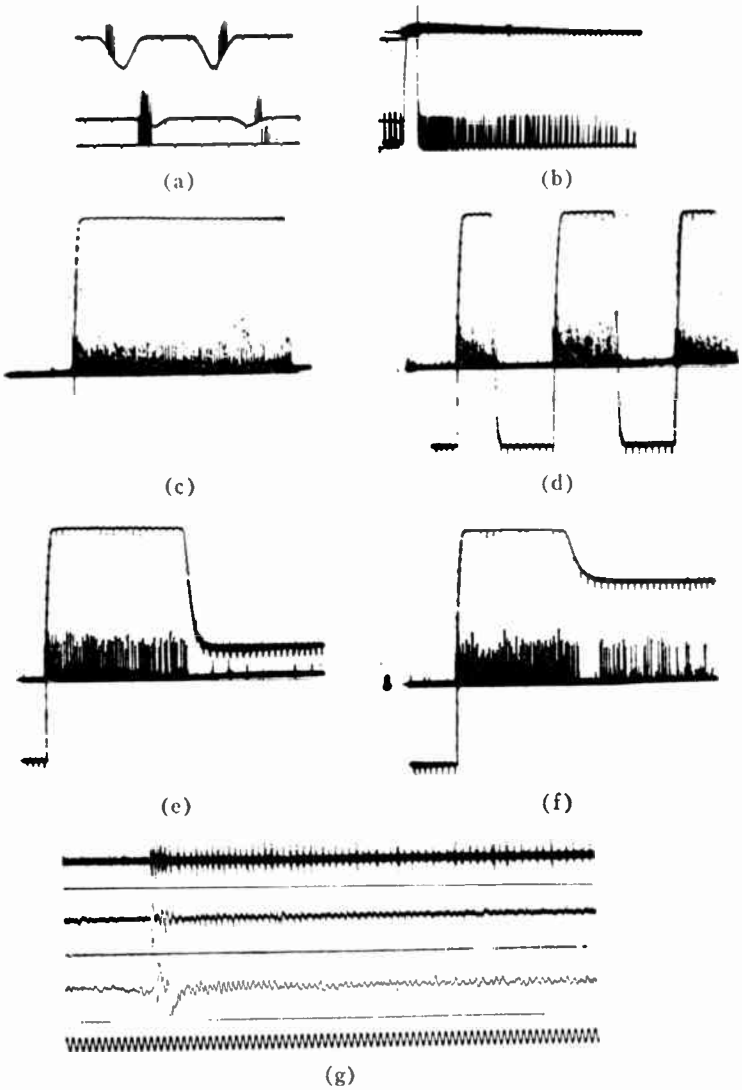


Fig. 5—Operation 4)—dimming detectors. (a) This and the next frame are taken from a single fiber in the optic nerve. Here we see the response to a 7° black disk passing through the receptive field. The three records are taken at three illumination levels over a 300:1 range. In the phototube record, darkening is a downward deflection. (Time marks, 5 per second.) (b) OFF and ON of light. The OFF was done shortly after one sweep began, the ON occurred a little earlier on the next sweep. The fiber is silenced completely by the ON. (Time marks, 5 per second.) (c) In this and the next three frames, we are recording from the fourth layer of endings in the tectum. This frame shows the response to turning OFF the general illumination. (d) OFF and ON of light at regular intervals. (e) OFF then ON of the light to a lesser brightness. (f) OFF then ON of the light to a still lesser brightness. The level to which the ON must come to abolish activity decreases steadily with time. (g) The synchrony of the dimming detectors as described in the text. At the top are three or four fibers recorded together in the optic nerve when the light is suddenly turned off. The fibers come from diverse areas on the retina. In the second line are the oscillations recorded from the freshly cut retinal stump of the optic nerve when the light is suddenly turned off. In the third line are the oscillations recorded on the surface of the tectum, the visual brain, before the nerve was cut. Again the light is suddenly turned off. The last line is 20 cps. These records of synchrony were obviously not all made at the same time, so that comparing them in detail is not profitable.

Others

Lastly, there is a small group of afferent fibers which does not seem to have distinct receptive fields. They each measure the absolute degree of darkness over a wide area with a long time constant. That is, the frequency of discharge is greater the darker it is. They have a complement in that some of the moving edge detectors have a resting discharge of very low frequency if the illumination is extremely bright.

Discussion

Let us compress all of these findings in the following description. Consider that we have four fibers, one from each group, which are concentric in their receptive fields. Suppose that an object is moved about in this concentric array:

1) The contrast detector tells, in the smallest area of all, of the presence of a sharp boundary, moving or still, with much or little contrast.

2) The convexity detector informs us in a somewhat larger area whether or not the object has a curved boundary, if it is darker than the background and moving on it; it remembers the object when it has stopped, providing the boundary lies totally within that area and is sharp; it shows most activity if the enclosed object moves intermittently with respect to a background. The memory of the object is abolished if a shadow obscures the object for a moment.

3) The moving-edge detector tells whether or not there is a moving boundary in a yet larger area within the field.

4) The dimming detector tells us how much dimming occurs in the largest area, weighted by distance from the center and by how fast it happens.

All of the operations are independent of general illumination. There are 30 times as many of the first two detectors as of the last two, and the sensitivity to sharpness of edge or increments of movement in the first two is also higher than in the last two.

Results in the Tectum

As remarked earlier, the optic nerve fibers are all disordered in position through the nerve. That is, the probability that any two adjacent fibers look at the same region in the retina is very small. However, when the fibers terminate in the superior colliculus they do so in an orderly way such that the termini exhibit a continuous map of the retina. Each optic nerve crosses the base of the skull and enters the opposite tectum [Fig. 1(b)] via two bundles—one rostromedial, the other caudalateral. The fibers sweep out over the tectum in the superficial neuropil in what grossly appears to be a laminated way [Fig. 1(c)]. The detail of ending is not known, and there is some reason to think Pedro Ramon's drawing [9] is too diagrammatic [Fig. 1(d)], however well it fits with our data.

In any case, the outer husk of neuropil, roughly about half the thickness of the optic tectum, is formed of the endings of the optic fibers mixed with dendrites of the deeper lying cells, and in this felting lie few cell bodies.

We have found it singularly easy to record from these terminal bushes of the optic fibers. That is, if an electrode is introduced in the middle of one bush, the external potential produced by action currents in any branch increases in proportion to the number of branches near the electrode. Since the bushes are densely interdigitated everywhere, it is not difficult to record from terminal arbors anywhere unless one kills or blocks them locally, as is easily done by pressure, etc.

One may inquire how we can be sure of recording from terminal arbors, and not from cells and their dendrites. The argument is this. First, there are about four layers of cells in the depths of the tectum [Fig. 1(d)], and only their dendrites issue into the superficial neuropil wherein lie very few cells indeed. There are about 250,000 of these cells in all, compared to 500,000 optic fibers. In the outer thickness of the tectum, among the terminating fibers, almost every element we record performs one of the four operations characterizing the fibers in the nerve, and has a corresponding receptive field. Now as the electrode moves down from the surface in one track, we record 5 to 10 cells in the deepest half of the tectum. Not a single cell so recorded shows activity even remotely resembling what we find in the superficial neuropil. Among the cells, none show optic nerve operations, and the smallest receptive fields we find are over 30° in diameter. If the active elements in the upper layers are cells (one will see about 20 to 30 active elements in one electrode track before reaching the cell layer), which cells can they be? Or if they are dendrites, to what neurons do they belong? We regard these considerations as conclusive enough.

Figs. 2–5 show that the four operational groups of fibers terminate in four separate layers of terminals, each layer exhibiting a continuous map of the retina (we confirm Gaze's diagram of the projection [10]) and all four maps are in registration. Most superficial lie the endings for the contrast detectors, the slowest fibers. Beneath them, but not so distinctly separate, are the convexity detectors. Deeper, and rather well separated, are the moving-edge detectors admixed with the rare and ill-defined axons that measure the actual level of darkness. Deepest (and occasionally contaminated with tectal cells or their axons) lie the dimming detectors. Thus the depth at which these fibers end is directly related to their speed of conduction.

Such an arrangement makes experiment easy, for all the fibers of one operation performed on the same field in the retina end in one place in the tectum and can be recorded as a group. It is very useful to see them this way, for then the individual variations among similar units cancel one another and only the common properties remain. We made the tectal records shown in the accompanying figures with a single electrode in two penetrations (to get decent separation of contrast and convexity detectors which lie just below the pia), to show how clear-cut the arrangement is.

Confirmation of Sperry's Proposal

The existence of a fourfold map of the retina in the tectal neuropil led us, naturally, to repeat Sperry's initial experiment on the regeneration of cut optic nerve [11]. Since the nerve is as scrambled as it can be originally, we saw no point in turning the eye around 180° but simply cut one nerve in a few frogs, separated the stumps to be sure of complete severance, and waited for about 3 months. At the end of this time, after it was clear that the cut nerves were functioning again, we compared the tectal maps of the cut and uncut nerves in some of them. We confirmed (as did Gaze [12]) Sperry's proposal that the fibers grew back to the regions where they originally terminated in mapping the retina [13]. But we also found a restoration of the four layers with no error or mixing. In one frog, after 90 days, the fibers had grown back best at the entrance of the two brachia to the colliculus, and least at the center, yet there were no serious errors. The total area of retina communicating with one point of the collicular neuropil (*i.e.*, the sum of the receptive fields of the fibers recorded from that point) had increased three or four times, from a diameter of about 15° to a diameter of about 30°. But there was no admixture of fibers with receptive fields in widely separated regions. In another frog, after 120 days, the area seen from one point was barely twice normal.

General Discussion

What are the consequences of this work? Fundamentally, it shows that the eye speaks to the brain in a language already highly organized and interpreted, instead of transmitting some more or less accurate copy of the distribution of light on the receptors. As a crude analogy, suppose that we have a man watching the clouds and reporting them to a weather station. If he is using a code, and one can see his portion of the sky too, then it is not difficult to find out what he is saying. It is certainly true that he is watching a distribution of light; nevertheless, local variations of light are not the terms in which he speaks nor the terms in which he is best understood. Indeed, if his vocabulary is restricted to types of things that he sees in the sky, trying to find his language by using flashes of light as stimuli will certainly fail. Now, since the purpose of a frog's vision is to get him food and allow him to evade predators no matter how bright or dim it is about him, it is not enough to know the reaction of his visual system to points of light. To get useful records from individual receptors (the rods and cones), assuming that they operate independently and under no reflex control, this stimulus may be adequate. But when one inspects responses that are a few nervous transformations removed from the receptors, as in the optic nerve, that same choice of stimulus is difficult to defend. It is equivalent to assuming that all of the interpretation is done further on in the nervous system. But, as we have seen, this is false.

One might attempt to measure numerically how the response of each kind of fiber varies with various properties of the successions of patterns of light which evoke them. But to characterize a succession of patterns in space requires knowledge of so many independent variables that this is hardly possible by experimental enumeration of cases. To examine the effect of curvature alone we should have to explore at least the response to all configurations of three spots moving in a large variety of ways. We would prefer to state the operations of ganglion cells as simply as possible in terms of whatever *quality* they seem to detect and, instead, examine the bipolar cells in the retina, expecting to find there a dissection of the operations into combinations of simpler ones performed at intermediate stages. For example, suppose that there are at least two sorts of rods and cones, one increasing its voltage with the log of light at one color, the other decreasing its voltage with the log of light at some other color. If bipolars connect to several contiguous rods or cones of opposing reactions and simply add voltages, some bipolars will register a large signal only if an appropriate contrast occurs. We have in fact found something of the sort occurring, for it seems that the inner plexiform layer of the retina is stratified to display several different local properties, one layer indicating local differences in intensity of light. Some of Svaetichin's [14] data can be adduced here. The different dendritic distribution of the ganglion cells, as in Fig. 1(a), may signify that they extract differently weighted samples of simple local operations done by the bipolars, and it is on this that we are now working.

But there is another reason for a reluctance to make accurate measurements on the activity of ganglion cells in the intact animal. A significant efferent outflow goes to the retina from the brain. We now know to a certain extent how the cells in the tectum handle the four inputs to them which are described in this paper. There are at least two distinct classes of these cells, and at least one of them issues axons back into the optic nerve. We have recorded this activity there. Such axons enter the retina and we think some effects of their activity on the sensitivity of ganglion cells are noticeable.

The way that the retina projects to the tectum suggests a nineteenth century view of visual space. The image on the retina, taken at the grain of the rods and cones, is an array of regularly spaced points at each of which there is a certain amount of light of a certain composition. If we know the position of every point and the values of light at every point, we can physically reconstruct the image, and looking at it understand the picture. If, however, we are required to establish continuities within the picture only from the numerical data on position and light at independent points, it becomes a very difficult task. The retina projects onto the tectum in four parallel sheets of endings, each sheet mapping continuously the retina in terms of a particular areal operation, and all four maps are in registration. Consider the dendrite of a tectal cell extending up through the four sheets. It is looking at a point in the image on the retina, but that point is now seen in terms of the properties of its neighborhood as defined by the operations. Since the overlap of receptive fields within any operation is very great, it now seems reasonable to erect simple criteria for finding continuities. For example, if an area over which there is little change in the fourfold signature of a moving object is bounded by regions of different signature, it seems likely that that area describes the image of a single object.

By transforming the image from a space of simple discrete points to a congruent space where each equivalent point is described by the intersection of particular qualities in its neighborhood, we can then give the image in terms of distributions of combinations of those qualities. In short, every point is seen in definite contexts. The character of these contexts, genetically built in, is the physiological synthetic *a priori*. The operations found in the frog make unlikely later processes in his system of the sort described by two of us earlier [15], for example, dilatations; but those were adduced for the sort of form recognition which the frog does not have. This work is an outgrowth of that earlier study which set the question.

Conclusion

The output from the retina of the frog is a set of four distributed operations on the visual image. These operations are independent of the level of general illumination and express the image in terms of: 1) local sharp edges and contrast; 2) the curvature of edge of a dark object; 3) the movement of edges; and 4) the local dimmings produced by movement or rapid general darkening. Each group of fibers serving one operation maps the retina continuously in a single sheet of endings in the frog's brain. There are four such sheets in the brain, corresponding to the four operations, and their maps are in registration. When all axonal connections between eye and brain are broken and the fibers grow back, they reconstitute the original retinal maps and also arrange themselves in depth in the original order with no mistakes. If there is any randomness in the connections of this system it must be at a very fine level indeed. In this, we consider that Sperry [11] is completely right.

We have described each of the operations on the retinal image in terms of what common factors in a large variety of stimuli cause response and what common factors have no effect. What, then, does a particular fiber in the optic nerve measure? We have considered it to be how much there is in a stimulus of that quality which excites the fiber maximally, naming that quality.

The operations thus have much more the flavor of perception than of sensation if that distinction has any meaning now. That is to say that the language in which they are best described is the language of complex abstractions from the visual image. We have been tempted, for example, to call the convexity detectors "bug perceivers." Such a fiber [operation 2] responds best when a dark object, smaller than a receptive field, enters that field, stops, and moves about intermittently thereafter. The response is not affected if the lighting changes or if the background (say a picture of grass and flowers) is moving, and is not there if only the background, moving or still, is in the field. Could one better describe a system for detecting an accessible bug?

Acknowledgment

We are particularly grateful to O. G. Selfridge whose experiments with mechanical recognizers of pattern helped drive us to this work and whose criticism in part shaped its course.

Bibliography

[1] H. R. Maturana, "Number of fibers in the optic nerve and the number of ganglion cells in the retina of Anurans," *Nature*, vol. 183, pp. 1406–1407; May 16, 1959.
[2] H. K. Hartline, "The response of single optic nerve fibres of the vertebrate eye to illumination of the retina," *Amer. J. Physiol.*, vol. 121, pp. 400–415; February, 1938.
Also, "The receptive fields of the optic nerve fibers," *Amer. J. Physiol.*, vol. 130, pp. 690–699; October, 1940.
[3] H. B. Barlow, "Summation and inhibition in the frog's retina," *J. Physiol.*, vol. 119, pp. 69–88; January, 1953.
[4] S. W. Kuffler, "Discharge patterns and functional organization of mammalian retina," *J. Neurophysiol.*, vol. 16, pp. 37–68; January, 1953.
[5] H. B. Barlow, R. Fitzhugh, and S. W. Kuffler, "Change of organization in the receptive fields of the cat's retina during dark adaptation," *J. Physiol.*, vol. 137, pp. 338–354; August, 1957.
[6] A. M. Andrew, "Report on Frog Colliculus," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept., pp. 77–78; July 15, 1955.
[7] A. M. Andrew, "Action potentials from the frog colliculus," *J. Physiol.*, vol. 130, p. 25P; September 23–24, 1955.
[8] H. R. Maturana, "The Fine Structure of the Optic Nerve and Tectum of Anurans. An Electron Microscope Study," Ph.D. dissertation, Harvard University, Cambridge, Mass.; 1958.
[9] Pedro Ramon Cajal, "Histologie du Systeme Nerveux," Ramon y Cajal, Maloine, Paris, France; 1909–1911.
[10] R. M. Gaze, "The representation of the retina on the optic lobe of the frog," *Quart. J. Exper. Physiol.*, vol. 43, pp. 209–214; March, 1958.
[11] R. Sperry, "Mechanisms of neural maturation," in "Handbook of Experimental Psychology," S. S. Stevens, Ed., John Wiley and Sons, Inc., New York, N. Y.; 1951.
[12] R. M. Gaze, "Regeneration of the optic nerve in *Xenopus laevi*," *J. Physiol.*, vol. 146, p. 40P; February 20–21, 1959.
[13] H. R. Maturana, J. Y. Lettvin, W. S. McCulloch, and W. H. Pitts, "Physiological evidence that cut optic nerve fibers in the frog regenerate to their proper places in the tectum," *Science*; 1959 (in press).
[14] G. Svaetichin and E. F. NcNichol Jr., "Retinal mechanisms for chromatic and achromatic vision," *Ann. N. Y. Acad. Sci.*, vol. 74, pp. 385–404; November, 1958.
[15] W. S. McCulloch and W. H. Pitts, "How we know universals. The perception of auditory and visual forms," *Bull. Math. Biophysics*, vol. 9, pp. 127–147; June, 1947.

Repetitive Analog Computer for Analysis of Sums of Distribution Functions*

F. W. NOBLE†, SENIOR MEMBER, IRE, J. E. HAYES, JR.†, AND M. EDEN†

Summary—**Many experimental procedures yield curves which are sums of distribution functions. Examples of such curves include electrophoretic, diffusion, and ultracentrifugal patterns, absorption spectra, and curves from countercurrent distribution and from partition chromatography in either liquid or vapor phase. In a given type of curve, each of the component functions is identical to the others in form (for example Gaussian) but can have very different values of the parameters governing height, width, and position along the abscissa. We wish to determine the parameters for each component by an analysis of the sum curve. The computer to be described performs this analysis by synthesizing a number of distribution functions of the desired form, each with adjustable parameters, and presenting, on an oscilloscope, the sum of these functions for comparison with the experimental curve being analyzed. A match is made visually by adjustment of the various parameters. When a match has been obtained, the parameters of the component functions are read out, following a switching procedure which presents the individual functions in sequence.**

Introduction

In numerous physico-chemical techniques, the result of an experiment is obtained as a curve which is the sum of a number of statistical distribution functions of a given type; for example, the normal and Lorentz distributions. In such cases, it is often of importance to obtain accurate values for the mean of each such function and for the area under it. For example, in a light absorption spectrum, the mean for a given one of the component absorption bands is a measure of the energy of the transition producing the absorption, while the area of the band is a measure of the probability of occurrence of that transition. Similarly in an electrophoretic or ultracentrifugal pattern, the mean of a given peak affords a measure of the rate at which the corresponding component moves in the force field applied while the area under the peak is proportional to the concentration of the component.

Consequently, it became of interest to design and construct a device capable of resolving such curves into their components. Since for a given experimental technique the shape of these component bands is assumed to be known at the outset, it is desirable that the analysis be made in terms of these known curves rather than, for example, sine and cosine functions as in the case of the Fourier analysis. Although a Fourier analysis could, of course, be performed, it is a more complicated procedure, and the result would not be in terms of functions having any direct significance in the systems being studied.

To our knowledge, previous attempts at solution of this problem have originated in two laboratories. The apparatus of French, *et al.*,[1] is an analog device using electromechanical generation of the functions to be used in fitting. That of Rogoff[2] is a digital machine which fulfills a somewhat different function. Its mode of operation is the linear combination of whole spectra to fit the spectrum of a mixture of substances rather than the resolution of a single spectrum into its component bands.

The machine to be described here generates in each of ten channels a distribution function of predetermined equation with parameters corresponding to height, width, and x-position continuously and independently variable. The sum of these functions is presented by means of an oscilloscope and the parameters of the individual curves are adjusted until their sum fits the experimental curve within the desired limits of error. The curves may then be presented individually for readout.

General Description

The two distribution functions of main interest are the Gaussian and the Lorentz. Both of these curves are mirror symmetric with respect to the ordinate through the peak but differ markedly in contour. In order to produce both types of waveforms in a single apparatus, we use a common generating waveform and provide an adjustable distortion device or function generator to shape the generating function into an approximation of the desired distribution function. The requirements of the generating function are that:

1) It must be of constant frequency so as to remain stationary on an oscilloscope synchronized to the master oscillator.
2) It must be continuously adjustable in x-position on the oscilloscope; *i.e.*, its phase must be continuously adjustable.
3) It must rise monotonically to a peak value and then descend mirror symmetric with respect to the ordinate through the peak.
4) It must be of constant height.
5) It must be of continuously adjustable width. The width variation must consist of a symmetrical expansion about a fixed x-position of the peak.
6) It must be simple to produce.

* Original manuscript received by the IRE, July 23, 1959. This paper was presented in part at the Second Internatl. Conf. on Medical Electronics, Paris, France, June 26, 1959.

† Lab. of Tech. Dev., Natl. Heart Inst., Natl. Institutes of Health, Bethesda, Md.

[1] C. S. French, *et al.*, "A curve analyzer and general purpose graphical computer," *Rev. Sci. Instr.*, vol. 25, pp. 765–775; August, 1954.

[2] M. Rogoff, "Automatic analysis of infrared spectra," *Ann. N. Y. Acad. Sci.*, vol. 69, pp. 27–37; September, 1957.

Possibly the simplest generating function is an isosceles triangle. Such a wave may be obtained as follows (see Fig. 1). The sinusoidal master oscillator voltage is passed through a continuously adjustable phase shifter to satisfy the continuous x-position requirement. The emerging sine wave is converted to an equal-interval square wave by conventional means, and this wave is integrated. The integrator output is a triangular wave of equal rising and falling slopes. The positive peak of the triangle is maintained at a fixed potential, and all values of the triangle which are negative with respect to a second fixed potential are clipped. The emerging wave then consists of a series of isosceles triangles of constant height with their bases connected by horizontal lines. Now if the amplitude of the input triangle is increased, the width of the output triangle will decrease in the required fashion. The frequency and height, of course, remain unchanged. If this generating waveform is connected to a function generator in which the output slope can be made to be the correct function of the input amplitude, the desired distribution function will be produced. As the width of the generating function is varied, the width of the distribution function will vary in direct proportion. The height variation of the distribution function is made by means of a gain adjustment on an amplifier following the function generator.

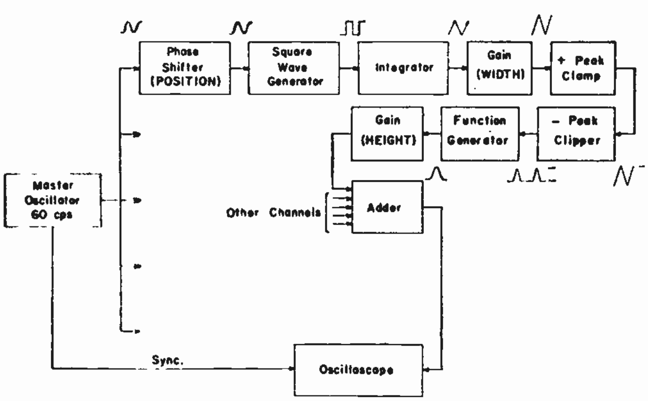


Fig. 1—Block diagram of the computer showing basic operations and waveforms.

A separate channel, as described above, is required for each component distribution function. Each channel is provided with a switch to connect or disconnect that channel from an adding network. The output of the adding network is connected to the oscilloscope, which accordingly presents the sum of all active channels. Since each channel has continuously and independently variable adjustments for height, width, and position of one component distribution function, it is possible to fit the sum curve appearing on the oscilloscope to an experimental curve by adjustment of the appropriate controls. When a suitable fitting has been made, the component curves may be presented one at a time and traced upon the experimental curve or otherwise read out. A special projection oscilloscope is used with the computer. The cathode-ray tube is mounted vertically above the operating table with the screen down. A standard enlarging lens focuses the CRT spot onto the table top. The focal length of the lens and the geometry are selected such that the linear span of the CRT covers an area on the table top roughly eight by ten inches. This arrangement permits the fitting of experimental curves on opaque paper directly without the usual photographic processing. Parallax is nonexistent. In addition, by adjustment of geometry, the projection oscilloscope permits utilization of the entire linear portion of the CRT screen for a curve of any size, thus permitting an optimum ratio of spot size to pattern size independent of the scale.

Circuits

To minimize flicker and provide adequate computing speed with respect to the speed of the human servomechanism, the 60-cycle line voltage is used as the master oscillator. All channels are driven by a single large 2.5-volt center-tapped filament transformer. As shown in Fig. 2, the sine wave can be reversed in phase by a switch and shifted continuously through 180° by the RC phase shifter. The output voltage of the shifter varies somewhat with shaft rotation, but this effect is not important because only the region near the axis crossing is used in succeeding circuits. The outstanding advantages of this circuit are that a full 180° range is obtained and that the phase shift vs shaft rotation is reasonably linear. V1 is a linear amplifier which decreases the rise time and phase jitter of the following square wave generator. Double triode V2 is a symmetrical clipper, the function of which is to produce a square wave of precisely equal positive and negative intervals. The potentiometer in its grid circuit adjusts the duty cycle of the square wave over a limited range. Diode V3a conducts for all values of its input signal which are negative with respect to ground, but does not conduct for positive values. Triode V3b cuts off long before the maximum negative value of its input is reached. Accordingly, the waveform at its plate is a fast-rising square wave. V4 (Fig. 3) decreases the rise time of the square wave and improves its flatness. In addition, V4b operates with the aid of V5 to stabilize the amplitude of the rectangle at the cathode of V5. Note that the return of V4 is to -150 volts. Throughout the half cycle when V4b is at zero bias, the potential of its plate is negative with respect to ground. Since the cathode of V5 returns to ground, it is cut off, making the potential at the cathode zero. Throughout the other half cycle when V4b is cut off, V5 conducts, thereby forming a voltage divider consisting of the plate resistor of V4b and the cathode-to-ground resistor of V5. This action sets the potential of V5 cathode at a very definite positive value. The absolute values of the positive and negative potentials of the square wave at V5 cathode are thus independent of the circuitry preceding V4. V6 and

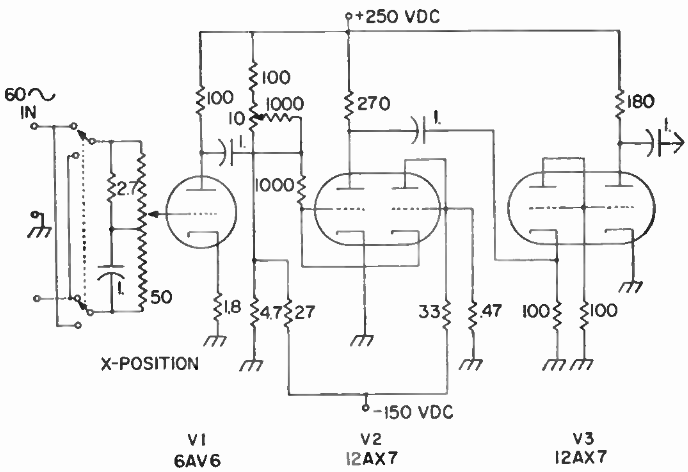


Fig. 2—Phase shifter and square wave generator circuits. In this and all following circuits resistor values are in thousand-ohm units and capacitor values are in microfarad units.

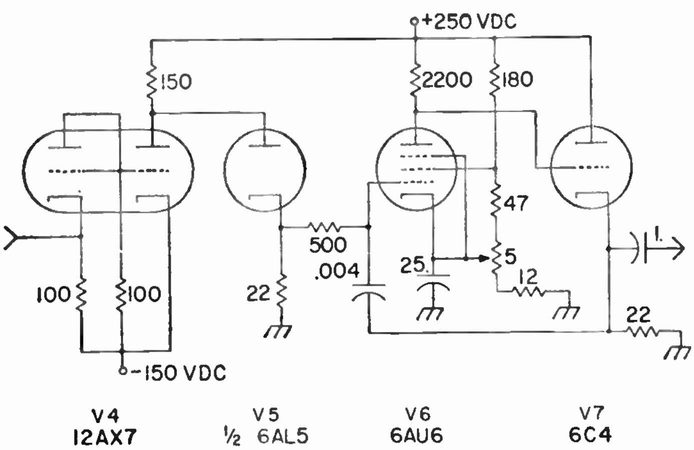


Fig. 3—Amplitude stabilizer and integrator circuits.

V7 together form a Miller integrator having excellent stability and large gain. The pentode is "starved," so that a very large plate resistor can be used to obtain gain. The bandwidth of the integrator amplifier without feedback must approach 10 kc if the distortion of the triangle is to be within acceptable limits. To this end, the capacitance in the plate circuit of V6 is minimized, and cathode follower V7 is used for its high input impedance. An inherent characteristic of the Miller integrator is that its output amplitude is nearly independent of gain changes in the amplifier, provided that the gain is large. Since we were careful to stabilize the amplitude of the input square wave, the output triangle from the integrator is very stable in amplitude. This is required by the following circuitry.

Diode V8a (Fig. 4) is a dc restorer which sets the negative peak of the triangle at the potential of the arm of the "ac/dc" control. This type of dc restorer invariably distorts the region of the curve where the diode conducts. The negative peak is restored rather than the positive peak because the region near the positive peak is later used to form the distribution function. To the

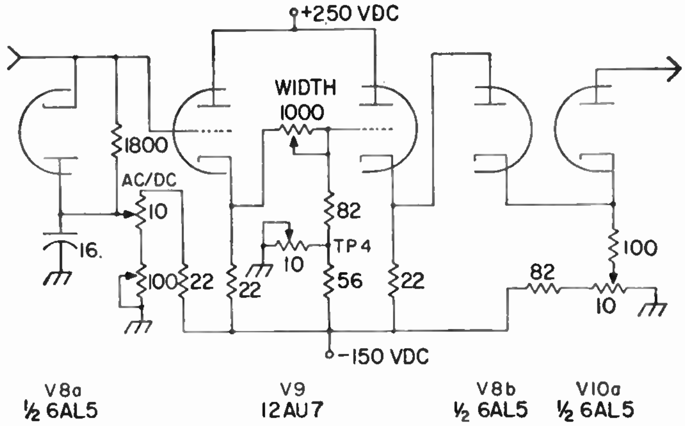


Fig. 4—Width control and negative peak clipper circuits.

extent that the input amplitude is fixed, the restorer sets the absolute potential of all points of the triangle at the grid of V9a. Triode V9a is a cathode follower which isolates the high-impedance dc restorer from the relatively low-impedance width control circuit. The object of the circuits of V8a and V9 is to hold the positive peak of the triangle at the cathode of V9b at a fixed potential, independent of the position of the width control. The signal at the grid of V9b may be considered to be the sum of a dc voltage and a triangular voltage. The positive maximum of this sum must at all times remain constant no matter where the width control is set. Now as the resistance of the width control is decreased, the amplitude of the triangle at the grid of V9b will increase. If the positive maximum is not to change, the dc voltage must go negative by precisely the same number of volts as the triangle amplitude has increased. This requires that the cathode of V9a be set to a potential which is negative with respect to TP4 by the correct amount. The potential at the cathode of V9a is adjusted by means of the "ac/dc" control, which is so named because it adjusts the ratio of ac to dc flowing in the width control. This procedure avoids the necessity of tracked potentiometers, one carrying the triangle and the other carrying dc. The width control performs both functions and has zero tracking error. The unconventional connection of the width control is used to make the width variation more nearly linear with shaft rotation than would be the case with the standard gain control circuit.

Diodes V8b and V10a set the potential at which negative peak clipping of the triangle occurs. The waveform at the cathode of V9b is an isosceles triangle occupying one half cycle for its rise and the other half cycle for its fall and having its positive maximum potential fixed, independent of the width control. We wish to clip off all values of the triangle which are negative with respect to a fixed value and to substitute this fixed value throughout the clipping interval. This is done by biasing the cathode of V8b at the clipping level. The diode will conduct and pass all values which are positive with respect to the clipping level, and will cut off and thereby disconnect the input signal for values which are negative with respect to this level. The clipping, or minus level, is set by the potentiometer. Diode V10a conducts for all values of its input which are negative with respect to ground and cuts off for positive inputs. It is used to prevent the grid of the following stage from being driven positive during preliminary adjustments. Its more important use is to round off the positive peak of its output triangle, starting the forming procedure.

Pentodes V11 and V12 (Fig. 5) perform the major part of the distortion required to shape a triangle into a distribution function. Use is made of the fact that the plate current vs grid voltage characteristic of a vacuum tube is highly nonlinear in the vicinity of cutoff, and of the fact that in a pentode the character of the nonlinearity is affected by the screen voltage. The grid of V11 is run from zero to near cutoff bias by the input triangle. Accordingly, its plate voltage resembles an inverted Lorentz with an exaggerated point at the peak. V12 inverts this signal and provides further shaping in the vicinity of the positive peak of the distribution function.

Triode V13a is a cathode follower used to drive the dc restorer, V10b. This diode sets the base line at a potential just sufficiently negative with respect to ground so that the baseline at the cathode of V13b is zero. The height control adjusts the amplitude of the distribution function without varying the base line potential. V13b is a cathode follower which provides the low output impedance needed to drive the adding network. A resistor is connected to the adding network when the channel switch is off. It has approximately the same resistance as the cathode follower output resistance and thus prevents change in the adder scale factor as channels are switched in and out. There are two controls in the dc restorer circuit in addition to the height control. One is the zero control which is adjusted with the height control set at zero. Its function is to make the baseline at the cathode of V13b equal zero. The "Diode" control is set with the height control fully up. It provides the retarding field required to hold V10b out of conduction when both ends of the height control are at the same potential. It is adjusted to make the base line zero for all positions of the height control.

The adding network is an ordinary parallel resistive type employing 100 K precision resistors for all elements.

The high voltage power supplies must be carefully regulated for two reasons. First, the stability of all reference voltages used throughout the circuit can be no better than the stability of the high voltage supply, and second, reliance is placed upon the extremely low source impedance of regulated supplies to prevent coupling through the supply leads.

Fig. 6 indicates the machine-dictated decomposition of the ultraviolet-visible absorption spectrum of sodium 2,4-dinitrophenolate, fitted with Gaussian components. Although there is no good theoretical reason for believing that absorption bands in this region should be of accurately Gaussian shape, in our hands, as well as in those of other workers, this curve seems to afford the best fits. It will be seen that the spectrum is decomposable into absorption bands which can each be assigned a frequency and an intensity.

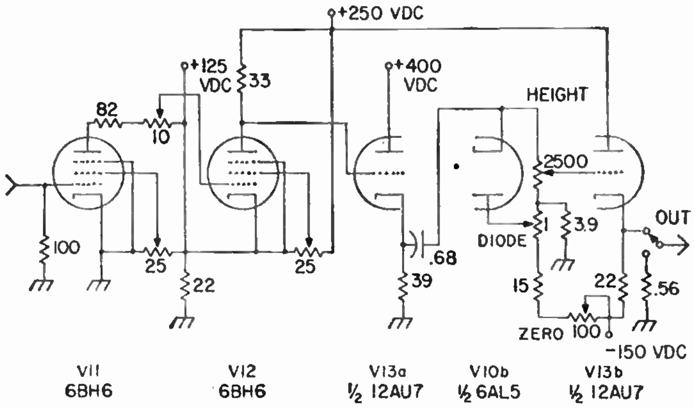


Fig. 5—Function generator and height control circuits.

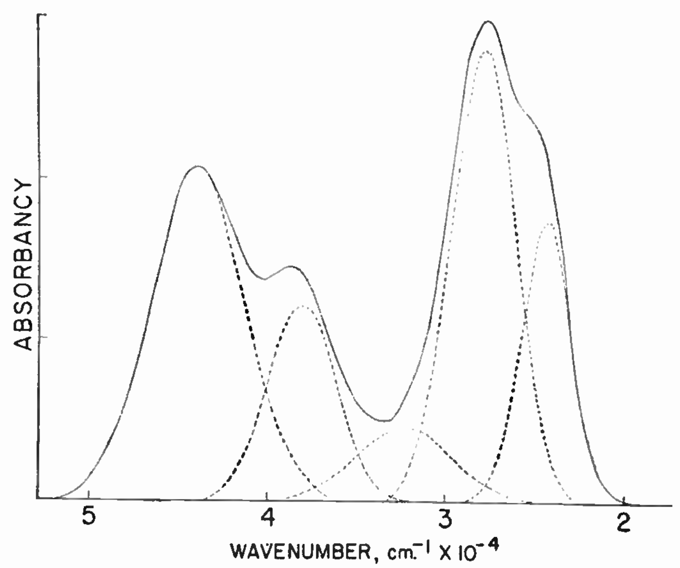


Fig. 6—Analysis of the ultraviolet-visible absorption spectrum of sodium 2,4-dinitrophenolate in terms of Gaussian components.

Fig. 7 shows a similar decomposition of the spectrum of iso-octane (2,2,4-trimethyl-pentane) in the three-micron region of the infrared. It will be noticed that in this case, the fitting function used is the Lorentz function. This function is generally considered to be the proper one for the shape of an infrared absorption band in the absence of instrumental artefacts.

Discussion

The fitting of optical absorption curves has been carried out for many years by laborious computations and plottings. Because of the trial and error nature of this fitting procedure, the time required to obtain a

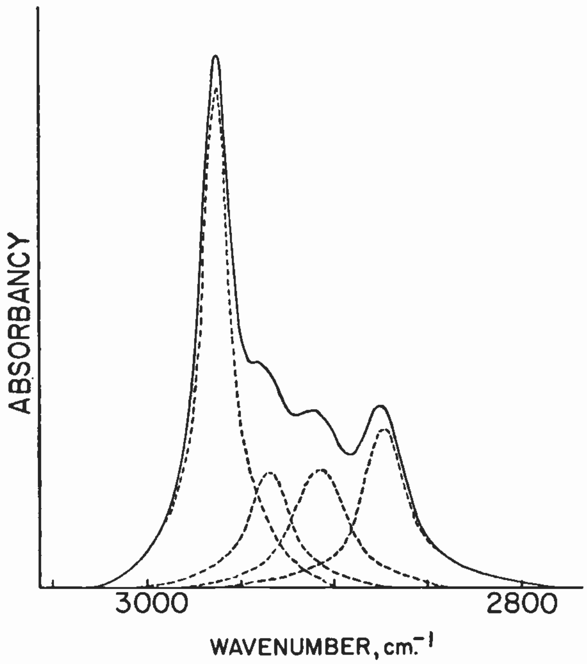


Fig. 7—Analysis of the spectrum of iso-octane (2,2,4-trimethylpentane) in the three-micron region of the infrared. The components in this case are Lorentz distributions.

reasonably accurate fit is so great that it is doubtful that the method will ever be applied to the reduction of large numbers of spectra. The method used in this computer is very similar in principle in that it is really a trial and error procedure. The contribution of the machinery is in making computations and hand plotting unnecessary, thus greatly reducing the labor and increasing the speed of the operation. The computer presents a sum curve many times faster than the operator can change a single parameter of a component curve. Hence the time required to accomplish a good fit is determined by the adroitness of the human operator in setting a group of knobs.

The uniqueness of solutions obtained by the use of the computer is theoretically identical to that obtainable by computation and plotting procedures if these procedures were carried to their limit. It is not clear to the writers by what means the uniqueness of solutions might be formally proven. Some measure of the sensitivity of the device to known alterations in wave form would seem helpful in this regard. To this end, a series of waveforms consisting of sums of two Gaussian functions were plotted on graph paper and fitted on the computer.

In the data below, σ stands for the variance and H for the height, and the mean is the x position of the peak. The conditions are for the detection of the presence of a second peak.

for $\sigma_1 = \sigma_2$:

$H_1 = H_2$, means must be greater than 1 σ_1 apart,
$H_1 = 3H_2$, means must be approximately 1 σ_1 apart,
$H_1 = 10H_2$, means must be greater than 1 σ_1 apart;

for $\sigma_1 = 1.5\,\sigma_2$:

$H_1 = H_2$, means can be identical,
$H_1 = 3H_2$, means must be greater than 0.4 σ_1 apart,
$H_1 = 10H_2$, means must be greater than 0.4 σ_1 apart;

for $\sigma_1 = 3\sigma_2$:

$H_1 = H_2$, means can be identical,
$H_1 = 3H_2$, means can be identical,
$H_1 = 10H_2$, means can be identical.

Objective methods for obtaining a fit and for judging its quality are under consideration. The experimental curve can be generated at the proper repetition rate by means of a photoformer. This function can then be subtracted from the sum of distribution functions coming from the computer. The difference can be read by a square-law voltmeter and the reading of this meter may be minimized by adjustment of the various parameters. When a minimum has been obtained, the data curve has been fitted in the least squares sense and the meter reading is a measure of the quality of the fit. Alternatively, the absolute value of the difference can be minimized, at which point the data curve is fitted in the minimum area difference sense and the reading of the indicator is a measure of the quality of this type of fit. At the present time it does not appear to be possible to generate the data curves with enough accuracy so that fitting by such means can be expected to improve the accuracy obtainable by a visual fit.

An artefact called slit error in infrared absorption spectrometers leads to spectral lines which are intermediate between the theoretically correct Lorentz and a Gauss function. The computer can be made to produce distribution functions of a variety of contours subject to the following restrictions: the function must be capable of being generated by means of a triangle fed to a function generator in which the output amplitude is a single valued function of the input amplitude. The generating triangle may have unequal rising and falling slopes, so that this type of skewed distribution may be generated. The equation of the desired function need not be known, since the function generator is adjusted to fit a drawn curve of the desired contour.

Sometimes the experimental data is presented in terms of sums of derivatives or sums of integrals of distribution functions. In these cases, fitting is accomplished by either differentiating or integrating the sum curve coming from the computer. The readout may be in terms of the derivative form, the function form itself, or the integral form of the component distribution functions merely by performing the appropriate operation on the sum curve coming from the computer.

In summary, we feel that we have demonstrated the practicability of obtaining information inherent in several types of experiments but heretofore unobtainable. In the near future we hope to present more concrete examples of the application of this device to the analysis of absorption spectra.

MEDICAL ULTRASONICS*

Introduction**

J. F. HERRICK†, MEMBER, IRE

THIS PAPER will attempt to give our readers a reliable report of the present status of ultrasound in the broad field of medical research. The reliability is based on the fortunate circumstance that each of the selected subdivisions of this paper was written by an investigator actively engaged in the particular aspect reported upon. Moreover, each has been doing active research for several years in the aspect in question.

No attempt will be made to give a historical account of the development of medical ultrasonics or a complete review of the literature. References pertinent to general topics are already available,[1-3] and may be found in the accompanying bibliography. Like some other agents which have been introduced into the armamentarium of clinical medicine, medical ultrasonics passed through the early stages of enthusiasm followed by a reactionary stage of pessimism before it achieved the status presently accorded it. Currently, there are promising developments and interesting applications of ultrasound for medical diagnosis, for therapy and for biologic measurements.

The interestingly intrinsic characteristics of ultrasound give this agent a unique position in medical applications. No other agent appears to be so conveniently sensitive to the discrete interfaces which occur so abundantly in biologic systems. In fact, it was this characteristic which prompted Professor Freundlich and his associates to publish a paper[4] suggesting the possibility of ultrasonic diathermy. Experiments in our laboratory at the Mayo Clinic and Mayo Foundation which were prompted by Freundlich's suggestion have resulted in an extensive study of the effect of ultrasound on bone. Bone presents a particularly interesting interface when ultrasound is being propagated through bodily tissues. As a result of some 10 years of investigations, we now entertain the hope that ultimately ultrasound can be used for inhibiting the growth of certain types of bone tumors. The Russian reports already describe the application of high-intensity ultrasound in the treatment of tumors.

Obviously, the introduction of any new agent into medical practice should be preceded by a thorough and extensive study of the biologic effects of that agent. Such studies in respect to ultrasound have been active ever since the early reports in 1927 by Wood and Loomis.[5] The literature in this field may be confusing to the uninitiated reader, since all possible results—from no effect to highly destructive effects—have been reported without an adequate description of the conditions for achieving some of these results. Obviously, a given result depends upon several experimental factors, such as the intensity of the incident acoustic energy, the duration of exposure to ultrasound, the frequency of the ultrasound, the coupling agent employed, the technique of application (continuous or pulsed, and also a means of directing the energy correctly), and acoustic characteristics of the tissues as well as physiologic factors. The measurement of the intensity of ultrasound *in vivo* is particularly difficult. Many of the early results which were reported could not be confirmed because of lack of adequate information about the conditions for obtaining those results.

* It is regretted that the manuscript "Fundamental Neurological Research and Human Neurosurgery Using Intense Ultrasound," describing the important contributions to Medical Ultrasonics by Prof. William J. Fry was not available for publication in this special issue of the PROCEEDINGS. However, the following reference is recommended highly: William J. Fry, "Intense Ultrasound in Investigations of the Central Nervous System," Cornelius A. Tobias and John H. Lawrence, "Advances in Biological and Medical Physics," Academic Press, Inc., New York, N. Y., vol. 6, pp. 282–347; 1958.

** Original manuscript received by the IRE, September 1, 1959.
† The Mayo Foundation, University of Minnesota, Rochester, Minn.

[1] R. Pohlman, "Die Ultraschalltherapie; Praktische Anwendung des Ultraschalls in der Medizin," Verlag Hans Huber, Bern, Switzerland; 1951.

[2] P. A. Nelson, J. F. Herrick, and F. H. Krusen, "Ultrasonics in medicine," *Arch. Phys. Med.*, vol. 31, pp. 6–19; January, 1950.

[3] J. F. Herrick and F. H. Krusen, "Ultrasound and medicine. A survey of experimental studies," *J. Acoust. Soc. Amer.*, vol. 26, pp. 236–240; March, 1954.

[4] H. Freundlich, K. Söllner, and F. Rogowski, "Einige biologische Wirkungen von Ultraschallwellen," *Klin. Wochschr.*, vol. 11, pp. 1512–1513; September, 1932.

[5] R. W. Wood and A. L. Loomis, "The physical and biological effects of high-frequency sound-waves of great intensity," *Phil. Mag and J. Sci.*, vol. 4, pp. 417–436; September, 1927.

A quantitative understanding of the biologic effects of ultrasound is a major objective of any scientific inquiry in this field. Measurements of a number of physical variables must be accomplished. These include the acoustic impedance of bodily tissues, the velocity of sound as a function of temperature of the tissues, and attenuation characteristics, including the depth of penetration as a function of frequency and the radiation pattern of the acoustic field in the tissue. Dr. Herman Schwan has devoted considerable effort to the study of the acoustic characteristics of tissues and their constituents. His work on the mechanisms of absorption of ultrasound in cellular suspensions and by subcellular elements is outstanding. He reports here on the major contributions to these aspects of medical electronics accomplished by his laboratory.

Professor Fry and his associates have spent 10 years investigating the effects of focused high-energy ultrasound on the tissue components of the central nervous system. They have discovered that selective irreversible (permanent) and reversible (temporary) changes can be produced in brain structures without affecting other tissue by focusing the sound into the region of interest. To accomplish this work, Fry and his collaborators have designed and developed instrumentation for accurately controlling and applying high-intensity ultrasonic radiation to biologic systems. As the direct result of their researches, high-intensity ultrasound is now being applied to the study of basic brain mechanisms in experimental animals, and Professor Fry has, in collaboration with neurosurgeons, initiated work on the treatment of a number of neurologic disorders of human beings.*

The application of ultrasound for medical diagnosis may in certain instances prove to be superior to the well-known X rays. John Reid has spent many years in the development of instrumentation for medical diagnosis and in the application of this instrumentation with his medical associates. He will report the results of his studies and also will review the results of similar studies by other investigators.

Generally speaking, ultrasound was introduced into clinical medicine for use in therapy. Departments of physical medicine and rehabilitation became interested in the claims which were publicized for this relatively new agent. In fact, the earliest equipment available commercially was designed for use in therapy. The literature contains numerous reports by many physicians in several countries. Some of these reports described unusual beneficial results, some of which any physician would be happy to confirm. During the wave of enthusiasm for the technique, it was reported that 5350 ultrasonic devices were being employed by physicians in western Germany alone. If this were true, there was one ultrasonic machine in medical use for every six physicians in that country. Annual international congresses entitled "Ultrasound in Medicine" were being held and much enthusiasm was evident at these meetings. Manufacturers of equipment were promoting ultrasonotherapy by supplying tabular data on the intensity of ultrasound to be used in specific diseases, on the duration of exposure and on the number of exposures required. The frequencies employed for therapy ranged from approximately 800 kc to 1 mc. The early equipment permitted rather high outputs. Following the wave of enthusiasm, a Standards Committee reduced the output permitted for use in therapy, and now manufacturers are producing standard equipment in accordance with the specifications which the excellent Standards Committee has prepared so carefully.

During the wave of enthusiasm for ultrasonotherapy, there is reason to believe that some operators of the equipment did not appreciate the propagation characteristics of ultrasound. This circumstance may have resulted in delivering little or no ultrasonic energy at the site at which the application was desired. Results, which were reported under these conditions, engendered even greater confusion among those who were attempting to review the literature.

Finally, this report will describe the application of ultrasound to biologic measurements. Physicists and engineers who are trained in the art and science of measurement may become discouraged when applying their skills and experience to living organisms. Nature can play unforeseen tricks when the instrumentation for measurement upsets her pattern of performance. In the early attempts at measurement, some investigators forced the tools already available for measurement upon the living organism and obtained good repeatable data —but under abnormal conditions. How much better it would be to spend time in the careful observation and analysis of normal processes so that equipment and techniques could be designed, developed and finally applied under conditions which preserve as closely as possible the natural processes. Thus, a more reliable quantitative understanding of the amazing biologic mechanisms could be attained. To cite a specific example: for several years (in fact for more than 100 years), knowledge of the correct relationship between the pressure of the blood and the flow of blood has been desired. Even today there is no ideal method for measuring either pressure or flow continuously in such a way that this relationship can be ascertained under conditions which do not interfere in any way with the normal pattern of performance.

We entertain the hope that ultrasound may prove useful in making intravascular measurements without disturbing the normal pattern, because ultrasound can be coupled to the blood through the wall of the unopened blood vessel. Methods of measuring the velocity of blood by means of ultrasound will be reported in some detail. Other measurements employing ultrasound will be described.

Absorption of Ultrasound by Tissues and Biological Matter*

H. P. SCHWAN†, FELLOW, IRE

Summary—**General principles which determine the frequency dependence of both absorption and velocity of ultrasound in matter are outlined and applied to cell suspensions and tissues. The mechanisms which are responsible in the biological case for the experimentally observed frequency dependence of ultrasonic properties are described. They relate predominantly to macromolecular components. Finally, the relationships which pertain to the propagation of ultrasound in heterogeneous tissue complexes are discussed, and consequences for the medical application as a therapeutic tool are considered.**

Both electromagnetic radiation and ultrasound have found increasing use for therapeutic purposes after World War II. While the electrical properties of body tissues and pertinent applications have been summarized in another place, it is intended in this chapter to outline our present knowledge about absorption of ultrasound in tissues and to see how such knowledge pertains to medical applications.[1]

Absorption and Velocity of Ultrasound in Tissue and Biological Matter

The mode of propagation of sound in matter is completely characterized by its coefficient of absorption and its velocity. The absorption coefficient is best defined by

$$I = I_0 \exp(-\mu x) \tag{1}$$

where I_0 stands for initial energy, I for energy at a distance x from the initial site, and μ is the absorption coefficient. The equation assumes plane wave propagation, *i.e.*, considers only reduction of intensity due to true absorption processes and not due to spread of energy caused by divergent characteristics of the beam. The absorption coefficient is usually expressed either in terms of neper per cm in accordance with (1) or in decibels per cm. Changes of velocity and absorption with frequency reflect the presence of processes of energy transfer which require time. If such a process is characterized only by one relaxation time constant T,[2] both absorption and velocity follow the simple relationships

$$\alpha\lambda = a\omega + 2(\alpha\lambda)' \frac{\omega T}{1 + (\omega T)^2} \tag{2}$$

$$c^2 = c_0^2 + \frac{2}{\pi} (\alpha\lambda)' c_0 c_\infty \frac{(\omega T)^2}{1 + (\omega T)^2} \tag{3}$$

where ω is the angular frequency $2\pi f$. The absorption is expressed in terms of absorption per wavelength $\alpha\lambda$, a is indicative of the possible presence of frequency independent absorption processes, not related to the existence of a relaxation process, and $(\alpha\lambda)'$ is the highest possible value of $\alpha\lambda$. The indices $_0$ and ∞ in the velocity term c refer to very low or very high frequencies, respectively. The corresponding frequency dependence of the absorption per wavelength $\alpha\lambda$ is illustrated in curve a of Fig. 1. For a broad spectrum of time constants, representative of a uniform distribution of activation energies ranging from one value to another, both velocity and absorption coefficient appear less frequency dependent, as indicated by the curves b and c.[3] It is apparent that for a sufficiently broad distribution func-

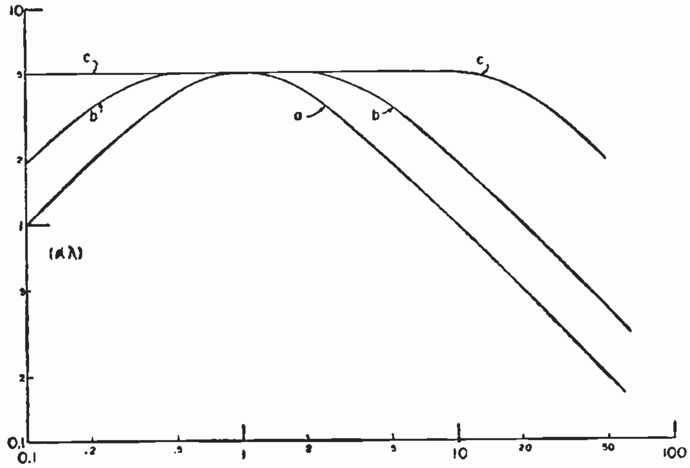


Fig. 1—Sonic absorption per wavelength ($\alpha\lambda$) as a function of frequency f. Assumed are relaxation mechanisms, which are characterized by distribution functions of activation energies of constant magnitude. (a) One time constant; (b) a moderate range of activation energies; (c) a wide range of activation energies. Both ordinate and abscissa are in arbitrary units.

* Original manuscript received by the IRE, September 1, 1959. This work was supported by the Office of Naval Research, Contract No. 119-289.

† Electromedical Div., The Moore School of Elec. Engrg., University of Pennsylvania, Philadelphia 4, Pa.

[1] H. P. Schwan, "Alternating current spectroscopy of biological substances," this issue, p. 1841.

[2] The time constant T can be defined, in general, as the time necessary for the response to a step function cause to cover $(1-1/e)$ of the difference between immediate and final response value.

[3] E. L. Carstensen and H. P. Schwan, "Acoustic properties of hemoglobin solutions," *J. Acoust. Soc. Amer.*, vol. 31, pp. 305–311; March, 1959.

tion of activation energies an almost linear frequency dependence of the absorption coefficient can be obtained over a very wide range of frequencies.

The frequency dependence of the absorption coefficient of most investigated tissues closely follows this behavior. It is characterized by a power function whose exponent varies between 1 and 1.3. Actual values of absorption have been reviewed by Goldman and Hueter[4] and are summarized in Table I for the three major classes of tissues.

TABLE I

ABSORPTION COEFFICIENTS AND VELOCITIES FOR MAJOR CLASSES OF BODY TISSUES

1) Absorption coefficient (neper per cm)	1 mc	2 mc	4 mc
Muscle (representative of all soft tissues with high water content)	0.3	0.5	1
Fat	0.07	0.2	0.6
Bone	0.6–2		
2) Velocity			
Muscle	1570 ± 20 m/sec		
Fat	1440		
Bone	3360		

The muscle data are representative of all soft tissues with high water content.

Application of (3) shows that the frequency dependence of the velocity of sound should be extremely small in comparison with the velocity itself. The validity of this statement should not be impaired if a one-time constant relation of the type given in (3) is replaced by a relationship which includes many time constants.[3] Indeed, experimental data support this point of view: The velocity is independent to a fraction of a per cent from frequency.[3] Actual values are also given in Table I. For a more detailed discussion of velocity data in tissue we refer particularly to Frucht's work.[5] The data in Table I demonstrate that the velocity of all soft tissues is near that of water. Differences up to 10 per cent reflect the influence of the solid components involved in the structure of tissue. The similarity of the velocity values with those of water is due to the high water content of most soft tissues.

MECHANISM OF ULTRASONIC ABSORPTION IN TISSUE AND SUSPENSIONS OF BIOLOGICAL CELLS

The statements formulated above, as related to the frequency dependence of absorption characteristics, explain observed data in terms of a broad spectrum of different processes and related activation energies. The observed frequency dependence thus appears well explained and reasonable, considering the complexity of biological structure. However, the advanced mathematical argument does not say anything about the origin of the related relaxation processes. Pertinent work will be summarized in this section.

The absorption of sound by blood has been investigated throughout the frequency range from 0.7 to 10 mc extensively by Carstensen and Schwan.[6,7] The results of this investigation demonstrate:

1) Most of the absorption of ultrasonic waves in blood arises from the presence of the protein content of the blood. This could be verified by showing that the predominant part of the absorption is in proportion to the over-all protein content and not to the cellular concentration.[6]
2) A small but noticeable fraction of the sonic absorption is due to the cellular organization of blood, *i.e.*, due to the presence of erythrocytes and, consequently, is not of molecular origin. This absorption contribution is due to the relative movement of the cells against the surrounding fluids which result from differences of specific weight. An observed decrease of this absorption contribution at high cellular concentration levels reflects the fact that the possibility of relative movement becomes inhibited in the immediate neighborhood of other cells.

Further data concerning the mechanism of absorption were obtained from an investigation of the acoustic properties of albumin.[8] The data obtained with albumin and hemoglobin seem to support the concept of an absorption which is independent of the particular macromolecule involved, if expressed in terms of absorption per gram protein matter in a given volume of water. The detailed hemoglobin study mentioned above[3] lends further support to the formulated mathematical analysis in terms of the presence of a broad spectrum of time constants. It was also possible to demonstrate that the small but measurable change in velocity with frequency agrees with values predicted from the absorption behavior of hemoglobin on the basis of a relaxation theory which assumes a broad spectrum of time constants.[3]

On the other hand, it became obvious that the blood data alone could not explain the mechanism of sonic absorption in tissues. Either another mechanism, other than the molecular one, which is not very noticeable in blood, contributes strongly to the tissue properties, or the tissue proteins must be assumed, specifically, to absorb much more strongly than albumin and hemoglobin in order to account for observed high tissue absorption

[4] D. E. Goldman and T. F. Hueter, "Tabular data of the velocity and absorption of high-frequency sound in mammalian tissues," *J. Acoust. Soc. Amer.*, vol. 28, pp. 35–37; January, 1958. "Errata," vol. 29, p. 655; May, 1957.

[5] A. H. Frucht, "Die Schallgeschwindigkeit in menschlichen und tierischen Geweben," *Z. ges. exp. Med.*, vol. 120, pp. 526–557; May, 1953.

[6] E. L. Carstensen and H. P. Schwan, "Absorption of sound arising from the presence of intact cells in blood," *J. Acoust. Soc. Amer.*, vol. 31, pp. 185–189; February, 1959.

[7] E. L. Carstensen, K. Li, and H. P. Schwan, "Determination of the acoustic properties of blood and its components," *J. Acoust. Soc. Amer.*, vol. 25, pp. 286–289; March, 1953.

[8] H. P. Schwan and E. L. Carstensen, "Advantages and limitations of ultrasonics in medicine," *J. Amer. Med. Assoc.*, vol. 149, pp. 121–125; May, 1952.

values (Table I). The latter conclusion found support in studies concerned with the absorption of cell nuclei in solutions.[9] In this case, the absorption per weight percentage solid compound (proteins and other nuclear molecules) was found to be very much higher than in the case of albumin and hemoglobin. A detailed study of liver tissue substantiated this result.[10] Liver tissue was subjected to repeated fractionation so that, at first, a rather coarse separation of the tissue into smaller cell complexes was accomplished. Then the division was further carried down in progressive steps, so that, eventually, even subcellular components, such as mitochondria and cell nuclei, were destroyed. Finally, a a solution of liver tissue proteins, which is free of all cellular or subcellular organization, was obtained. Throughout this process, ultrasonic absorption and velocity were monitored under a variety of conditions. The following conclusions were reached:

1) At least 80 per cent of the total tissue absorption is of molecular origin, *i.e.*, remains unchanged as the destruction of tissue structure is advanced step by step. The small change of about 20 per cent of the absorption takes place at the very initial step of tissue grinding and therefore is due to the initial existence of relatively large structures, such as pieces of connective tissue, etc.
2) Changes of the molecular constituents may be affected by variation in hydrogen ion concentration (pH). While the absorption coefficient responds to these changes in a very pronounced way, the characteristic of the frequency dependence is not affected. This indicates that the affected changes do not alter the fact that a broad spectrum of activation energies participates in the absorption mechanism. Other changes in the absolute value of the absorption coefficient can be affected by thermal denaturation of the tissue proteins.

In summary, by far the largest part, of the ultrasonic absorption of tissues arises on a macromolecular level. It is due to the existence of a broad spectrum of thus far unspecified transfer processes associated with the presence of the macromolecules. While the specific absorption per weight percentage macromolecule content is similar for albumin and hemoglobin, it is much larger for tissue proteins. It is also subject to considerable change with the chemical environment of the proteins and denaturation.

Propagation of Sound in Heterogeneous Tissue Complexes

One of the foremost applications of ultrasound is in physical medicine. It is utilized to affect heating and evoke consequent physiological responses well below the surface of the human body. This is found to be beneficial in the treatment of a variety of rheumatic and arthritic conditions. The superiority of ultrasound for this purpose in comparison with other forms of diathermy *i.e.*, ultrashort wave and microwave therapy, may be best illustrated by means of Fig. 2.[11] In it the

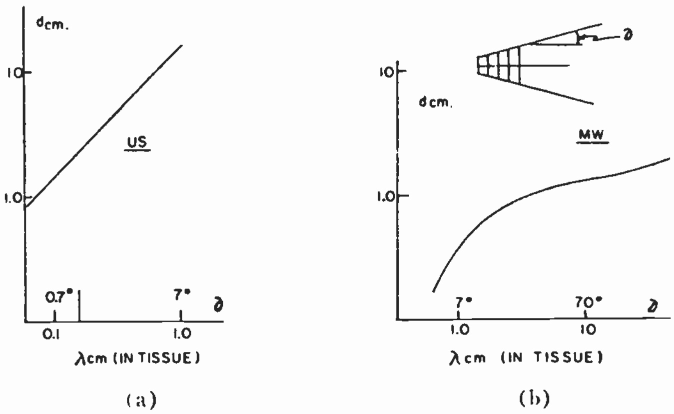


Fig. 2—Depth of penetration vs beam divergence for ultrasonic (US) and microwave (MW). The data pertain to muscular tissue as the medium of propagation. A reflector diameter of 5 cm is assumed in both cases. (Reprinted by courtesy of the *American Journal of Physical Medicine.*)

penetration, which can be achieved with either of these forms of energy, is plotted on the ordinate and the angle of divergence, which characterizes the beam of either form of radiation, on the abscissa. The divergence pertains to the "distant field." It is only well defined sufficiently far from the source so that it is meaningful to characterize a monotonously declining field in terms of its constant spreading characteristics.[12] The data refer to "piston" sources, *i.e.*, are based on the assumption of a plane source of radiation from whose individual parts wavetrains originate all equal in intensity and phase. It is assumed further, that the size (diameter) of the field spread, *i.e.*, the smaller angle of divergence, is achieved with larger diameter, or aperture, of the source of radiation. The penetration data pertain to muscular tissue as a representative for all tissues with high water content. Fig. 2 demonstrates that it is much easier to combine good penetration with practical beam definition with ultrasound than with electromagnetic radiation.

Another major advantage of ultrasound results from the aforementioned observation that the velocity of sound is nearly the same for all soft tissues. The characteristic impedance Z, which is determined by the ratio of sound pressure and particle velocity, is predom-

[9] A. Smith and H. P. Schwan, "Ultrasonic absorption and velocity of sound of cell nuclei," *Proc. Natl. Biophysics Conf.*, Columbus, Ohio; March, 1957. (Abstracts, p. 66.)

[10] H. Pauly, "Absorption of Ultrasound in Biological Media," presented at the Internatl. Conf. on Ultrasonics in Medicine, Los Angeles, Calif.; September 6–7, 1957. (Abstract.)

[11] H. P. Schwan, "Biophysics in diathermy," in "Therapeutic Heat," S. Licht., Ed., E. Licht, New Haven, Conn., vol. 2, pp. 55–115; 1958.

[12] In the "distant" field the intensity decreases inversely with the square of the distance, due to the constant spread of the beam. Interference of different waves, originating from different parts of the radiant source, cause periodic fluctuations in the "near" field. The separation between near and distant field may be defined by the distance D^2/λ where D is the diameter or the effective aperture of the radiant source, and λ the wavelength of the radiation.

inantly determined by the velocity and is only to a minor extent dependent on the absorption coefficient, for the values found to be typical for tissues. Hence, any fractional reflection of energy

$$\frac{Z_1 - Z_2}{Z_1 + Z_2},$$

which arises at the interface separating two different types of soft tissue, will be a very minor one for sound propagated into the human body.[13] This, by no means, holds for the case of electromagnetic radiation, as pointed out in another place in this volume.[1] The distribution of heat sources resulting from the application of a plane wave of ultrasonic energy traveling through various tissues, such as skin, subcutaneous fat and muscle tissue, consequently, can be predicted immediately from the absorption coefficients of these tissues. This is done simply by allowing for the attenuation of the sound wave which takes place in each tissue and realizing that the differentiated form of (1)

$$-\frac{dI}{dX} = I\mu I_0 \cdot \exp \cdot (-\mu x) \qquad (4)$$

is equal to the heat rate. Pertinent calculations have been carried out by Schwan, Carstensen and Li.[14] Fig. 3 summarizes some typical data which result from such calculations and pertain to the ratio of the heat developed in tissues beneath the subcutaneous fat layer to the total heat made available from the radiant source to the body, *i.e.*, including the heat developed in the subcutaneous fat layer. This ratio has been termed "depth efficiency"[11] since it characterizes to what extent the desired goal of developing as much energy as possible in the "deep" tissues beneath the subcutaneous fat is realized. The heating in the fat is considered unimportant and undesirable since it cannot be utilized to evoke physiological responses, such as dilatation of blood vessels to the same extent possible in muscular tissues. Fig. 3 demonstrates that an ultrasonic frequency of about 1 mc, as presently applied in clinical practice, has an excellent depth efficiency, while the microwave therapy used presently, operating at 2450 mc, fares very poorly in comparison.[1]

The application of physical forms of energy for heating purposes in medicine demands that the power indicated and produced by the generating equipment can

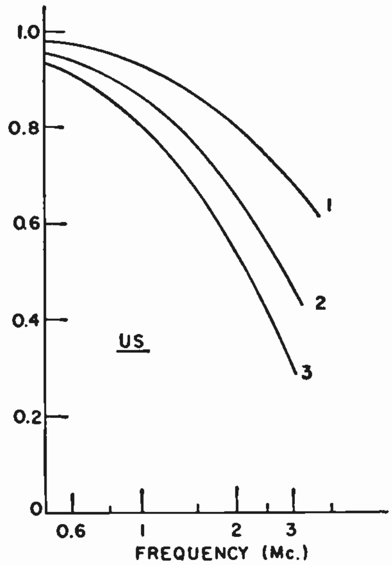


Fig. 3—Depth efficiency of ultrasound (ratio of energy which reaches deeper situated tissues to the total energy absorbed by the body). The parameter indicates the thickness of the subcutaneous fat layer in cm. (Reprinted by courtesy of E. Licht, publisher of "Therapeutic Heat," S. Licht, Ed., New Haven, Conn., 1958.)

be identified, or is, at least, simply related to the energy absorbed by the patient. If this is not so, difficulties arise in relating instrument indicated and biological effective power. Reproducibility of patient dose and dose rate may then be seriously impaired, making it impossible to operate under controlled conditions. It has been shown that the amount of electromagnetic radiation which is absorbed by the human body is, in a very complicated manner, related to the totally available power.[15] This is a consequence of the fact that, depending on frequency and thickness of such tissues as skin and subcutaneous fat, the input wave impedance of the human body can vary between the wave impedance of air and values which are different from this by as much as one order of magnitude. Ultrasonic energy, on the other hand, if properly applied, is completely transmitted to the human body,[16] establishing another advantage of ultrasound in comparison with microwave therapy.

[13] This does not apply to the boundary between soft and hard tissues such as bone. Indeed, strong reflections and selective boundary heating take place under such circumstances. They are also partially due to the transformation of normal and longitudinal sound waves into rapidly absorbed transversal "shear" waves.

[14] H. P. Schwan, E. L. Carstensen, and K. Li, "Heating of fat-muscle layers by electromagnetic and ultrasonic diathermy," *AIEE Trans. on Commun. & Electronics*, pp. 483–488; September, 1953.

[15] H. P. Schwan and K. Li, "The mechanism of absorption of ultrahigh frequency electromagnetic energy in tissues, as related to the problem of tolerance dosage," IRE TRANS. ON MEDICAL ELECTRONICS, no. PGME-4, pp. 45–49; February, 1956.

[16] To achieve this, it is necessary to avoid having even very thin layers of air present between the sound transducer and the body surface. The sonic impedance of air is very different from that of the tissues and would establish a major impedance discontinuity and, thereby effectively hinder the transfer of sonic energy. The necessity of maintaining good mechanical contact throughout the application of sonic energy is a disadvantage of sonic therapy. For a more detailed comparison we refer to Schwan, *op. cit.*, footnote 11. Also see H. P. Schwan, "The biophysical basis of physical medicine," *J. Amer. Med. Assoc.*, vol. 160, pp. 191–197; January, 1956.

Diagnostic Applications of Ultrasound*

JOHN M. REID†, MEMBER, IRE

Summary—**Current research on the use of low-power ultrasound as a diagnostic tool is reviewed. Ultrasonic waves can be used to investigate soft tissue structures which are opaque to light or invisible to X rays. Continuous-wave techniques can picture absorbing or reflecting areas in tissue specimens. Doppler frequencies related to the motion of the heart have also been investigated. Pulse reflection techniques can detect the small echoes reflected from interfaces between and within tissue structures. Radar scanning tecnhiques are used to form reflection pictures of accessible tissues. The significance of such pictures of the breast, bowel, eye, extremities, kidney, liver, and neck areas in the living human is being studied. The position and velocity of the walls of the heart during the cardiac cycle also can be recorded.**

Introduction

A number of investigators have sought to obtain information of diagnostic value by the use of ultrasound. Their methods are based upon the effects of absorption, reflection and refraction of ultrasonic waves by body tissues. At present, X-ray radiation is being used in this way to produce pictures of absorbing tissues. Ultrasound can be used to form pictures of absorbing and reflecting tissues which are invisible to the X ray, since the propagation constants of the two forms of energy differ fundamentally.[1] The picture detail obtainable with ultrasound generally will be inferior to that with X rays or even with visible light because the wavelength at usable frequencies is much longer. In some cases, a true pictorial display may be discarded and only a meaningful representation sought.

The medical application is similar to the now-familiar use of ultrasound in industrial nondestructive testing. The nondestructive advantage is retained, since the power levels used are far below those used in therapy or any level known to produce damage to tissue.[2] In pulse apparatus with an average power less than one watt per square centimeter, the peak power can reach several tens of watts per square centimeter for a microsecond without any demonstrable effect. The damage threshold at duty cycles of the order of 10^{-3} has not yet been found. Because of the safety of the methods, most of the applications have been directed at diagnosis of the intact human being. It has not been necessary for the patient to ingest or to have injected additional material.

Continuous wave transmission and reflection and pulse reflection techniques have been used to date. We confine "pulse techniques" to the use of short pulses to obtain time separation of the reflections of interest. Pulsing may be used with continuous waves to obtain free field conditions without the use of anechoic chambers. The choice between methods is made on the basis of the information desired and the properties of the tissues involved.

Acoustic Propagation in Tissue

Small reflections are produced between tissue structures of different acoustic impedances (ratio of excess pressure to particle velocity) in the same manner that reflections are found at impedance discontinuities on a transmission line. The acoustic impedance of most soft tissue approximates that of water, and the differences in impedance are not large. However, pulse apparatus in use can detect a signal about 100 db below the transmitted pulse. A signal of this magnitude would be reflected by a plane interface between tissues whose acoustic impedances differ by two parts in 10^5. The average velocity of propagation is near 1500 meters per second for soft tissue, so that the wavelength ranges from 1.5 to 0.1 mm over the frequency range of 1.0 to 15 mc. Some refraction may be caused by body fat since its propagation velocity is about 10 per cent less than other tissues. Bone has about twice the average velocity, but its principal effect on ultrasound is to reflect most of the incident wave and absorb the energy that does penetrate. The absorption for ultrasound is about 1 db per centimeter per megacycle in soft tissue. This rather high value limits the frequency that can be used for a given depth of penetration. The highest practical frequency generally is used to obtain the best resolving power. The absorption in fat and body fluids is lower than that in muscle and organs, and transmission methods have been tried to detect areas of differing absorption. The absorption for shear waves is higher than for compressional waves so that if shear waves, which travel at a different velocity, are produced at reflecting interfaces they are damped out within a short distance.

Continuous Wave Techniques

The first application to diagnosis was made by Dussik using continuous wave transmission.[3] The apparatus was used to scan the head in a manner directly analogous to X-ray transmission. The sound was transmitted by one directive transducer and, after passing through the head, received by another. Mechanical motion moved the pair to cover the area of interest while the received signal strength was plotted to form a picture of regions with greater sound transmission. It was

* Original manuscript received by the IRE, September 1, 1959.
† Electromedical Div., The Moore School of Elec. Engrg., University of Pennsylvania, Philadelphia, Pa.
[1] H. P. Schwan, "Absorption of ultrasound by tissues and biological matter," this issue, p. 1959.
[2] L. A. French, J. J. Wild, and D. Neal, "Attempts to determine harmful effects of pulsed ultrasonic vibrations," *Cancer*, vol. 4, pp. 342–344; March, 1951.
[3] K. T. Dussik, "Possibility of using mechanical high frequency vibrations as a diagnostic aid," *Z. Neurol. Psychiat.*, vol. 174, pp. 153–168; 1942.

hoped that the picture would show the shape of the liquid-filled ventricles in the brain, since brain tumors can be located if they are large enough to displace the ventricles. The ventricles can be visualized with the X ray only by draining them and injecting air to obtain sufficient contrast, which entails some risk to the patient. This work was continued by a group at M.I.T., but was finally abandoned.[4] Their conclusion was that the signal-strength variations caused by the absorption and refraction of the skull bones were so large that the variation due to the ventricles was completely masked, despite elaborate compensating means.

Satomura has reported success in diagnosing heart ailments with a continuous wave reflection technique.[5] Sound waves at a frequency of 3 mc are transmitted into the heart between the ribs where the heart contacts the chest wall. The receiving equipment is connected to a transducer coaxial with that used for transmission and detects the Doppler frequency shift caused by motion of reflecting structures in the direction of propagation of the sound waves. He identifies shifts of 500 cps and less (velocity about 10 cm per second) as arising from the motion of the heart walls. Wall velocities in this range have also been found using pulse reflection apparatus.[6] Satomura also found a band of frequencies centered at 1000 cps, and in experiments with dogs, he concluded they were caused by motion of the valves. An interesting Doppler component covering a wide band of frequencies was found to be especially strong in diseased hearts. The different frequencies are separated by filters and displayed in synchronism with the electrocardiogram. Diagnosis is made on the basis of the relative times of occurrence of the Doppler components.

A continuous wave system of considerable interest has been investigated by Suckling, although it has not been applied to any specific diagnostic problem.[7] It can be either a reflection or a transmission system, but it is most nearly an analog of an optical imaging system. A transmitting transducer serves to illuminate the specimen, and either the transmitted or reflected sound waves are focused by an acoustic lens onto a receiving quartz crystal which has only one face covered by an electrode. The acoustic pressure distribution across the face of the quartz is thus representative of the energy passing through or reflected from a plane inside the specimen. The charge distribution on the bare face of the crystal was found to be directly related to the pressure distribution. The acoustic spreading within the crystal, covering a few square millimeters of area, determines the resolution of the system, although the resolving power of the lens must also be considered. The picture is read out by scanning the back of the crystal with a small capacitive probe and is displayed on a cathode-ray tube with synchronized sweep. Several biological specimens have been pictured by this means. An acoustic transducer using the same principles with readout by means of an electron beam has been described by Sokoloff.[8] This tube offers advantages in speed and convenience. Since the generation of ultrasound in the kilomegacycle frequency range has been reported, the construction of an "ultrasonic microscope" may be possible. At 3 kmc, the wavelength will be in the same range as visible light if the velocity remains near 1500 meters per second. To avoid excessive loss, the specimens would have to be extremely thin as is the case with the optical microscope. Dunn and Fry have reported an ultrasonic microscope which does not use a lens but has a small thermocouple next to the specimen for detection of the transmitted energy.[9]

Pulse Reflection Techniques

A number of investigators have found that pulsed ultrasound can be used to distinguish various soft tissue structures from each other. Such structures produce small reflections at the interfaces between and within them. These reflections are so small that they have little effect upon the total attenuation, but they can be observed with the time separation afforded by pulsing. It is rather fortunate that the reflection coefficients are small, since the structure of soft tissues can be extremely complicated, and any multiple reflections between the interfaces must be less than the direct reflection by a factor approximately equal to the reflection coefficient. Transducers of reasonable size have directive radiation patterns which can be made even narrower by the use of weak focusing lenses. The directivity allows the spatial position of reflecting structures to be determined. The averaging ranging time for soft tissue is 13 μsec per centimeter, so that radar circuitry and display techniques must be used. Complete cross-section pictures of the locations of reflecting interfaces can be obtained by scanning the tissue with the sound beam and displaying the echoes of the face of an intensity modulated cathode-ray tube.

Fig. 1 is an example of the pictures produced by the 15-mc apparatus used by Wild and Reid in the investigation of human breast cancer.[10] This is a true "B" scan display produced by translating the transducer and

[4] H. T. Ballantine, Jr., T. F. Hueter, and R. H. Bolt, "On the use of ultrasound for tumor detection," *J. Acoust. Soc. Amer.*, vol. 26, p. 581; July, 1954.

[5] S. Satomura, "Ultrasonic Doppler method for the inspection of cardiac functions," *J. Acoust. Soc. Amer.*, vol. 29, pp. 1181–1185; November, 1957.

[6] S. Effert, H. Erkens, and F. Grosse-Brockhoff, "The ultrasonic echo method in cardiological diagnosis," *German Med. Monthly*, vol. 2, pp. 325–328; November, 1957.

[7] E. E. Suckling and W. R. Maclean, "Sonic images," *J. Acoust. Soc. Amer.*, vol. 29, pp. 146–148; January, 1957.

[8] L. D. Rozenberg, "Survey of methods used for the visualization of ultrasonic fields," *Soviet Phys.-Acoust.*, vol. 1, pp. 105–116; 1955. (Translation.) This paper contains descriptions and references for Sokoloff's tube and other devices which might be used in this application.

[9] F. Dunn and W. J. Fry, "Ultrasonic absorption microscope," *J. Acoust. Soc. Amer.*, vol. 31, pp. 632–633; May, 1959.

[10] J. J. Wild and J. M. Reid, "Progress in the techniques of soft tissue examination by 15 mc pulsed ultrasound," in "Ultrasound in Biology and Medicine," E. Kelley, Ed., Amer. Inst. of Biological Sciences, Washington, D. C., publ. no. 3, pp. 30–45; 1957.

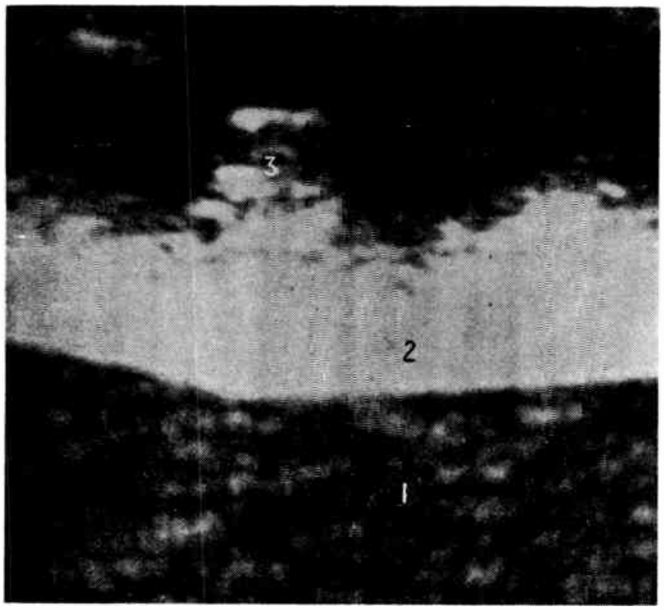


Fig. 1—Type "B" scan cross section of human breast showing the reflection of 15 mc ultrasound from a cancerous tumor. 1, Water chamber within which the transducer moved; 2, Echoes from normal tissue; 3, Echoes from cancerous tumor (Adeno-carcinoma). (Wild and Reid.[10])

sound beam in the horizontal direction while the beam was directed into the tissue. The range sweep is plotted vertically. The dark space, 1, represents the water chamber within which the transducer moves, and the white band, 2, the echoes from many small breast structures. These structures are smaller than the spatial extent of the pulse, so they are not resolved. The echoes, 3, returned from the greatest depth within the tissue are from a site later determined at operation to contain a cancerous tumor. The echo strength decreases rapidly with increasing range because of the high attenuation at 15 mc. It was found that 89 per cent of the malignant tumors studied "stood out" from the background echoes, and that 82 per cent of the nonmalignant tumors were indistinguishable from or had smaller reflections than the background.

The principal aim of this work was the detection, rather than the diagnosis of cancer. The diagnostic approach was used to obtain data on tumors which were subsequently removed and examined by a pathologist so that the ultrasonic findings could be correlated with clinical data. Cancers as small as 1 mm in diameter have been detected in the area of the nipple. Wild is continuing work on detection with equipment for scanning the entire breast at different frequencies. Pictures of breast cancer which have a very different appearance have been obtained by Howry using a frequency of 2.5 mc.[11] Apparently, the small structures in the normal breast do not scatter as much energy at the longer wave length as they do at 15 mc. The tumor structures are clearly seen against the blank space representing normal tissue. Although diagnosis might be difficult at this frequency, this display should be ideal for the detection of tumors.

Instruments suitable for scanning the prostate, cervix, and lower bowel after insertion in the appropriate body opening are being developed by Wild. Although these special tools are not of general applicability elsewhere in the body, they do simplify the diagnostic problem at these sites of high incidence of cancer. The lower bowel, for example, executes an "S" curve deep within the pelvic bones. A rotating transducer and PPI display has produced a series of cross-section pictures of the last 17 inches of bowel in which the wall structure as well as echoes from outside the wall can be seen.[12]

To record the size and shape of anatomical structures faithfully, the pulse length and width must be smaller than the smallest feature of interest. The pulse length can be reduced with the proper acoustic terminations for the piezoelectric element,[13] but even with acoustic lenses to narrow the beamwidth, the pulse is usually wider than it is long. For very strong signals, the limit level of the receiver can be 40 to 60 db down from the peak of the beam pattern, resulting in severe "pulse stretching" in azimuth. Such strong signals arise from the specular reflection at smooth interfaces, as at major organs, bones, and large blood vessels. To map all the reflections from such an area, the relatively weak scattered reflections from interfaces that are not normal to the axis of the sound beam must be used, and these are easily masked by the wide specular reflections.

The only reported method which satisfactorily solves this problem for tissues which are stationary was developed by Howry.[14] He worked on the general problem of visualizing body tissues and developed a system whereby the sound beam is moved in such a manner that each interface is effectively scanned from many angles. This is termed "compound scanning," since it is easily accomplished by the superposition of two simple scanning motions. One component moves the average position of the transducer around the periphery of the scanned area on a circular path; the other is an oscillation along a line tangent to the circle. The range sweep on the display oscilloscope must follow the motion of the sound beam in both starting position and angle. If the display system is perfectly linear, only those portions of the echo signals which are not affected by refraction, multiple reflection, and azimuth stretching will plot in the same position on the tube face. The signals which do appear in the same place will be added at the film of the oscilloscope camera and present a true re-

[11] D. H. Howry, D. A. Stott, and W. R. Bliss, "The ultrasonic visualization of carcinoma of the breast and other soft tissue structures," *Cancer*, vol. 7, pp. 354–358; March, 1954.

[12] J. M. Reid and J. J. Wild, "Current developments in ultrasonic equipment for medical diagnosis," *Proc. Natl. Electronics Conf.*, vol. 12, pp. 1002–1015; October, 1956.

[13] E. G. Cook, "Transient and steady-state response of ultrasonic piezoelectric transducers," 1956 IRE CONVENTION RECORD, pt. 9, pp. 61–69.

[14] D. H. Howry, "Techniques used in ultrasonic visualization of soft tissues," in "Ultrasound in Biology and Medicine," E. Kelley, Ed., Amer. Inst. of Biological Sciences, Washington, D. C., publ. no. 3, pp. 49–63; 1957.

flection picture of the structures. The pulse length is made very short by using a pulse consisting of only 1 or $1\frac{1}{2}$ cycles of the "center frequency," and it determines the effective resolving power of the system. Howry has used a center frequency of 2.5 mc and weak lenses to produce a narrow beam capable of differentiating structures to a depth of at least three inches. Fig. 2 is a picture of a normal leg obtained by Howry, and Fig. 3, a normal neck. These pictures show a surprising amount of detail and illustrate the complexity of soft tissue.

The current status of this work and the problems encountered in the interpretation and use of the new information made available are outlined in the following note contributed by Dr. Joseph H. Holmes of the University of Colorado Medical Center.

The diagnostic use of the ultrasonic technique in medicine requires further work before it will be possible to obtain reliable consistent, and repeatable diagnostic results in a large variety of patients with different body structure and pathology. At present the equipment seems to be improved to the point where it is possible to get good tissue detail in a variety of anatomical areas under ideal conditions. The diagnostic significance of some of the structures seen will require further correlation with autopsy and biopsy material, operative examination or special studies. This will require time and considerable experience with the examination of various anatomical areas by many investigators. For example, although a number of abnormal liver pictures have been obtained on patients, it has not yet been possible to verify the pathology in many instances and obtain correlation with the clinical data. Our own group has just succeeded in obtaining good somagrams on dogs; thus it will now be possible to produce experimental pathological lesions and follow their course with serial somagrams and confirm the findings pathologically.

In examining patients, a horizontal scan indicates only the gross presence of the lesion but no anatomical data regarding its nature or magnitude. The circular scan will give additional detail regarding the structure of the lesion. It, of course, visualizes ideally in circular or semi-circular structures such as the leg, the neck, and the liver, whereas greater difficulty is encountered in visualizing the kidney since the circular scan will not always be equidistant from the skin surface throughout the necessary scanning area. Because of this factor and also because of considerable individual variation in anatomical contours, it will be necessary to set up ideal positioning and scanning conditions which can then be applied repeatedly to a large variety of patients and problems. This is one of our particular objectives at the present time. It is also necessary to develop various supportive braces and appliances to support the sick patient comfortably without movement during the examination. Further effort is being devoted to development of examining techniques which do not require immersing the patient in a tub. To date, this has been best accomplished by a semicircular aluminum pan cut away sufficiently to permit insertion of the patient's side. The fluid within the pan is contained on the cut side by polyvinyl plastic sheeting. When the patient's side is rubbed with K-Y Jelly and pressed against the plastic sheeting a good sound seal has been obtained, and it is possible to get good pictures of the hepatic and splenic areas.

To date, our own diagnostic studies have been confined to the neck, the liver, breast, spleen and kidney areas. In the case of the neck, it has been used diagnostically to delineate lumps and nodules with a fair degree of success when correlated with the surgical findings. In the case of the liver, cirrhotic processes and tumors have been visualized. In the case of cirrhosis, sufficient data are not yet available to state fully the meaning of the various patterns obtained. In the case of the breast, fibrotic lesions visualize well, while homogeneous tumors present a more difficult diagnostic problem. In the leg, studies have been made of edema, various muscular and joint lesions, muscular dystrophy and vascular lesions. In some instances of muscular dystrophy, quite different patterns have been obtained, but their significance has not been ascertained. Some good pictures have been obtained in the edematous leg. Thus while diagnostic studies are being run daily with this type of equipment at this institution, it may still require some time before ideal examining techniques are developed and definite diagnostic patterns are established.

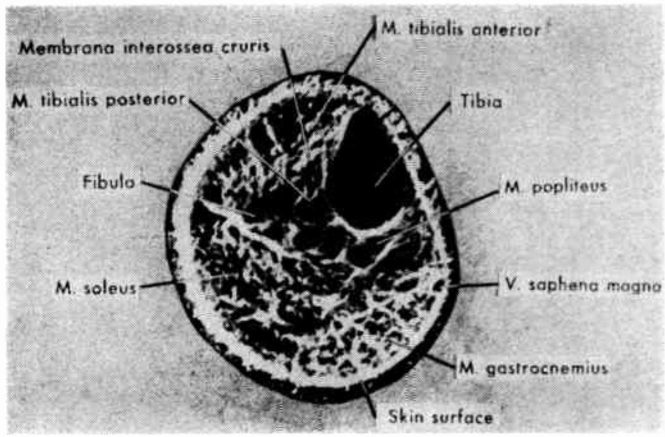


Fig. 2—Compound scan cross section of normal lower leg showing details of structure obtained with 2.5-mc ultrasound in short pulses and a focused beam. The transducer travel followed a "B" scan motion which was moved around the circumference of the leg. M, muscle; V, vein. (Howry and Holmes.)

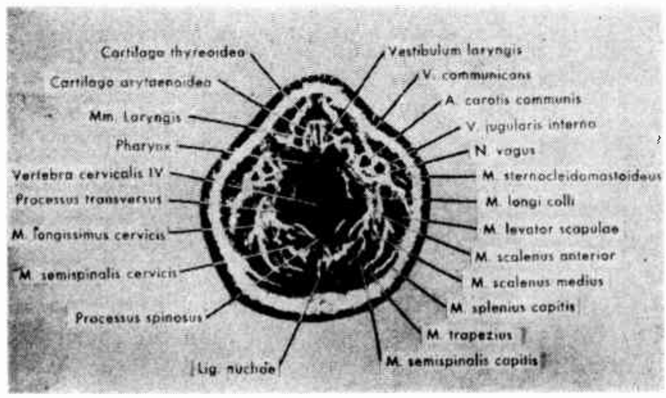


Fig. 3—Compound scan of normal neck as in Fig. 2. M, muscle; V, vein; A, artery. (Howry and Holmes.)

Many diseases and accidents to the eye are complicated by hemorrhage, exudate, cataracts or corneal damage rendering the eye opaque so that visual examination cannot determine the extent of the injury. In addition, tumors of the muscles behind the eye cannot be seen since the wall of the eye is opaque. The soft tissues of the eye and many foreign bodies cannot be seen with the X ray. The application of ultrasonic echo-ranging techniques to the problem is being investigated by Baum and Greenwood.[15] The eye is an almost ideal site for examination by this means since it is on the body surface, transparent to sound waves, and of simple and regular geometry. Fig. 4 is a sector scan of a human eye with total retinal detachment, verified by ophthalmoscopy. The cornea, where the sound entered the eye, is at the top. The "V" shaped structure, 3, is the retina which normally is attached to the rear of the eye, shown at 4. The lens of the eye is not visualized. The transducer was immersed in an isotonic saline bath contained in an open-ended tank sealed to the patient's face by a rubber mask.

[15] G. Baum and I. Greenwood, "The application of ultrasonic locating techniques to ophthalmology," *A.M.A. Arch. Ophthalmol.*, vol. 60, pp. 263–279; August, 1958.

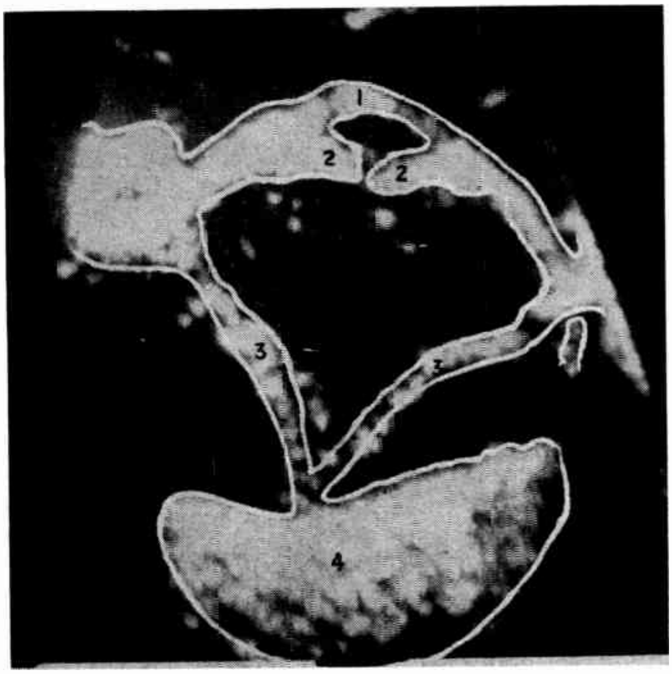


Fig. 4—Sector scan cross section of the eye as seen from above, showing the reflection of 15-mc ultrasound from a totally detached retina. 1, cornea; 2, iris; 3, retina; 4, posterior sclera and orbital fat. (Baum and Greenwood.[15])

The ability to differentiate between hemorrhage, tumors, and foreign bodies (wood) as well as retinal detachment has also been demonstrated. Currently, compound scanning is being applied to an eye scanner to help visualization of the entire orbit and to obtain better detail.

The application of ultrasonic echo-ranging to the diagnosis of heart disease has been made by several investigators. The heart is not a particularly promising site for examination. In the intact human, it is hidden within a rib cage offering propagation paths through half-inch slits, and it is continuously in motion. That the recording of the motion alone may be of importance has been demonstrated by Hertz and Edler.[16,6] By echo-ranging through the space between the ribs, they obtained strong echoes from the walls of the heart and recorded their instantaneous position on a film record from which the wall velocity at different parts of the cardiac cycle can be obtained.

The velocity of the left atrial wall has been shown to indicate the degree of mitral stenosis. With stenosis, the mitral valve opening is too small, and when the atrium empties into the ventricle the atrial wall velocity is less than normal.

Echoes from other heart walls have been demonstrated, but their significance is not yet clear.

Work on this problem is being continued by Schwan and Reid.[17] Propagation through the lung is being studied in the expectation of obtaining heart echoes from regions where some lung tissue is in the propagation path. Since low frequencies are required to overcome the scattering by the lung, considerable effort is being expended on overcoming the excessive beamwidth resulting from the small aperture between the ribs. One-way transmission of 1-mc pulses through the entire lung has been achieved.

[16] C. H. Hertz and I. Edler, "Die Registrierung von Herzwandbewegungen mit Hilfe des Ultraschall-Impulsverfahrens," *Acustica*, vol. 6, pp. 361–364; 1956.

[17] Supported by research grant H-3882 from the Natl. Heart Inst., Natl. Institutes of Health, Bethesda, Md.

Applications of Ultrasound to Biologic Measurements*

J. F. HERRICK†, MEMBER, IRE

Summary—**The achievement of reliable and reproducible measurements in the broad field of the life sciences demands the highest level of competency among the experts who know how to measure. The new technology called "sonics" encompasses heretofore problematic industrial measurements with ultrasound. The application of sonics for measurement in the biologic sciences gives promise. Methods for measuring the velocity of blood, the viscosity of blood and certain dimensions of living bodily tissues are described in this report.**

* Original manuscript received by the IRE, September 1, 1959.
† The Mayo Foundation, University of Minnesota, Rochester, Minn.

UNTIL reliable measurements can be achieved in any aspect of scientific information, the considered opinion of the experts in a particular field will exert a dominating influence. When the opinions of two or more experts in a given field are not in agreement, arguments arise, the literature becomes cluttered, and a state of dissatisfaction develops. A quantitative understanding of a mechanism not only leads to a better control of the mechanism but also permits predictions to be made. It would be interesting to cite instances in the development of our knowledge where simple and direct measurements of familiar phenomena have modified greatly the interpretation of previous qualitative observations.

It is not necessary to elaborate further on measurement because the majority of our readers—physicists and engineers—are dedicated to the art, the skill and the science of measurement. Physicists and engineers have become experts in measurements because the phenomena included in their disciplines lend themselves well, on the whole, to measurement. Elegant techniques have been conceived and successfully applied for obtaining quantitative understanding of processes hidden to the limited human senses.

At the time of this report, the experts in measurement are being challenged to apply their potentialities in the fields of biology and of medicine. If measurements could have been made as easily in the field of the life sciences as in the field of physics, biologists would also be regarded as skilled in the use of tools for measurement. Reliable measurements of the many parameters characteristic of living things are extremely difficult. Many biologic measurements have been made *in vitro* where certain variables interfering with the desired measurements may be controlled or excluded. However these *in vitro* measurements do not permit extrapolations readily into *in vivo* situations.

It is not sufficient for a physicist or an engineer to bring his tools and know-how into the biology laboratory and proceed to make measurements. The science of biophysics is not to be defined as the immediate and direct application of the tools and the laws of physics to living phenomena. The laws of physics have their origin in things that are not alive. An understanding gained by thorough and careful analysis of the biologic problem is required before the methods for measurement are applied so that the application of the tools for measurement will not modify the desired quantitative information. The true biophysicist hopes to discover the laws of living phenomena!

In the author's opinion, this ultimate objective of discovering the laws of the living can be achieved most hopefully by a suitable grouping of competent investigators within a research institute as described so well by Sir Charles Harrington.[1] Cross fertilization of the skills and the know-how of outstanding scientists in the fields required for solving a given biologic problem will be the secret to the successful solution. Diametrically opposed to this opinion is another opinion based on the following thesis: for creativity, ignorance helps.

The potentialities of ultrasound for biologic measurement are especially promising. Acoustics has been put to work in industry with such success that a new area of technology called "sonics" has been born. A description of successful applications of ultrasound for biologic measurements follows.

[1] C. Harrington, "The place of the research institute in the advance of medicine," *Lancet*, vol. 1, pp. 1345–1351; June, 1958.

Measurement of Flow of Blood

The application of ultrasound to measurement of blood flow is relatively recent. From an armchair point of view, the ultrasonic method for measuring the flow of blood in vivo appeared to be almost ideal for the following reasons. 1) Measurements could be made without opening the vascular system (Fig. 1). Ultrasound can be coupled to the blood through the wall of the blood vessel. 2) The method is fast. Reliable quantitative observations on changes in flow within the interval of a cardiac cycle can be made. 3) The observed index of flow (an ultrasonic signal having a frequency of 380 kc or higher) is foreign to any naturally occurring parameter within the living animal so far as is known at this time. 4) There is a linear relation between the flow of blood and the observed index (Fig. 2).

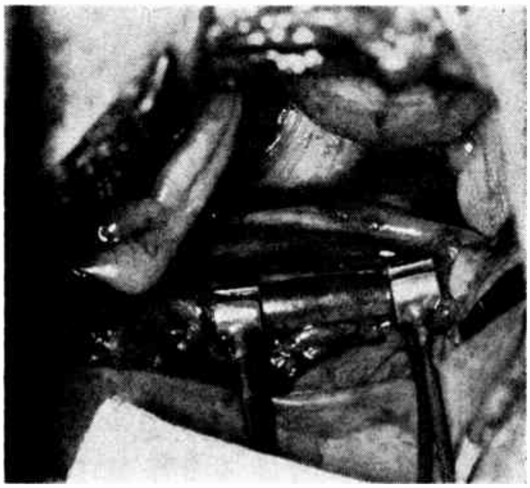


Fig. 1—Cylindrical ceramic transducers applied *in vivo* to blood vessel. Note slit in each transducer which permits application on blood vessel.

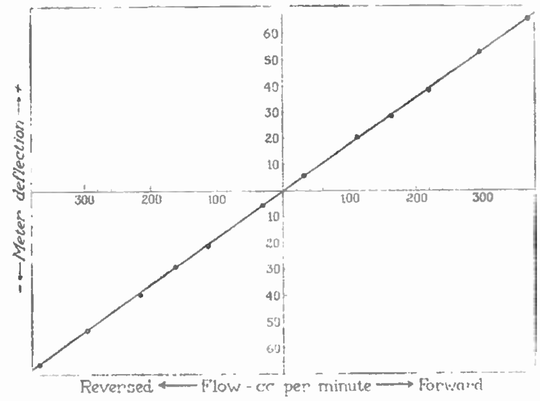


Fig. 2—Illustration of linearity obtained from calibration of ultrasonic flowmeter *in vitro*.

Several separate groups of investigators have either developed or are presently developing ultrasonic blood flowmeters. The flowmeter of each group is of somewhat different type, so that ultimately, the best type of ultrasonic flowmeter should be achieved.

At the Mayo Clinic laboratory, it was decided to reproduce, with slight modifications, the ultrasonic flowmeter developed by Kalmus and his associates.[2] The disk type of ceramic transducer has been replaced by one which is cylindrical in shape, and an automatic phase-shift control has been introduced into the circuitry. A previous publication[3] presents details of the various circuits, and therefore, the circuitry will not be reproduced in this report.

The purpose of this paper is to report the achievements which presently have been attained, not to present a detailed and lengthy report on any particular application of ultrasound for biologic measurement. The particular achievement in measuring blood flow is considered to be electronic, because the sensitivity and the stability of the required instrumentation for measuring the small flows of blood made demands which were beyond the physically realizable limits of the ultrasonic flowmeters already available for use in industry.

The particular ultrasonic flowmeter developed in the clinic laboratory requires the detection of a phase difference between signals traveling upstream and downstream between two given transducers. The frequency used presently is 380 kc and the distance between the transducers is 2.54 cm. If an arbitrary value for the velocity of blood is considered to be 1 meter per second, and since the velocity of sound in blood is about 1500 meters per second, then the difference in transit times of the signal traveling up and downstream will be 2×10^{-8} seconds. The phase shift in the signals is less than 3 degrees. Actually, the velocity of the blood in the larger vessels of the resting dog is less, so that much smaller phase shifts occur *in vivo*. The measurement of such small phase shifts has been so difficult to achieve that the physiologists at the University of Washington decided to develop an ultrasonic flowmeter designed to measure differences in transit times directly with the use of the pulse technique.[4] The physical orientation of the transducers in their flowmeter is similar to that described by Considine.[5] This author is happy to note that this issue contains a paper[6] which includes a detailed description of the ultrasonic flowmeter which has been developed by Professor Rushmer and his associates at the University of Washington.

Other laboratories are developing ultrasonic blood flowmeters which make use of the Doppler principle. Still other groups are working on designs which operate on different basic principles. Further work must be done before these more recent models of ultrasonic flowmeters can be reported in the literature. During an interesting symposium on blood flowmeters which was held at the University of Nebraska School of Medicine in June, 1959, it was apparent that considerable interest in ultrasonic blood flowmeters was aroused. If the various groups which are developing these ultrasonic flowmeters can share their experiences and techniques, ultimately an excellent method for measuring the flow of blood should be available for medical research.

Measurement of the Viscosity of Blood

An instrument known as the "ultra-viscoson,"[7] which measures the viscosity of blood continuously, is available commercially and has been used successfully for several years by medical research groups. Changes in blood viscosity can be measured automatically and recorded graphically. Such continuously recorded profiles of instantaneous changes in the viscosity of blood permit the collection of highly valuable information. The ultra-viscoson incorporates an ultrasonic probe and an electronic automatic computer. The sensing element of the probe is a thin blade of steel alloy which functions as an ultrasonic transducer having a frequency of 28 kc. The maximal amplitude of vibration is less than 1 μ. The ultrasonic power intensity is of the order of microwatts. Such small power intensities eliminate any effects of ultrasonic energy on the characteristics of the blood.

The absolute accuracy of the method for Newtonian liquids, in the range for which the instrument is designed, is ± 2 per cent. Blood is a non-Newtonian liquid and, hence, the ultra-viscoson measures the apparent viscosity with a reproducibility of ± 2 per cent. The reproducibility of the measurements is described as excellent by those medical investigators who use the ultra-viscoson. They have employed this ultrasonic device for studying changes in the viscosity of blood during coagulation. The viscosity profiles were found to be useful and informative indicators of changes in blood coagulability in patients with thrombo-embolism.[8] These profiles frequently seemed to indicate heparin requirements sooner than did clotting times. The ultra-viscoson may be used for studying changes in viscosity of the blood in patients having various pathologic conditions.

Measurements of Dimensions of Living Tissues

Perhaps one of the earliest measurements of the thickness of body tissues by ultrasonic pulses was done

[2] H. P. Kalmus, A. L. Hedrich, and D. R. Pardue, "The acoustic flowmeter using electronic switching," IRE Trans. on Ultrasonics Engineering, vol. UE-1, pp. 49–62; June, 1954.

[3] M. G. Haugen, W. R. Farrall, J. F. Herrick, and E. J. Baldes, "An ultrasonic flowmeter," *Proc. Natl. Electronics Conf.*, vol. 11, pp. 465–475; October, 1955.

[4] D. K. Franklin and R. M. Ellis, "A pulsed ultrasonic flowmeter," *Federation Proc.*, vol. 17, p. 49; March, 1958. (Abstract.)

[5] D. M. Considine, "Process Instruments and Controls Handbook," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 4–89—4–90; 1957.

[6] D. Baker, R. M. Ellis, D. L. Franklin, and R. F. Rushmer, "Some engineering aspects of modern cardiac research," this issue, p. 1917.

[7] R. Yesner, A. Hurwitz, S. R. Rich, W. Roth, and M. E. Gordon, "Preliminary observations on blood coagulation utilizing ultrasonics for continuous measurement of viscosity," *Yale J. Biol. and Med.*, vol. 24, pp. 231–235; December, 1951.

[8] H. W. Harrower, A. Hurwitz, and R. Yesner, "The treatment of thromboembolism with aqueous heparin," *Surg. Gynecol. and Obstet.*, vol. 106, pp. 293–305; March, 1958.

by Wild[9] at the University of Minnesota Medical School. Dr. Wild measured the thickness of segments of the human bowel and dog bowel that had been placed in a suitable chamber containing a quartz crystal having a frequency of 15 mc. The duration of the pulse was about $\frac{1}{2}$ μsec. The ultrasonic pulses were presented on a synchroscope which permitted time measurements for the echoes from the interfaces at either side of the segment of bowel. The thickness of the tissue could be determined from these data.

Rushmer and his associates[10] are measuring the thickness of certain organs of the dog's body (spleen, liver and others) by placing transducers in apposition (on opposite sides at the level for the desired measurement) and recording the time for the ultrasonic signal to traverse the unknown distance. The frequency of the ultrasound in most of these experiments is 3 mc. Some observations were made using a frequency of 10 mc. Measurements of ventricular dimensions by this technique have been published.[11]

Techniques for the visualization of body tissues have developed greatly in the last few years. By means of standard radarlike techniques, so-called ultrasonograms of living tissues can be displayed on a cathode-ray-tube screen. These visualization techniques permit quantitative observations on body structures. The portion of this paper by Mr. Reid gives valuable information about ultrasonograms. As better ultrasonograms are produced, more reliable measurements will be achieved.

[9] J. J. Wild, "The use of ultrasonic pulses for the measurement of biologic tissues and the detection of tissue density changes," *Surgery*, vol. 27, pp. 183–188; February, 1950.

[10] R. F. Rushmer, personal communication to the author.

[11] R. F. Rushmer, *et al.*, "Left ventricular dimensions recorded by sonocardiometry," *Circ. Res.*, vol. 4, pp. 684–688; November, 1956.

The Use of Electronic Computers to Aid in Medical Diagnosis*

R. S. LEDLEY†, MEMBER, IRE, AND L. B. LUSTED‡, FELLOW, IRE

Summary—**With the use of computers several mathematical techniques can be applied to aid certain aspects of medical diagnosis. However, much work remains to be accomplished in trying these methods under practical conditions. Although wide interest has been expressed, few studies have been reported in the literature. Among the potential advantages of computer aids are: making available to the physician quantitative methods in areas related to data analysis and differential diagnosis; assisting in the evaluation of the best alternative courses of action during stages of the diagnostic testing processes; and periodic recording and evaluating of individual physiologic norms for more sensitive determination of an individual's health trend relative to disease prevention. Communication between the physician or researcher and the computer is presently technically feasible but much research and planning are still required for realistic application.**

Introduction

Recently it has been recognized that the use of electronic computers may be able to aid in certain aspects of medical diagnosis. For example, research on this possibility has shown that a computer can 1) produce a list of possible diagnoses for a hospital case by analyzing the symptoms[1] presented with respect to data characteristic of certain diseases; 2) indicate further diagnostic tests which best differentiate between remaining disease possibilities; 3) calculate probabilities for the alternate diagnostic possibilities; and 4) aid in an analysis, based on hospital case data, of value decisions which lead to treatment planning.[2,3] Since the above functions must be based on extensive medical data, it is possible that computers could simultaneously 5) compile statistics that relate symptom combinations to disease states, and treatment to prognosis.[2] In addition, computers can aid data recording and analysis of certain diagnostic procedures. For example, computers can 6) tabulate quantitative criteria derived from electrocardiograms and electroencephalograms, and perform calculations based on such data. Finally, the computers can 7) retrieve current information relative to the above functions, and 8) record and

* Original manuscript received by the IRE, September 1, 1959.

† Dept. of Electrical Engineering, The George Washington University, Washington, D. C.

‡ University of Rochester School of Medicine and Dentistry, Rochester, N. Y.

[1] For the purposes of this paper, we use the term symptom, or test, to include medical history, signs, symptoms, laboratory test results, etc.; that is, these terms include all the information that can be obtained with respect to the patient's medical state.

[2] R. S. Ledley and L. B. Lusted, "Reasoning foundations of medical diagnosis," *Science*, vol. 130, pp. 9–21; July, 1959.

[3] M. Lipkin and J. D. Hardy, "Mechanical correlation of data in differential diagnosis of hematological diseases," *J. Amer. Med. Assoc.*, vol. 166, pp. 113–125; January 11, 1958.

recall desired aspects of a particular patient's total medical record (such as total radiation dosage received, previous allergic reactions, individual biochemical and physiologic norms and deviations, etc.) which might be useful in a current evaluation of the patient's status.

The successful use of computers is necessarily limited by the validity of the input information. It must be remembered that one cannot get out more than one puts into a computer; that is, a computer can do no more than it is *programmed* to do. If a certain disease-symptom relationship is not recorded in the computer memory, the relationship cannot enter into any computations. Poor history taking, incorrectly executed laboratory tests, incorrect interpretation of results of diagnostic procedures, etc., will all contribute to incorrect computer outputs. In fact, these remarks apply to any method of making the diagnosis and formulating the treatment plan; the computer has no magical properties from this point of view.

While it is important to understand the limitation of the computer, it is also unwise to underestimate its capabilities. For example, a computer can be made to 1) "check" the input information for logical inconsistencies; 2) "learn" from the statistical experience of current data that may be continually fed into its memory; 3) "organize" current information for coordinate retrieval; 4) "read" an article and extract pertinent diagnostic information; 5) carry out complicated "reasoning" processes; 6) "evaluate" an optimum or near-optimum course of action for given circumstances, and so forth.[4] But in the final analysis it is the scientific ingenuity of the human mind that "tells" the computer how to perform these remarkable feats.

There are other problems associated with the use of computers besides those of input data accuracy and data omission. Standardization of nomenclature is a practical necessity. While the computer can be made to recognize different words as denoting the same idea, it obviously cannot distinguish between different ideas denoted by the same word. It is highly advisable to standardize the coding procedures so that, for example, medical records from different hospitals will have directly compatible codes. There is also the ever-present problem of standardization of test interpretations. Observe that these problems arise simply because the use of a computer requires logical self-consistency, together with the preliminary analysis of all possible relevant outcomes.

The advantage of computer aids could arise from many sources. First, computers might assist the physician by making available the use of mathematical methods of logic, probability, and statistics in areas related to data analysis and differential diagnosis, and of value and utility theory in areas related to planning decisions for patient treatment. The medical profession is greatly concerned about the effectiveness of present diagnostic processes and is continually striving for improvement—witness the extensive clinical pathological discussions and case histories that appear in medical journals.[5] The medical literature also reflects the importance physicians attach to the exceedingly difficult value decisions that may frequently arise.[6]

Second, after the history and physical examination have been completed, it frequently happens that the physician may desire additional laboratory tests, often within a specific class (*i.e.*, blood tests, liver function tests, etc.). With the aid of a computer, the minimum necessary tests for a particular patient may be calculated. This might result in substantial savings for the patient, not only in money but also in discomfort in many instances. Of equal importance is the substantial saving that may result in the case of frequently overburdened hospital laboratory services. Of course, it must be noted that eliminating redundant tests means that the correct interpretation of any one test becomes crucial. Thus if a certain set of tests is difficult to make or interpret a redundant set of tests should be used.

Third, physiological indexes (*i.e.*, blood pressure, blood sugar level, etc.) measure the state of health. The variation or standard deviation of such an index for an individual is smaller than the variation of the index for the population in general, and in fact the mean for an individual may not coincide with the mean for the population. Thus, an index value which falls within the normal range for the population may not be normal for an individual, and conversely. A computer may be used to record the values of such indexes taken at periodic examinations for individuals. In this way the test results for an individual may be accurately evaluated at any time with respect to that individual's physiological normal-indexes, rather than with respect to the norm for the entire population. Also, trends leading to pathological conditions might be observed with greater sensitivity, thus permitting earlier initiation of preventative measures.

Previous Studies Related to Differential Diagnosis

Several studies have been made on the use of computer-type aids for listing diagnostic possibilities. The first was Nash's "diagnostic slide rule."[7] Here the dis-

[4] See, for example, R. S. Ledley, "Mathematical foundations and computational methods for a digital logic machine," *J. Operations Res. Soc. Amer.*, vol. 2, pp. 249–274, August, 1954; *Proc. Internatl. Conf. on Scientific Information*, Washington, D. C., 1958; H. P. Luhn, "The automatic creation of literature abstracts," *IBM J. Res. & Dev.*, vol. 2, pp. 159–165; April, 1958, and others.

[5] See, for example, A. H. Douthwarte, "Mistakes in diagnosis," *Medical World*, vol. 79, pp. 113–115, August, 1953; A. H. Douthwarte, "Pitfalls in medicine," *Brit. Med. J.*, vol. 2, pp. 895–900, 958–967, October 20 and 27, 1956; "Misdiagnoses," (editorial) *Lancet*, vol. 1, p. 1034, May 23, 1953; L. Clendening and E. H. Hashinger, "Methods of Diagnosis," C. V. Mosby Co., St. Louis, Mo., pp. 73–76, 1947. Clinicopathological exercises taken from the case records of the Massachusetts General Hospital are published each week in the *New England Journal of Medicine*.

[6] W. L. Sperry, "The Ethical Basis of Medical Practise," Paul B. Hoeber Inc., New York, N. Y.; 1950.

[7] F. A. Nash, "Differential diagnosis: an apparatus to assist the logical faculties," *Lancet*, vol. 1, pp. 874–875; April 24, 1954.

eases are listed along one edge of the rule. For each symptom there is a symptom stick which can be placed adjacent to the disease listing; lines are drawn on each symptom stick opposite the positions of those diseases having the symptom. Then those sticks corresponding to a patient's symptoms are placed in the rule and the diseases corresponding to the most symptom lines are the diagnostic possibilities.

The second and third studies used marginal-punched cards.[3,8] In the first of these Lipkin used cards each representing a disease; the marginal positions represented symptoms. Those marginal positions of a card that were punched represented symptoms associated with the disease of that card. Then, for a set of symptoms presented by a patient, the deck of disease cards was sorted for those cards that were associated with *all* of the given symptoms. The cards obtained from this sorting indicated the diagnostic possibilities. In addition, for each disease certain items of data and their places on the margin of the punched card were noted. These were the items that had to be present in order to establish a diagnosis of the disease. Therefore, after sorting, the items of data needed to establish each diagnosis were automatically returned with the diagnostic possibilities. One could automatically determine if a diagnosis could be established based on the patient's data. If no diagnosis was established, the information returned stated which further tests were needed to differentiate among the disease possibilities. Although the above principles were established in the marginal punched card study, the coding procedure did not include negative items of data, nor the ability to handle more than one diagnosis or atypical finding for a single patient without eliminating the correct diagnosis. The second of these card studies, by Ledley, used essentially this same technique.

Finally, Lipkin used a digital computer to compare the findings from a hospital case with data characteristic of the class of diseases under consideration.[9] The computer printed out every disease which corresponded to at least all the findings of the hospital case. For each disease that was printed out, all data corresponding to that disease were also printed out so that the additional differential symptoms could be observed. This was essentially an information retrieval procedure.

Listing Diagnostic Possibilities

Symptom-Disease Complex and Medical Knowledge

Mathematical logic (the propositional calculus) may be employed in order to make a computer list diagnostic possibilities for the symptoms presented by a patient. The use of logic for this purpose is closely related to the concept of symptom-disease complexes. A symptom complex is a list of the symptoms that a patient does and does *not* have; a disease complex is a similar list of diseases. A *symptom-disease complex* (hereafter referred to as the sdc) is a list of both the symptoms *and* diseases that a patient does and does *not* have. For example, consider Fig. 1, where for simplicity our attention is limited to two symptoms, $S(1)$ and $S(2)$, and two diseases, $D(1)$ and $D(2)$. Each column represents an sdc where a unit in the row signifies that the patient has the corresponding symptom or disease, and a zero signifies that the patient does not have the symptom or disease. Thus the column enclosed in the rectangle of Fig. 1 represents the sdc of the patient having symptom $S(1)$, not having $S(2)$, having $D(1)$, and having $D(2)$. The columns of Fig. 1 represent all conceivable sdc's (symptom-disease complexes) that can be formed from two symptoms and two diseases. If p symptoms and q diseases are under consideration, then clearly there are 2^{p+q} conceivable sdc's. For our example, $p=2$, $q=2$, whence $2^{2+2}=16$ columns.

S(1)	1111	1111	0000	0000
S(2)	1111	0000	1111	0000
D(1)	1100	1100	1100	1100
D(2)	1010	1010	1010	1010

Fig. 1—All conceivable symptom-disease complexes.

We use the phrase "all *conceivable* sdc's," but not all of these are *possible* or actually occur. It is available *medical knowledge* that informs us which sdc's do or do not occur. For example, suppose that for the particular symptoms and diseases under consideration medical knowledge informs us that if a patient does not have $D(1)$ and does not have $D(2)$, then he cannot have $S(1)$ and he cannot have $S(2)$, and conversely. Such a statement can be symbolized as

$$\overline{D(1)}\cdot\overline{D(2)} = \overline{S(1)}\cdot\overline{S(2)}$$

where the bar represents "not," and the dot represents "and." The effect of such a statement of medical knowledge is to eliminate from consideration some of the conceivable sdc's which are not possible. For example, the following columns

$S(1)$	110	000
$S(2)$	101	000
$D(1)$	000	110
$D(2)$	000	101

represent the sdc's that cannot happen according to our rule of medical knowledge, for the three columns on the left above represent sdc's where the patient has no diseases but does have the symptoms, and the three columns on the right represent sdc's where the patient has no symptoms but does have the diseases. In Fig. 2 these columns have been crossed off.

[8] R. S. Ledley, "Logical aid to systematic medical diagnosis," *J. Operations Res. Soc. Amer.*, vol. 3, no. 3; August, 1955 (abstract).
[9] In the *Proc. Conf. on Medical Diagnosis*, Rockefeller Institute, New York, N. Y., 1959 (in preparation).

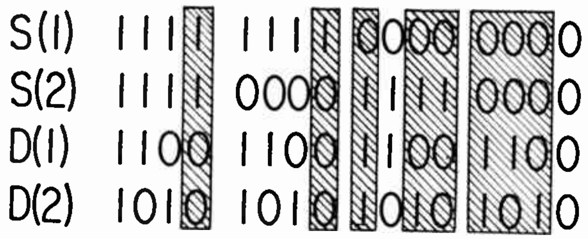


Fig. 2—Effect of medical knowledge: If $\overline{D(1)}\cdot\overline{D(2)}=\overline{S(1)}\cdot\overline{S(2)}$, and $D(2)\rightarrow S(1)$, then the cross-hatched columns are omitted and the remaining possible symptom-disease complexes are shown.

Suppose, for example, that another assertion of medical knowledge is that if the patient has disease $D(2)$, then he must have symptom $S(1)$, which can be symbolized as $D(2)\rightarrow S(1)$. This means, in addition, that the sdc's represented by the following columns cannot occur for they are contrary to the assertion of medical knowledge:

$S(1)$	0	0
$S(2)$	1	1
$D(1)$	1	0
$D(2)$	1	1

These columns have also been crossed off in Fig. 2.

The logical effect of medical knowledge is to reduce the totality of all conceivable sdc's to those which are possible or compatible with the assertions embodied in medical knowledge. Fig. 3 represents all possible sdc's for our illustrative situation. This method may be used as a basis for computer analysis of information pertaining to differential diagnosis.

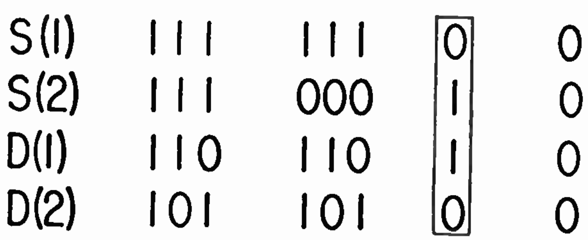


Fig. 3—Making the diagnosis:
Patient presents $\overline{S(1)}\cdot S(2)$; the diagnosis is $D(1)\cdot\overline{D(2)}$.

Listing Diagnostic Possibilities

Suppose a particular patient presents a symptom complex as follows: The patient does not have symptom $S(1)$ but does have symptom $S(2)$. Written symbolically, this is $\overline{S(1)}\cdot S(2)$. To list all the diagnostic possibilities within our framework, we simply look at Fig. 3, which contains all possible sdc's, for those columns which contain the symptom complex $S(1)$. There is only one such column:

$S(1)$	0
$S(2)$	1
$D(1)$	1
$D(2)$	0

Thus the logical analysis informs us that the patient with symptom complex $\overline{S(1)}\cdot S(2)$ has disease complex $D(1)\cdot\overline{D(2)}$; *i.e.*, he has disease $D(1)$ but not disease $D(2)$. Suppose another patient presents $S(1)\cdot\overline{S(2)}$. The columns of Fig. 3 containing this symptom complex are

$S(1)$	111
$S(2)$	000
$D(1)$	110
$D(2)$	101

Thus the patient may have both $D(1)$ and $D(2)$, or else $D(1)$ and not $D(2)$, or else $D(2)$ but not $D(1)$. This means that the symptoms considered are insufficient in this case to distinguish between $D(1)$, $D(2)$, or both $D(1)$ and $D(2)$. Thus, by means of logical analysis[10] the computer can produce a list of possible diagnoses consistent with the symptoms presented by the patient.

Determining Further Tests

The Minimum Necessary Additional Tests

The ease with which medical symptoms or test results can be obtained varies greatly with the complexity of the examination, cost, inconvenience to the patient, and so forth. Clearly it is not feasible first to make all possible tests and then to determine the diagnosis. The more realistic method uses iterated states: collecting information; determining how precise a diagnosis can be made from the results so far obtained; determining what further tests and evaluations should be made; determining how precise a diagnosis can be made from the additional results obtained; determining what further tests should be made, and so forth, until a satisfactory diagnosis has been accomplished.

This problem can be formulated in terms of our logical considerations of the previous section. The formulation will be described by means of an illustration. Fig. 4 shows a set of sdc's associated with 14 symptoms and 8 diseases. For this illustration the symptoms have been classified into history, physical examination, laboratory blood tests, and bone marrow tests, arranged in order of increasing difficulty of performance. Notice that $S(1)$ actually contributes nothing to the diagnosis because, having a unit in every column, it does not distinguish between any columns. Suppose, for illustration, that the patient has both $S(2)$ and $S(3)$. Then the diagnosis so far is that the patient has a disease complex associated with one of the first eleven columns. Thus, due to the patient's symptoms, we eliminate from consideration

[10] A more formal statement of our logical analysis is as follows: Let $D(1), D(2), \cdots, D(n)$ be the diseases under consideration, and let $S(1), S(2), \cdots, S(m)$ be the symptoms under consideration. Then medical knowledge appears as a *Boolean* function E.

$$E(D(1), \cdots, D(n), S(1), \cdots, S(m)).$$

The symptom complex presented by the patient is the Boolean function $P=P(S(1), \cdots, S(m))$, and the logical aspect of the medical diagnostic problem is to determine a function $f=f(D(1), \cdots, D(n))$ of the list of diagnostic possibilities such that the following Boolean equation is satisfied:

$$E\rightarrow(P\rightarrow f)$$

	Column Numbers:	1 2 3 4	5 6 7 8	9 10 11 12	13 14 15 16	
History	S(1)	1111	1111	1111	1111	
	→ S(2)	1111	1111	1111	1000	
	→ S(3)	1111	1111	1110	0110	
Physical examination	→ S(4)	1111	1111	1000	1100	Symptom Complexes
	→ S(5)	1100	0000	0111	0000	
	S(6)	1100	0000	0000	1111	
Blood tests	S(7)	1111	1101	0111	1110	
	S(8)	1111	1110	1110	0001	
	→ S(9)	1111	1010	0101	1001	
	→ S(10)	1111	0100	0000	0100	
	S(11)	1110	0000	1000	1000	
	→ S(12)	1110	1100	1100	0001	
Bone marrow tests	S(13)	1011	0000	0111	0000	
	S(14)	1100	1110	0000	1111	
	D(1)	0000	0000	1000	0000	Disease Complexes
	D(2)	0101	0000	0000	0010	
	D(3)	0010	0100	0001	1000	
	D(4)	0001	1000	0000	0100	
	D(5)	0000	1000	0000	0001	
	D(6)	0000	0100	0111	0000	
	D(7)	1000	0010	0011	0000	
	D(8)	0000	0001	0000	0010	

Fig. 4—What tests to make next.

the rest of the columns, as shown in Fig. 4. Of the physical examination symptoms $S(4)$, $S(5)$, and $S(6)$, only two are necessary, for clearly all three cannot differentiate between the remaining columns better than can any two. Suppose $S(4)$ and $S(5)$ are taken, and the patient has $S(4)$ and *not* $S(5)$; then only seven columns remain as possible diagnoses, *i.e.*, columns 3–9, as shown in Fig. 4. The choice among the blood tests $S(7)$, $S(8)$, $S(9)$, $S(10)$, $S(11)$, and $S(12)$ is not as easy to make. As we shall show below, only three of these six tests need be made, namely $S(9)$, $S(10)$, and $S(12)$. Suppose the patient presents positive results for tests $S(9)$ and $S(12)$, and a negative result for test $S(10)$. Then only a single column remains, namely column 5 as shown in Fig. 4. Thus the bone marrow tests need not be made, and the patient has a combination of both diseases $D(4)$ and $D(5)$ but none of the other diseases. Suppose, on the other hand, that at the physical examination the patient had both $S(4)$ and $S(5)$. Then only columns 1 and 2 would remain. The blood tests will not distinguish between them, and bone marrow test $S(13)$ would have to be made.

Minimizing the Number of Tests

It thus becomes clear that at any stage in the examination of the patient, the minimum number of tests of the next order of difficulty can be determined by a process such as the one described above. It can be shown that the problem of choosing the minimum number of tests at any stage is closely related to the problem of the simplification of Boolean functions.[11] In realistic situations, where there are many more tests and many more possible sdc columns, the aid of an electronic computer becomes essential. However, it is helpful to have a basic rule for choosing the minimum tests in relatively simple cases.

For example, consider the choice between $S(7)$, $S(8)$, $S(9)$, $S(10)$, $S(11)$, and $S(12)$ blood tests when only columns 3, 4, 5, 6, 7, 8, and 9 remain (see Fig. 4). The first step is to count the number of units that appear in the columns for those rows:

	34	5678	9	Number of units
$S(7)$	11	1101	0	5
$S(8)$	11	1110	1	6
$S(9)$	11	1010	0	4
$S(10)$	11	0100	0	3
$S(11)$	10	0000	1	2
$S(12)$	10	1100	1	4

The next step for finding the minimum set is to choose the test that most nearly separates the columns into two equal parts, *i.e.*, having 3 or 4 units, in our case. Test $S(9)$ satisfies this requirement. We now consider the unit columns of $S(9)$, and proceed to find another test that divides these equally. We do the same for the zero columns of $S(9)$; if we can choose the same test as that for the units of $S(9)$, so much the better. The process is repeated. For example, we have

		3	4	5	7	6	8	9
Columns separated by test $S(9)$	$S(9)$	1	1	1	1	0	0	0
	$S(7)$	1	1	1	0	1	1	0
	$S(8)$	1	1	1	1	1	0	1
	$S(10)$	1	1	0	0	1	0	0
	$S(11)$	1	0	0	0	0	0	1
	$S(12)$	1	0	1	0	1	0	1
Columns separated by test $S(10)$	$S(10)$	1	1	0	0	1	0	0
	$S(7)$	1	1	1	0	1	1	0
	$S(8)$	1	1	1	1	1	0	1
	$S(11)$	1	0	0	0	0	0	1
	$S(12)$	1	0	1	0	1	0	1
Columns separated by test $S(12)$	$S(12)$	1	0	1	0	1	0	1

[11] See, for example, S. H. Caldwell, "Switching Circuits and Logical Design," John Wiley and Sons, New York, N. Y., ch. 5 *et passim*, 1958; or R. S. Ledley, "Digital Computer and Control Engineering," McGraw-Hill Book Co., Inc., New York, N. Y., pt. 3 (in publication.)

Hence only the three tests $S(9)$, $S(10)$, and $S(12)$ are necessary:

	3	4	5	6	7	8	9
$S(9)$	1	1	1	0	1	0	0
$S(10)$	1	1	0	1	0	0	0
$S(12)$	1	0	1	1	0	0	1

Calculating the Alternative Diagnostic Probabilities

Probability in Medical Diagnosis

It frequently occurs that it is neither feasible nor desirable to make further tests in order to distinguish among alternative possible diagnoses. The problem then resolves itself into answering the question: In the light of the patient's present symptom complex, what is the probability that the patient has a particular disease complex? We are asking for a conditional probability: Given the patient's symptom complex s_j, what is the probability of his having the disease complex d_i? This is frequently denoted by $P(d_i | s_j)$, where the symbol to the right of the slash is the condition, and the symbol to the left of the slash is the occurrence whose conditional probability is desired.

The well-known Bayes' formula offers important information concerning the composition of $P(d_i | s_j)$. The formula is

$$P(d_i | s_j) = \frac{P(d_i) P(s_j | d_i)}{\sum_{\text{all } k} P(d_k) P(s_j | d_k)}$$

where $\Sigma_{\text{all } k}$ indicates a summation over *all possible disease complexes* d_k under consideration. Let us focus our attention on the two terms in the numerator, namely $P(d_i) P(s_j | d_i)$, since the denominator is merely a normalization factor. The term $P(s_j | d_i)$ is the probability of having the symptom complex s_j when it is known that the patient has the particular disease complex d_i. This is precisely what is described in medical textbooks when a particular disease is under consideration. Of course textbooks do not at the present time assign numerical values to $P(s_j | d_i)$, but they do state that associated with a certain disease particular symptoms are common, frequent, rare, and so forth. The reason that $P(s_j | d_i)$ is discussed [rather than $P(d_i | s_j)$] is that the etiology of the symptoms is related to, or stems from, the disease.

Thus $P(s_j | d_i)$, *i.e.*, *the symptoms resulting from a disease, is a constant for the particular disease complex under consideration, being relatively independent of other circumstantial factors.*

On the other hand, consider the term $P(d_i)$. This is the probability that of the particular population of patients involved any particular patient has this disease complex d_i. By "particular population of patients" we mean that population from which the patients can be considered as chosen *at random* (in the technical sense[12]). For example, the population may be that of the entire United States, or that of a particular clinic in a particular hospital, etc.; in each case $P(d_i)$ will, in general, be different. *This term relates the influence of geographical, seasonal, epidemiological, social, and other such factors to the diagnosis.* Both these factors—the circumstantial or local, and the constant relating to the disease-symptom syndrome—influence the diagnosis.

Properties of the Probabilities

The relation between the logical analysis and our probability considerations is as follows: When a logical analysis for a given symptom complex s results in a *single* diagnosis d, then $P(d | s) = 1$, and conversely. If the logical analysis results in alternative possible diagnoses, then the probabilities with which the patient has these possibilities can be calculated. The probabilities of the patient's having a diagnosis *not* given by the logical analysis are zero. Thus, the logical analysis eliminates from consideration those diagnoses that have zero probability in the light of the patient's symptom complex.

As a series of tests are performed on a given patient, each test result alters the probabilities that the patient has certain disease complexes. It is important to note (see Fig. 6) that these probabilities are not necessarily monotonic and, in general, that the probabilistic results based on partial test knowledge *in no way indicate how the probabilities will run for future test results.* The only exception is that when a probability has become zero, it will stay zero.

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	
S(1)	0	1	0	1	0	0	0	1	0	symptom complexes
S(2)	0	0	1	1	1	1	0	0	1	
S(3)	1	1	0	0	1	0	1	0	1	
S(4)	1	1	1	1	1	1	0	0	0	
D(1)	0	0	0	0	0	0	1	1	1	disease complexes
D(2)	1	1	1	0	0	0	0	0	0	
D(3)	0	0	0	1	1	1	1	1	1	
	d_1			d_2			d_3			
thousands of patients	5	30	10	10	5	5	20	10	5	

Fig. 5—Illustration of disease-symptom complexes for 100,000 patients.

The properties of the probabilities may best be observed by means of an illustration. Suppose, for a population of 100,000 patients, we consider the nine possible sdc's formed by $S(1)$, $S(2)$, $S(3)$, and $S(4)$, and $D(1)$ and $D(2)$, shown in Fig. 5. We have labeled the different symptom complexes by $s_1, s_2, \cdots, s_9$, and the disease complexes by d_1, d_2, and d_3. Let us suppose that we know how many of the 100,000 patients have each of the nine sdc's, for then we can compute all the important probabilities for our illustration. For instance, since

[12] See W. Feller, "Probability Theory and Its Applications," John Wiley and Sons, New York, N. Y., p. 25; 1950.

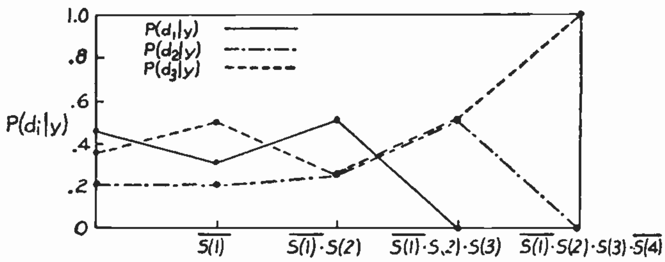


Fig. 6—Example of how increased knowledge changes the probabilities.

$$P(X) = \frac{\text{No. of patients with characteristic } X}{\text{Total no. of patients}}, \text{ and}$$

$$P(X \mid Y) = \frac{\text{Of patients with characteristic } Y\text{, the no. that also have characteristic } X}{\text{No. of patients with characteristic } Y},$$

then we can compute, for example,

$$P(d_1) = \frac{\text{No. of patients with } d_1}{\text{Total no. of patients}} = \frac{5 + 30 + 10}{100} = .45$$

$$P(d_1 \mid \overline{S(1)}) = \frac{\text{Of patients with } \overline{S(1)}\text{, the no. that also have } d_1}{\text{No. of patients having } \overline{S(1)}} = \frac{5 + 10}{5 + 10 + 5 + 5 + 20 + 5} = .30.$$

Now let us consider a particular patient chosen at random from our population and follow the effect that successively testing for $S(1)$, $S(2)$, $S(3)$, and $S(4)$ has on the probability of the patient's having d_1, d_2, or d_3. Before any tests are made, $P(d_1) = .45$, $P(d_2) = .20$, and $P(d_3) = .35$ (see Fig. 5). Let test $S(1)$ be made, and suppose the result is negative; *i.e.*, the patient has $\overline{S(1)}$. Then $P(d_1 | \overline{S(1)}) = .30$, $P(d_2 | \overline{S(1)}) = .20$, and $P(d_3 | \overline{S(1)}) = .50$. Similarly, suppose test $S(2)$ was made, and the patient has $S(2)$; next, that $S(3)$ is made, and the patient has $S(3)$; and, finally, that $S(4)$ is made, and the patient has $\overline{S(4)}$. In Fig. 6 we have graphed the successive probabilities for this sequence of test results. Note that the probabilities are far from monotonic.

Finally, suppose only $P(d_i)$ and $P(s_j | d_i)$ are given. We shall illustrate the use of Bayes' formula to find $P(d_i | s_j)$. Using the figures of Fig. 6, we find, for example, that $P(s_3 | d_1) = .22$, $P(s_3 | d_2) = .25$, $P(s_3 | d_3) = 0$, and $P(d_1) = .45$, $P(d_2) = .20$, $P(d_3) = .35$. The conditional probabilities are assumed to be a constant, independent of time and place; the $P(d_i)$ are determined locally and have current values. Thus, for example,

$$P(d_2 \mid s_3) = \frac{P(d_2)P(s_3 \mid d_2)}{P(d_1)P(s_3 \mid d_1) + P(d_2)P(s_3 \mid d_2) + P(d_3)P(s_3 \mid d_3)}$$

$$= \frac{(.20)(.25)}{(.45)(.22) + (.20)(.25) + (.35)(0)} = .33.$$

Computing Statistics

Relating Symptom-Disease Combinations

As we have seen above, in order to determine the probability with which a patient has a disease complex we must make use of symptom-disease complex statistics. The method of obtaining such statistics is suggested by our sdc structure. For each sdc under consideration, a tally is kept of the number of people having that sdc. In other words, for a particular population of patients (*i.e.*, particular clinic, hospital, etc.) and a particular collection of possible sdc's, each time a diagnosis is made with reasonable certainty or definitely confirmed, one more is added to the tally of the corresponding sdc. When a sufficiently large number of patients has been sampled in this way, the tallies obtained can be used to evaluate the probabilities as described above. It should be observed that in many cases statistics compiled in this way can mount up rapidly. For example, if 50 patients per day pass through a specialty hospital clinic, this amounts to 1000 patients per month, which is not unreasonable; if ten hospitals cooperate, this amounts to 10,000 patients per month.

The collection of such statistics, to generate the necessary probabilistic data of a realistic situation involving many diseases and symptoms, is a data-processing problem to be handled by electronic computers. In order to keep the statistics current, the data collection is a continuing process; new statistics are fed into the computer as they are generated, old statistics are dropped chronologically at the same rate, so that at any time the data in the computer represent results of current trends only. Frequently such a technique is referred to as "making the computer learn by experience."

So far we have been assuming that for each statistic (patient) the complete symptom-disease complex has been determined and confirmed. However, this is not always the case, and the problem arises as to the use of partial information in our data-collection scheme. Partial information refers to cases where it is not known whether or not the patient has a particular symptom $S(n)$, or a particular disease $D(m)$. In such cases, one more is added to the tallies of all those sdc's that correspond to either $S(n)$ or $\overline{S(n)}$ [or $D(m)$ or $\overline{D(m)}$], as well as those symptoms and diseases known for the patient. For example, referring to Fig. 5, suppose the present tallies are (for this example, not in thousands):

5 30 10 10 5 5 20 10 5.

If a new confirmed diagnosis came in as $s_3 \cdot d_1$, then the new tallies would be:

5 30 *11* 10 5 5 20 10 5.

Suppose, next, a confirmed diagnosis was made for partial information $S(3) \cdot S(4)$ and d_1, then the new tallies would be:

6 *31* 11 10 5 5 20 10 5.

Next, if a confirmed diagnosis was made for partial information $S(1) \cdot S(3)$ and $\overline{D(2)} \cdot D(3)$, then the new tallies would be:

6 31 11 10 *6* 5 *21* 10 *6*.

(In our example we did not drop the oldest statistic each time, as we should have.)

Problems in Computing Statistics

As we have noted above, the statistics in which we are interested, *i.e.*, $P(d_i|s_j)$, reflect local changes in place and time, and therefore must be collected locally. However, in order to be valid, "sufficiently large" data must be used, which may not be available locally. Bayes' formula comes to our aid under such circumstances, for clearly it is not necessary to have as large a sample to determine $P(d_i)$ as to determine $P(d_i|s_j)$. Thus, since $P(s_j|d_i)$ is a constant, not dependent on local conditions and independent of time, nation-wide statistics collected over a long period of time can be used for this purpose. Then, locally, only $P(d_i)$ need be derived from collected data, with Bayes' formula used as illustrated above to determine $P(d_i|s_j)$. Notice also that the local collection of statistics to calculate simply $P(d_i)$, where no symptoms enter, is a much simpler process than that we have just described. Thus the use of Bayes' formula presents important practical advantages.

Another important problem arises in connection with rare diseases. If we had a clinic for rare diseases, then for the population of this clinic they would not be rare and the problem would be solved. However, patients having rare diseases must be referred to this clinic; hence the rare diseases must be diagnosed even under conditions of "rarity." The columns or sdc's corresponding to rare disease complexes cannot be allowed to have a zero tally, for then they would be missed. Thus, a minimum but greater-than-zero tally must be kept for these rare sdc's at all times. This minimum tally will give wrong probabilistic results; *i.e.*, in general, it will give results that are too large, but the fact that the minimum nonzero tally is involved should be enough warning. Actually, the probabilities will not be misleading, in comparisons with nonrare diseases; but the probabilities will have no meaning when comparing rare diseases.

Finally, consider the problem of the delay with which our "current" statistics will reflect time-dependent changes. Obviously, even though we are dropping old statistics at the same rate as receiving new statistics, the initiation of an epidemic, for example, will be observed with a finite delay, as will the end of an epidemic. This is a practical problem that exists no matter how the statistics are taken; it reflects the delay in current information-gathering and processing. It might be observed that an "uncertainty principle" applies here, analogous to the well-known uncertainty principle that appears in quantum mechanics, *i.e.*,

$$(\Delta L)(\Delta t) > \epsilon.$$

For, if we wish to make our statistics more sensitive to time changes (*i.e.*, decrease the delay time Δt), then we must accept more statistics from a larger area, thereby increasing ΔL, the geographical locality factor. On the other hand, if we limit the area in which we collect statistics, *i.e.*, decrease ΔL, we shall have to compensate by waiting a longer time, *i.e.*, increase Δt. However, there are many techniques for handling such situations. For example, physicians are alerted at the time of the seasonal occurrence, say, of poliomyelitis. Similarly, the statistics can be tampered with by artificially raising certain tallies from previous epidemiological experience. In this way, the lead in incidence over statistical reflection can be reduced. After the initial priming, the usual tallies can then take over.

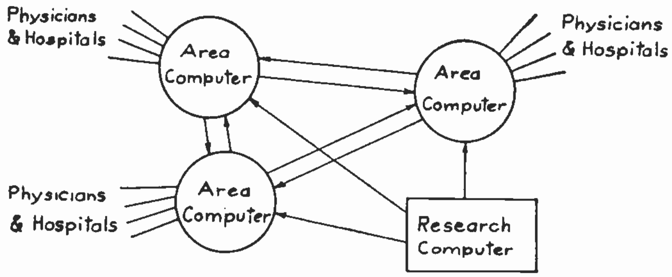


Fig. 7—Central health computing and records service.

Conclusion

It appears that with the use of computers several mathematical techniques can be applied to aid certain aspects of medical diagnosis. Much work remains to be accomplished in trying these methods in practical situations. Although much interest on the part of various medical organizations and individuals has been expressed, few studies have been made. The large-scale use of computers for medical record analysis and aids to diagnosis seems feasible in the near future. In fact, there are at present computers that carry out processes closely related to those we have been discussing. For example, computers exist that 1) automatically communicate with ticket and travel agencies and make nation-wide airline and hotel reservations; 2) automatically update, record, tally, etc., bank accounts and other financial records; 3) record, compute, update, and communicate bills and charge accounts, and so forth. A hypothetical health computing system might require a network of such computers, as appears in Fig. 7. Here each computer can communicate with individual physicians and hospitals within its area, receiving, transmitting, and computing medical information as required. All the area computers can communicate with a special research computer that can sample data as required for various research and public health control investigations. Of course, much research and planning will be necessary before such a health computer network can be a reality. The great significance and importance of such a health computer network cannot be overestimated as an aid to increasing individual good health and longevity, as well as providing a vast new source of medical information concerning mankind. To our knowledge, no such project is under investigation at present. This is surprising since the advantages of such a system are well recognized and present technological capabilities are more than adequate.

New Instrumentation Concepts for Manned Flight*

LAWRENCE J. FOGEL†

Summary—**The advent of modern aircraft has forced the recognition of three fundamental principles required to optimize human flight control:**

The first, kinalog attitude display, is an adaptive kinesthetic analog tracing the human orientation as *g* force is sensed, intended to inhibit the onset of vertigo through the maintenance of continued agreement between the instruments and the human operator's internal "up" vector.

The second, anticipatory display, describes information relative to some aspect of a future status of the vehicle, thus overcoming both the pilot's and the vehicle's response time lag. The speed of modern aircraft already leaves too little time for decision making. Anticipatory display may overcome this problem and significantly improve performance.

The third, modified pictorial display, presents an integrated pictorial view from which has been removed much of the irrelevant data which would be seen in the real world.

These concepts are embodied in proposed aircraft instrument designs which fall within the present state of the art. They are also extended to possible future spacecraft applications. Compatible quantitative instrumentation is also described to complete the cockpit panel. Cursory evaluation has been accomplished by ground simulation and some relevant data is presented. These initial experiments appear to offer a significant promise to increase the performance capability of future manned vehicles.

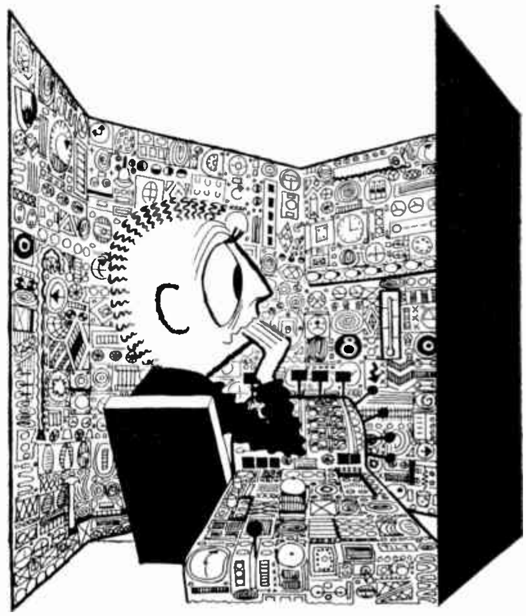


Fig. 1—The clock shop cockpit.

Introduction

The human operator is rapidly becoming the limiting element in manned aircraft. The increased performance capability of the vehicle can only be realized if the aircraft system is designed to be compatible with its human component. Compatibility is the key word; compatibility with the displays to facilitate integration of information so as to allow the required human decision making, and compatibility with the controls which accept the results of decision and translate these commands into the required flight maneuvers.

Present aircraft cockpits show an unfortunate resemblance to clock shops. The pilot is surrounded by dials, pointers and indicator lights and it is his job to monitor all of these, remain aware of the value of each of the parameters, and assimilate this multitude of parameters into the context of the aircraft flight status. Fig. 1 indicates the magnitude of this task. It is understandable why considerable training and experience is requisite to the safe control of present day aircraft. This paper addresses the subject of cockpit design through the recognition of three primary problems. It also establishes the ground rule that the proposed solutions must be feasible within the state of the art and that they must be mutually exclusive to ensure that any combination will improve flight control. At this point, each of the problems will be considered separately, with an indication of the logic for each solution. The proposed mechanization is reserved for the discussion section, which describes each of the instruments as a separate entity.

The first problem is that of vertigo.[1] Its importance is reflected in the inordinately large amount of aircraft accidents which are attributed to "pilot error."[2-4] If even a small percentage of these are truely attributable to vertigo, then elimination of this causal factor would be a distinct asset.

Vertigo is defined as the loss of orientation which can be caused by physiological or psychological factors. The former of these may be induced by rotating the vestibular canals of the ear in orthogonal planes causing dizziness and possible loss of orientation.[5] Proper human engineering practice prohibits placing instruments in a position which is likely to result in such a head motion. It is presumed that this practice will be followed in future cockpit design. The psychological cause may occur when a disparity of orientation information is received

* Original manuscript received by the IRE, June 8, 1959.

† Reliability Group, Convair, Div. of General Dynamics Corp., San Diego, Calif.

[1] W. E. Vinacke, "Aviator's vertigo," *J. Aviation Med.*, vol. 19, pp. 158–170; June, 1948.

[2] B. Clark and A. Graybiel, "Vertigo as a cause of pilot error in jet aircraft," *J. Aviation Med.*, vol. 28, pp. 409–478; October, 1957.

[3] B. Clark and M. A. Nicholson, "Aviator's vertigo: a cause of pilot error in naval aviation students," *J. Aviation Med.*, vol. 25, pp. 171–179; 1954.

[4] A. F. Zeller, "Human factors in selected multi-engine jet aircraft accidents," *J. Aviation Med.*, vol. 29, pp. 197–205; March, 1958.

[5] "Vertigo," *Approach, The Naval Aviation Safety Review*, NAVAER-00-75-510, vol. 4, pp. 28–33; September, 1958.

by the human brain. The pilot monitors his attitude in space by reference to the attitude indicator and at the same time receives attitude information through the vestibular canals of the ear and the kinesthetic interoceptors distributed throughout the body. If these two information channels do not agree, the pilot must make a decision as to which is correct.

Standard pilot training has always taught pilots to believe their instruments as opposed to their own inner sensations.[6] It is understandable that this is a difficult task for a man who has spent all of his life obeying body sensations and only a comparatively short time monitoring the attitude instrument. To make things even worse, the pilot, when faced with such a situation, is often under stress and fully realizes that an improper decision might cost him his life. In fact, even the time he takes to make the decision may prove to be too great a delay. It is felt that the instruction "always believe your attitude instrument" is certainly a negative approach to this problem.

Conventional attitude instruments are inside-out; that is, they show a moving horizon as if the pilot remained aircraft oriented. Fig. 2 indicates such an attitude display as seen in a right turn at 45° bank. (The aircraft symbol is conventionally viewed from behind the tail.) This mode of attitude display assumes that the aircraft "up" is the only worthwhile reference axis.

Alternative to this mode of display, it is possible to present the same information in outside-in mode as shown in Fig. 3. Here the aircraft symbol becomes the moving member and the earth's "up" is the only reference axis. A considerable number of experimental investigations have been undertaken to determine which of these two modes is superior. The controversy, however, still remains unresolved.[7–10]

It is contended that these are not the only two possible modes for attitude display. The above described displays may be considered to form the zero velocity extremes of a continuum wherein both the aircraft symbol and the horizon move. This additional degree of generality permits continual reference to display design, from the viewpoint of the human operator's internal kinesthetic orientation, and in this manner eliminates the disparity of information between the visual input and that of the kinesthetic channel. It is recognized that the human body cannot be redesigned; however, the instrument can be designed to form a kinesthetic analog (termed "Kinalog") of the adaptive human. Instead of being aircraft oriented or earth oriented, the display should be pilot oriented so as to code the information in a most understandable manner.

Consider the kinesthetic sensing (the body's "view" of a coordinated right turn) at a 45° bank angle, for example. As this maneuver is initiated the pilot feels the aircraft tilt toward the right, however his "up" reference is still that of the earth. Fig. 4 shows the aircraft symbol moving in agreement with the human body's input. By the time the 45° bank angle has been attained, adaptation to the colinear g force (due to the coordinated turn) becomes apparent. The pilot's "up" will agree with the display as shown in Fig. 5. Note that the adaptation is reflected in slow counter-clockwise rotation of both the aircraft symbol and horizon. The true bank angle remains shown as the difference angle between the two moving members (see Fig. 6). As adaptation asymptotically approaches completion, the display appears as shown in Fig. 7, remaining thus until some new control action is taken. While continuing in this steady-state turn (constant bank angle) maneuver, the pilot's "up" will agree with the aircraft's "up." Initially, the attitude display was almost completely outside-in, but as time passed, it gradually proceeded toward inside-out in an exponential manner. The position along this inside-out–outside-in continuum is dependent upon both the time passage as well as the magnitude of the g force sensed by the pilot.

To recover to level flight the pilot actuates the control stick until the aircraft symbol becomes parallel to the horizon. Level flight has now been achieved, however, "up," as sensed by the pilot, is as shown in Fig. 8. As time proceeds both moving members rotate together as shown in Figs. 9 and 10.

If the above example had involved a pitch change, then the same kind of adaptation to the new pitch would have taken place. After the nose rises to the new pitch, both the aircraft symbol and the horizon gradually settle, until finally the aircraft symbol is once again in the neutral position. A sudden dive would appear as a drop of the aircraft symbol to a fall below the horizon bar. As adaptation takes place, the two moving members gradually rise until the aircraft symbol again reaches the neutral position. A more complicated maneuver would show both members moving simultaneously. The important thing to remember is that the instrument is always telling the truth about the present attitude. This information is read from the difference between the two moving parts. Both members "adapt" together without changing the difference reading while they continue to maintain the pilot's knowledge of personal "up."

If the pilot makes a rapid sequence of maneuvers or encounters turbulence, the amount of adaptation to each portion of the maneuver is very small and the dynamic display appears primarily to be the moving aircraft,

[6] "Instrument Flying," Air Force Manual No. 51-45, USAF, Washington, D. C.; April, 1956.

[7] S. E. Angrist, "The Comparative Interpretability of Two Aircraft Attitude Indicators," Aviation Psychology Lab., University of Illinois, Urbana, Memo. Rept. 57-8; February, 1957.

[8] "Pilot Performance with Two Different Attitude Displays," prepared for Douglas Aircraft Co. by Dunlap Assoc., Inc., Stamford, Conn., under ONR BuAer Contract No. 1076(00); July 21, 1955.

[9] "Evaluation of 'Moving-Airplane' Display," Naval Air Test Center, Patuxent River, Md., Rept. No. 11, Final Rept., Project TED, No. PTR AE-7058, 3 ST 312-320; November 14, 1956.

[10] J. F. Garner and R. J. Lacey, "An Experimental Comparison of Five Different Attitude Indicators," Wright Air Dev. Center, Wright-Patterson AFB, Dayton, Ohio, Tech. Rept. No. 54-32; May, 1954.

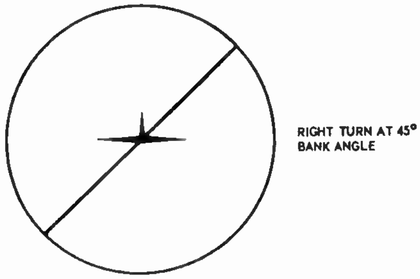


Fig. 2—Inside-out attitude display.

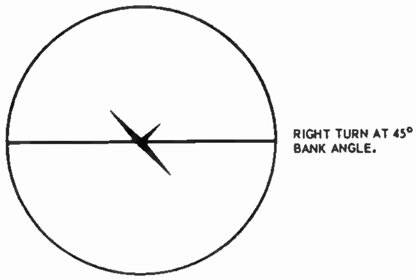


Fig. 3—Outside-in attitude display.

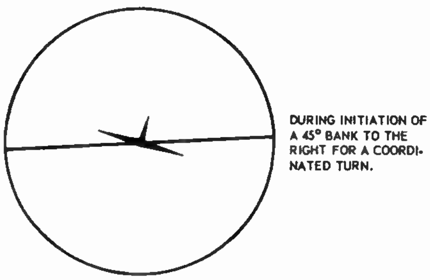


Fig. 4—Kinalog attitude display.

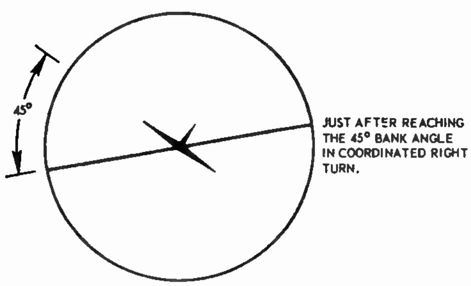


Fig. 5—Kinalog attitude display.

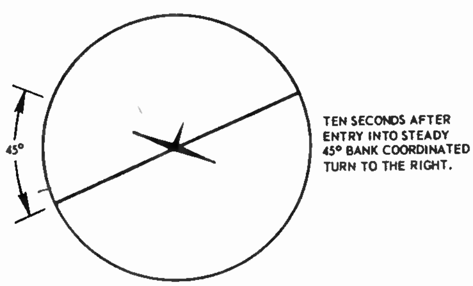


Fig. 6—Kinalog attitude display.

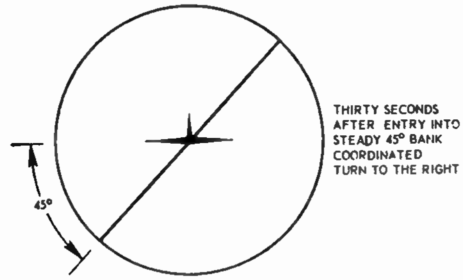


Fig. 7—Kinalog attitude display.

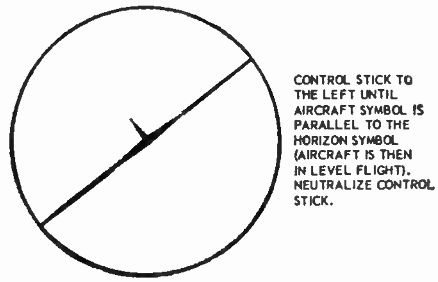


Fig. 8—Kinalog attitude display.

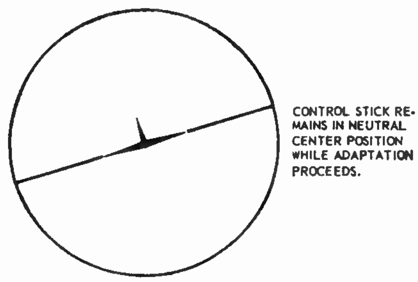


Fig. 9—Kinalog attitude display.

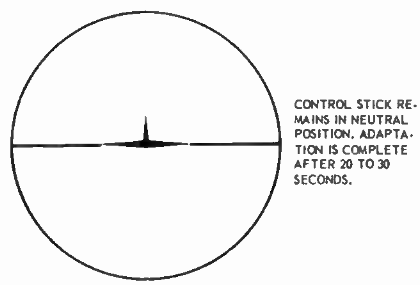


Fig. 10—Kinalog attitude display.

which is, after all, what is actually happening. On the other hand, if he maintains a constant attitude and a slow drift develops, the display will reflect what he might not ordinarily be aware of. Previous investigation has indicated the value of partitioning the frequency spectrum of the tracked variable so as to allow a combination of compensatory and the pursuit modes of tracking.[11] The kinalog mode accomplishes just this, in a most rational way, displaying the high-frequency portion of the spectrum in the movement of the aircraft symbol and the low-frequency portion in the adaptive response of both moving members.

It is claimed that all attitudes, and, in fact, all possible maneuvers, can successfully be displayed in the kinalog mode. It so happens that the view of an aircraft from a chase plane coincides with the kinesthetic "view" the pilot of the forward plane feels so long as the chase plane remains at the proper distance behind. Since the pilot of this fictitious chase plane can hypothetically chase the real forward plane through any maneuver, all attitudes and situations can be validly kinaloged. Use of this conceptual device allows visualization of the kinalog attitude display for *any* maneuver.

Determination of the consistency of human adaptation to various imposed *g* vectors has yet to be fully explored. Only such data can reveal whether a single adaptation response could prove satisfactory for all pilots. If not, the data might suggest that allowance be made for each pilot to adjust-in the proper response rate prior to flight. Past proposals for conversion from inside-out to outside-in were rejected because of the radical retraining which present pilots would be required to take, as well as because of the existence of a mixed inventory of aircraft. The generality of the kinalog technique forms a compromise between these two opposite display modes. Its fundamental compatibility with the human operator should remove the retraining requirement and allow all aircraft to be changed to kinalog.

In the near future man will be faced with new *g* and non-*g* conditions.[12] Inhibition of vertigo becomes of increased importance. It is felt that the kinalog technique for achieving time domain compatibility will prove to be a distinct positive approach to the long-neglected problem of vertigo.

The second problem relating to anticipatory display is a direct outgrowth of the increased speed and maneuverability of modern aircraft. This greater performance capability compresses the allowable reaction time to one where logical decisions and even conditioned reflex actions may no longer be possible. Even the simplest decisions require some finite time interval; for example, it is evident that the delay inherent in attitude decision-making may be sufficient to endanger life. To emphasize the immediacy of this problem, consider the following, as reported in the *Journal of Aviation Medicine*.[13]

> The pilot of a jet bomber was flying at 30,000 feet on a clear morning. He made a slow turn and was startled to see three other bombers approximately one mile away and on a collision course with him. Before he could react or alter the course of his aircraft he shot through the formation, missing the nose of the first aircraft, flying under the second, and over the third. As he went over the third bomber, one of his engines struck the upper part of this bomber's tail and knocked it off. The pilot who flew through this formation then returned to his home base, landed, and recounted his experience. Inasmuch as no report had been received from the formation he had flown through, it was called and requested to land. When it landed it was found that the formation consisted not of three aircraft, but of six. The aircraft whose tail had been hit was not significantly damaged, but what is amazing is that neither the pilot, the co-pilot, nor the observer in any of the six aircraft had seen the other bomber fly through the formation.

One way to combat this problem is through the use of anticipatory displays; that is, displays which present information relative to some future aspect of the observed parameters. In the simplest case, the display will indicate the predicted values for each parameter. The pilot "flys" this projection of his own aircraft's present status. Even though the values he observes may never come to exist, his control should prove more precise. Error is introduced by the prediction process; however, it is contended that this error, together with the pilot's error (which has been minimized by allowing sufficient time for decision making), will be far less than the error associated with the required almost-instantaneous human response.

Many questions immediately arise, such as whether or not all parameters should be projected to the same time in the future or whether each prediction time interval should be dependent upon the spectral properties of the signal and incident disturbances. Some empirical evidence is offered below in connection with the optimal prediction time interval for altitude under turbulent flight. Even the use of linear prediction enables naive subjects to accomplish altitude control comparable to that of experienced test pilots.

Anticipation may be accomplished at any point in the deterministic-stochastic continuum. The human operator has the ability to vary this point, even subconsciously. For example, the baseball outfielder initially perceives the flyball at a great distance from his glove and immediately starts to relocate his position on the field. This decision is based on experience; that is, reference to the statistical constants which have been committed to memory by repeated operation under similar circumstance. As the ball approaches, his final corrections can disregard more and more the effects of wind disturbance on the trajectory. His extrapolation of the airpath becomes more and more deterministic until, in the last instant, he positions his hand almost purely upon the ball's location and observed space derivatives. During the short time interval of the ball's flight, the player has gone from almost purely stochastic to almost purely deterministic prediction. It might certainly prove of value to utilize some type of stochastic prediction for the generation of anticipatory cockpit displays.

[11] J. W. Senders, "The Influence of Surround on Tracking Performance, Part I. Tracking on Combined Pursuit and Compensatory One-Dimensional Tasks with and without a Structured Surround," Wright Air Dev. Center, Wright-Patterson AFB, Dayton, Ohio, Tech. Rept. No. 52-229; February, 1953.

[12] H. J. von Beckh, "Weightlessness and space flight," *J. Appl. Rocket Soc.*, vol. 4, pp. 26–27; February, 1959.

[13] H. Moseley and V. Steinbridge, "The hostile environment as a cause of aircraft accidents," *J. Aviation Med.*, vol. 28, pp. 535–540; December, 1957.

Deterministic anticipation may be expressed as a series expansion of the weighted derivatives. The particular weightings may be chosen so as to be dependent upon the user; if this is done, it is certain that these will differ in some regard from ideal deterministic prediction represented by the Taylor Series Expansion. Happily, the instrument systems designer can choose any weighting set. In fact, by analysis of the imposed human data transformation, it is possible for him to choose weightings which will achieve control and minimize some error criterion. Such deterministic extrapolation is called "quickening."[14,15] When the derivate coefficients are intended to approximate the Taylor Series weightings, the special case of "prediction" occurs. Statistical prediction techniques are also mechanically feasible; however, they are usually much more costly in terms of equipment weight, space, and complexity.

The third problem is generated by the finite information channel capacity of the human operator. It is now recognized that there is an upper limit to the average information rate which a pilot can successfully process in accomplishing his flight control task.[16,17] Experimental evidence has shown that the error of human decision making rapidly increases if the display complexity offers an information rate which exceeds this channel capacity.[18] Today, a large portion of the pilot's time is spent in directly controlling the vehicle. Technology is providing him with more and more automatic equipments which he must monitor and program in order to direct the flight operation. Future pilot decisions may be fewer in number, but individually they will have much greater importance in accomplishing the mission. Less error will be tolerated. The instrument display must be designed to insure that its complexity will not exceed the channel-capacity of the human operator.

All consciously-perceived information derived from any display takes up a portion of the available capacity for information processing. It is clearly evident that all data considered to be irrelevant to the required decisions should be eliminated from the display. Such irrelevant information forms noise, in the sense that it can only result in degraded decision making.

The pilot's view of the real world is not necessarily the best display to allow flight control. Consider certain disadvantages which are inherent in the "natural" displays. From a psychological standpoint, riding in a closed skyscraper elevator is more satisfactory than riding in an open one. Benefit results from obstructing the view of the shaft walls which would move with extremely high velocity through the observer's field of vision. A similar situation prevails when the pilot endeavors to perform critical maneuvers at extremely low altitude. The earth moves through the field of vision so rapidly that, even if resolution remained, insufficient time would be available to allow execution of even emergency corrective measures. The world beneath is a blur. The only received information might stimulate confusion or fear. The pilot surely would be more effective if he were on meaningful, predictive instruments.

Another aspect which deserves critical attention is the value of information which a pilot may perceive even under the best of flight conditions. His discrimination of roll and pitch angle is poor, except under the special condition of nearly level flight (when he needs these least). Judgment of the aircraft altitude or the altitude of a mountain peak is definitely poor (being some inverse function of the aircraft altitude). The reliability of estimating weather from the observable cloud formation is even lower. Heading discernment is nil, except when flying level or directly into a bright sun. Position can be determined only when the terrain is familiar or extremely well marked by towns, rivers, etc., and it should be noted that this discrimination is also some inverse function of altitude. Crab angle and various other parameters are viewed with such poor discrimination and reliability as to make them unworthy of consideration as true human-sensed information.

A pilot is sometimes reluctant to initiate instrument flight as the weather becomes worse. It is under such conditions that he must assimilate the perceived real world display during intermittent and usually short intervals of continuous observation. The detail and complexity of the observed earth probably makes its transmission rate far in excess of the human channel capacity and thereby further lowers the effectiveness of the operator. It is this complexity that often contributes to the deceptive quality which may be encountered even under CAVU (ceiling and visibility unlimited) conditions.

Any assumption that a view of the real world is optimal tacitly assumes that the earth was formed in its natural state for the purpose of allowing pilots to fly over it. The noise due to rough or obscure terrain might best be eliminated by furnishing a fictitious earth to the pilot, one which would appear as a map of the surrounding region with only predominant landmarks and characteristics identified. For example, such a viewed fictitious earth could be in three dimensions wherein the mountains could be replaced by polyhedrons defining the unsafe zone, color coded in altitude. Such "modified pictorialism" would reduce the disturbing elements of nature's presentation and thereby enhance the information to the pilot. This, most probably, is not a practical display design; however, it does serve to point out that

[14] H. P. Birmingham and F. V. Taylor, "A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems," Naval Res. Lab., Washington, D. C., NRL Rept. 4333; April 7, 1954.

[15] H. P. Birmingham and F. V. Taylor, "A Demonstration of the Effects of Quickening in Multiple-Coordinate Control Tasks," Naval Res. Lab., Washington, D. C., Rept. No. 4380; June, 1954.

[16] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Exper. Psychol.*, vol. 47, pp. 381–391; June, 1954.

[17] J. R. Pierce and J. E. Karlin, "Reading rates and the information rate of the human channel," *Bell Sys. Tech. J.*, vol. 36, pp. 497–516; March, 1957.

[18] H. Quastler, "Human Performance in Information Transmission," Control System Lab., University of Illinois, Urbana, Rept. No. R-62; March, 1955.

a modified pictorial display can be superior to the exact representation of the real world.

The reduction in display complexity benefits decision-making by reducing time and increasing accuracy. Required training is minimized since the modified pictorial mode of display retains the interrelations among the parameters of the real world display and, by so doing, retains agreement with all of the pilot's past experience. The naturalness of such a display is the best promise of successful rapid response to a posed problem under the stress of life-or-death conditions.

The discussion section will present proposed flight control instruments which embody solutions to these three primary problems. Each instrument design will be introduced with an outline of the individual problem recognized as pertaining to decision-making on the particular parameters under consideration. The devised solutions are incorporated into features of the display designs, always remaining within the context of the capability of the human operator. The flight controls offer both visual and tactual information feedback to the pilot. In this sense, they are also a display. Aspects of control design will be reviewed in order to complete the design of a compatible information input source for the human operator.

Discussion

The "Window" Display

The most informative view from an aircraft cockpit is usually found looking forward and slightly downward. The forward view yields attitude information while the downward view offers earth-reference data. Here is a totally integrated display; that is, integrated in the sense that the same member is referenced to portray roll, pitch, altitude, heading, position, and various other flight parameters. This integration is performed in the most natural manner possible in that it relates to all the past experience of the human operator, even before he became a pilot.

Fig. 11 indicates the wide-angle "window" display configuration. Note the ease of interpretation. The particular situation presented shows the aircraft in an attitude of about 20° right bank with 10° pitch up (level flight would show the tail-viewed aircraft symbol superimposed on the horizon). The triangular aircraft shadow symbol represents the position relative to the terrain. Tilting of this triangle indicates the crab angle. These symbols remain connected by a fixed vertical line which is inscribed on the instrument bezel indicating the forward flight path over the ground. The terrain is almost undistorted under the shadow symbol and becomes increasingly foreshortened toward the horizon. The range to any point under the imminent purview of the aircraft may be judged in relation to the range rings which appear on the displayed terrain representation. Heading may be judged by reference to the compass roses which are placed periodically on the map to insure that at least one will always be easily readable. Terrain altitude may

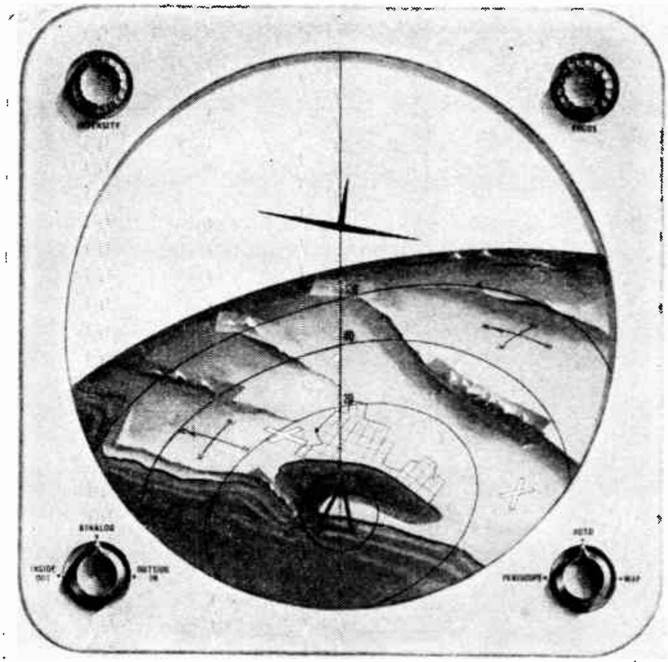


Fig. 11—The modified pictorial "window" display (in color).

be read from the indicated color coding designating mountainous areas; for example, a red zone might require flight above 15,000 feet while a yellow one might indicate a minimum acceptable altitude of 10,000 feet. Cities are simplified in pattern and almost all titles are omitted, since it is presumed that the pilot has previously become familiar with the terrain under the intended flight plan and can make reference to available maps should more detailed data be desired. Emergency runways are enlarged. A minimum of printed material appears on the map and all extraneous data are removed.

Velocity may be judged by noting the speed with which the terrain appears to pass under the shadow symbol in a similar manner to velocity judgment when looking out of the aircraft. Altitude may be judged both from the scale of the terrain beneath the aircraft shadow symbol and from the curvature of the horizon. An increase in altitude makes the earth appear to shrink away, and a dive apparently brings the earth closer.

Note that the display as shown in Fig. 11 is neither inside-out nor outside-in. The kinalog display indicates that a climbing right turn has just begun. As this attitude is maintained, the pitch and bank angles remain accurately displayed in terms of the fixed difference distance and angle between the tail-viewed aircraft symbol and the horizon, while both of these symbols gradually shift toward placing the aircraft symbol in a central horizontal position in accordance with the adaptation rate of the human pilot under the colinear g force which he senses. Reference is made to Figs. 12–16 which indicate a maneuver in kinalog "window" display as compared to more usual attitude instrumentation. A three-position switch could be offered, as shown at the lower left. This would permit older pilots to retain the

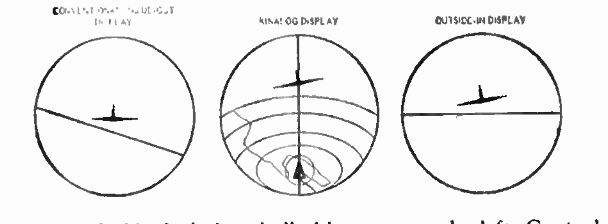


Fig. 12—Suddenly induced climbing turn to the left. Control stick is pulled back and to the left.

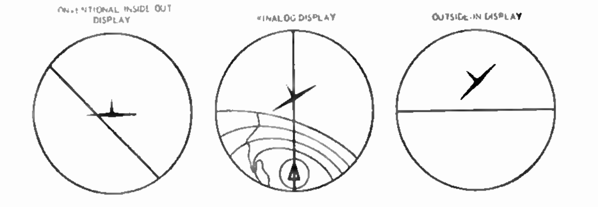


Fig. 13—Control stick is neutralized. Aircraft continues to climb at the same pitch and bank angle.

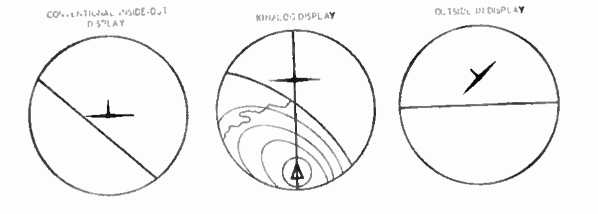


Fig. 14—Still climbing and turning several seconds later (after adaptation).

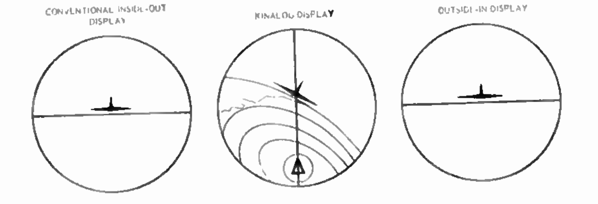


Fig. 15—Just as level flight is attained.

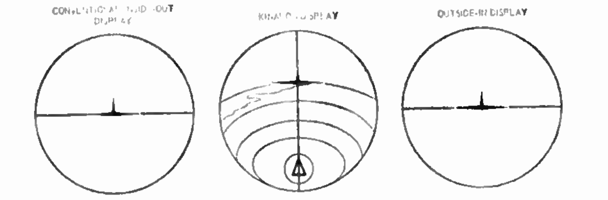


Fig. 16—After adaptation of kinalog (and pilot) has completed.

inside-out mode, should they so desire. The selection would be of particular value during initial flight evaluation.

Cursory evaluation of the display has been accomplished through the use of a jury-rig ground based simulator. "Flight" was permitted at several fixed altitudes and velocities utilizing any one of the three attitude display modes. Projection of a color motion picture of a scaled "flight" over a chosen ground path was remotely controlled by the "pilot" and permitted two degrees of freedom. The use of mirrors and damped spring "adaptation" displayed a modified pictorial view over the range of $\pm 60°$ pitch and $\pm 80°$ bank angle. Thirteen Convair test pilots were familiarized with the simulator and accomplished sufficient control to form a first judgment. Opinion was unanimously in favor of the general "window" display. Ease of interpretation was acknowledged. Difference of opinion was noted with respect to the value of the kinalog mode of attitude display. This is understandable in view of the fact that it was necessary for the "pilot" to imagine the *g* force as he went into a maneuver. This alone would be difficult to do. It is even more difficult, if not impossible, to imagine his own adaptation to *g* force. Only a centrifuge or flight environment would permit true pilot evaluation. The simulator display could be imagined to be anticipatory without any difficulty since such operation would in no way be apparent to the "pilot."

Fig. 17 indicates the conceptual design for the "window" display instrument. A light source is concentrated upon that particular portion of the map color transparency which corresponds to the domain under the immediate forward purview of the aircraft. Navigation devices such as VOR-DME or TACAN allow a computed heading and ground velocity. These quantities drive their respective servos to maintain proper position track. (Heading rotates the map, while forward ground velocity determines the velocity of the map linear movement.) The image thus formed passes through the concentric range rings which, conforming to the scale of the map, are inscribed on the succeeding transparency. The light then passes through the prismatic distortion lens which foreshortens that portion of the image corresponding to the pilot's visual horizon. The lens shape curves the horizon. Its position is the resultant of two independent servos which are driven respectively by sensed pitch and bank angle. Translation due to pitch change translates the image, while rotation of the cylinder rotates the image to the appropriate bank angle. The perspective distortion can equivalently be accomplished through the use of a nonplanar mirror. This might allow lower cost and reduced physical length of the display generating black box.

The image is then acted upon by the variable-magnification constant-focus lens system. This is the same lens system as is used on modern TV and motion picture cameras for "zoom" pictures. This lens system's servo is driven by the sensed altitude; thus it provides increased magnification with decreased altitude. The image proceeds to the viewing screen.

Separate light sources provide illumination for the aircraft symbol silhouette transparencies. The tail view symbol transparency is servo rotated by the sensed bank angle and raised or lowered on the screen by the servo tilted mirror corresponding to pitch. A coupled silhouette transparency shows a "wheels down" view. Actuation of the landing gear interchanges these to generate the display shown in Fig. 18. It would certainly be difficult to misinterpret such wheels status information. The

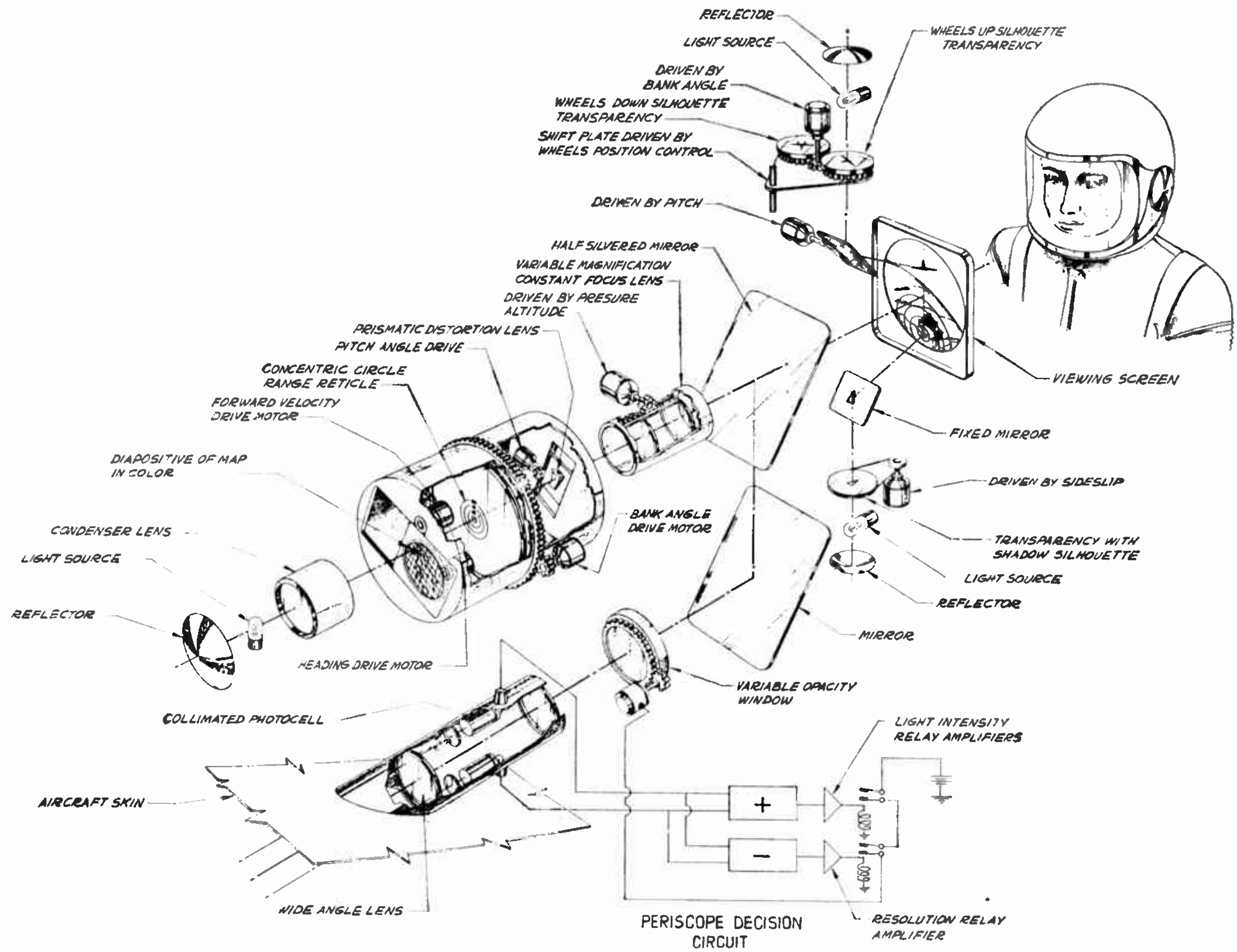


Fig. 17—Proposed optical projection system.

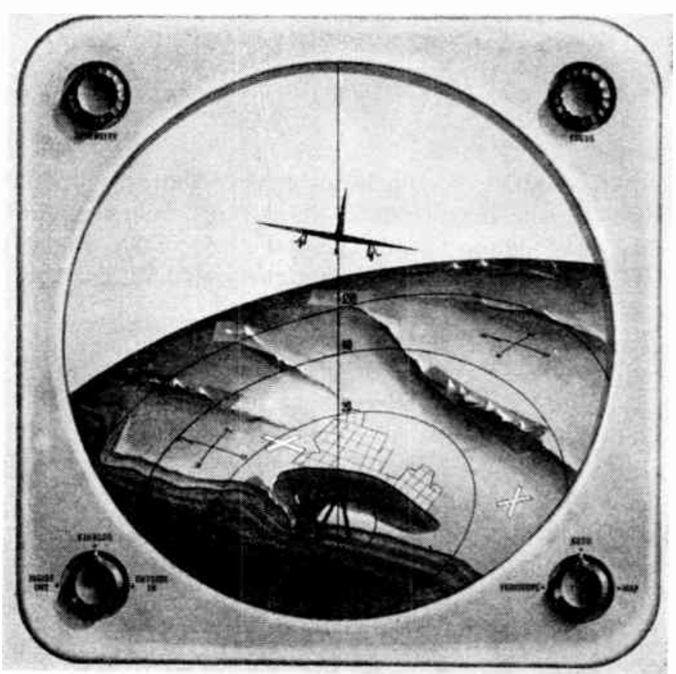


Fig. 18—The modified pictorial "window" display (wheels down).

lower aircraft shadow symbol silhouette transparency is rotated by the sensed crab angle. The final image is the resultant composite of the three images.

Should complete terrain realism be desired, it is only necessary to substitute a composite aerial photo for the map transparency. On the other hand, flight over ocean or unknown territory would project a grid map in terms of latitude and longitude. Particular points of interest could be indicated by a marker. The modified pictorial display shown in Figs. 11 and 18 is considered to be an optimal compromise between these two extremes.

It is important to note that the required display generation has utilized only the normally sensed variables. Computation is carried out by the independent successive transformations of the optical image. Here is relatively simple, reliable computation when compared to the average complexity of electronic coding, computation, and CRT display generation to accomplish a similar image with equal resolution and range of colors. The optical system appears to offer distinct advantages in cost, simplicity, information memory capacity, ease of servicing and, possibly most important, in-flight reliability.

Since the "window" display has the same general appearance as the real world, transition from visual flight to instrument flight can be facilitated by providing a periscope view of the outside world while under visual flight conditions. This periscope would utilize a wide angle lens so that the view of the real world would encompass that presented by the instrument flight artificially generated display. Fig. 11 shows a switch at the lower right permitting map, periscope, or automatic mode of display. In the automatic mode, two or more collimated light photo cells generate a voltage proportional to the intensity of the point illumination of the real world image. If the sum of the voltages is above a prechosen value, the view is sufficiently bright to be of possible interest, this, however, might also correspond to viewing a bright cloud or the sun. The difference voltage is also taken to remove this ambiguity. If this difference also exceeds its prechosen threshold value, it is indicative of sufficient resolution to provide an interesting observation. Closing both relay switches clears the variable opacity window or opens an aperture allowing light to pass to the viewing screen. At the same time, appropriate switching removes the transparency-generated image. Should either the real world average light intensity or the disparity between the individually-sensed light levels fall below some prechosen value, the aritficial display will be resumed. Some hysteresis might be added to the switching to insure that oscillatory transition will not occur. Incorporation of the periscope removes the necessity for the projecting canopy. A zero drag cockpit is often greatly desired.

A half-silvered mirror would permit superposition of a PPI radar image on the undistorted portion of the viewed map or terrain. Coincidence assures the validity of automatically computed navigation. The radar may portray targets of immediate concern such as ships, these appearing in the proper pictorial context. It is also possible to superimpose fire control or collision avoidance information portraying other aircraft relative angle, range, and attitude.

Since the "window" display presents a replica of the real world as the common denominator of primary interest, it is suggested that this instrument be placed in the center of the forward panel, tilted to about 20° to provide additional agreement with the direction of "observation." (See Fig. 19.)

The above described instrument concept may be modified for use in particular vehicle types or may even be designed to the specific intended mission. For example, consider a possible manned satellite primary display. Fig. 20 shows such a display-generating mechanism. A light source inside the transparency globe has a shield affixed to it, and thus illuminates only the "daylight" portion of the "world." This hemisphere light source is servo-rotated diurnally with respect to the globe. The "night" portion may be illuminated by ultraviolet or point sources on the surface. The globe itself is rotated about its own axis by a servo motor which is driven by the magnitude of the ground velocity. Another servo motor rotates the globe axis in accordance with the orbital heading so that the view will maintain proper ground track.

If the takeoff trajectory is as shown in Fig. 21, the desired "view" would initially be directly downward and would be gradually tilting forward as a direct function of the forward ground velocity. The orbital "view" would be forward and somewhat downward, much as in level aircraft flight. Thus it would provide a pictorial "window" display. Vertical translation of the globe support is provided by a servo so that the optical centerline may be shifted toward a parallel tangent to the "earth's" surface. This action tilts the view as desired. A stationary platform supports a zoom lens which is altitude controlled.

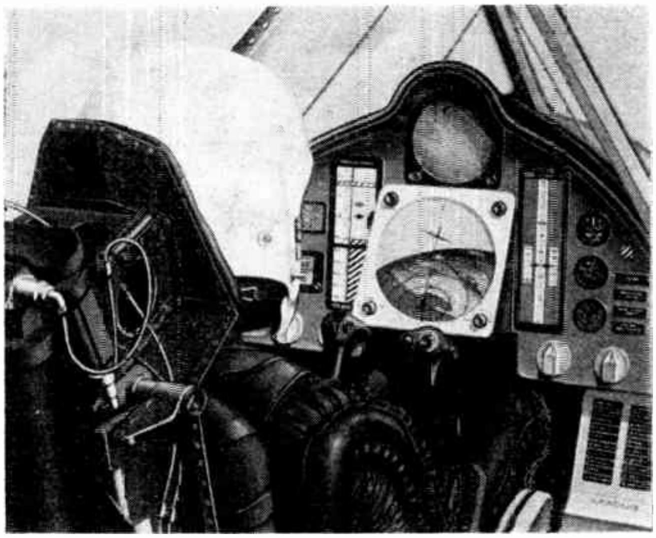


Fig. 19—The conceived cockpit display panel.

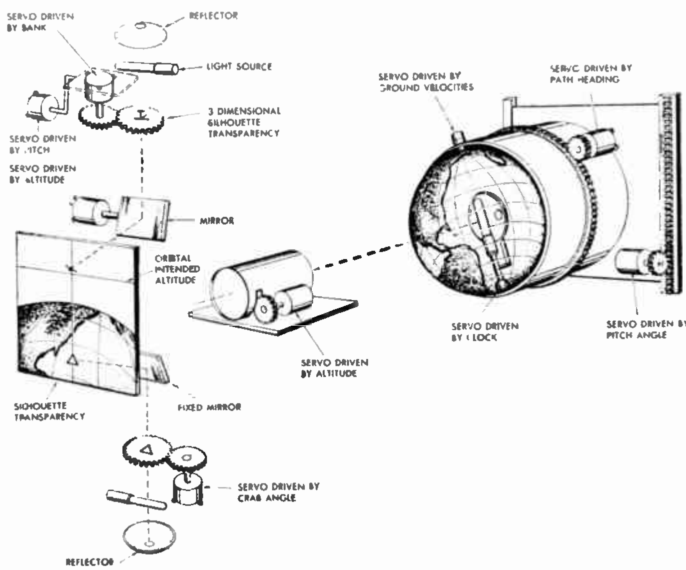


Fig. 20—Proposed satellite display generating system.

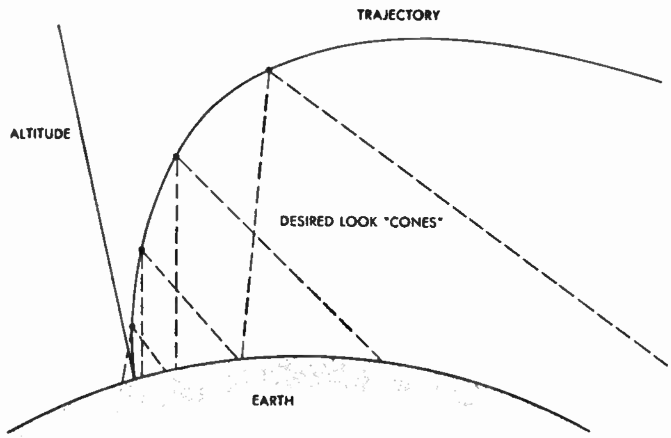


Fig. 21—Desired view angle vs trajectory.

Another light source is concentrated to pass through a three-dimensional silhouette transparency. This transparency is rotated about two axes in accordance with attitude information providing a pictorial tail-view aircraft symbol which portrays attitude. Appropriate electronic coupling would permit the kinalog attitude display mode. The image is reflected onto the viewing screen by a mirror which is tilted in accordance with altitude information. The aircraft shadow symbol is generated in an obvious manner and displays position over the terrain as well as crab angle. Such a display should certainly be superior to a true window view if the vehicle is rotating to create an artificial *g* force.

Although most flight control decisions are accomplished on the basis of the coalition of qualitative data on various parameters, certain situations arise in which specific numeric values for parameters must be available. Additional instruments are proposed to provide this information.

The Rate-of-Climb, Planning, and Predictive Altimeter

The display of altitude data presents an interesting problem. Traditionally the two types of quantitative displays available are moving pointer (outside-in) and moving tape (inside-out). Each of these has certain advantages and disadvantages as indicated in Fig. 22.

The primary advantage of a moving pointer display is its distinct pictorialism. The moving pointer directly represents the moving aircraft and, if a vertical linear altitude scale is utilized, direct real world analogy results. The meaning of the data is preserved in its topological representation. It is this property which inhibits the 1000- and 10,000-foot errors which have been characteristic of the three-pointer circular dial altimeter.[19] Unfortunately, an intrinsic disadvantage of this quantitative display mode is the low precision with which the numeric value can be read; this is particularly true for increased altitude capability aircraft.

ALTIMETER	OUTSIDE-IN (MOVING POINTER)	INSIDE-OUT (MOVING TAPE)	
ADVANTAGES	HIGHLY PICTORIAL	HIGH PRECISION	
DISADVANTAGES	LOW PRECISION	POOR DYNAMIC	POOR STATIC

Fig. 22—The altimeter problem.

Moving tape display mode overcomes this deficiency since the tape can be made long and the result is any degree of desired precision. The corresponding disadvantage in the dynamic situation is created when the numbers are conventionally arranged in increasing order with the largest number at the top. Increase in altitude is shown by descending numbers. The direction of display motion should obviously correspond to that of the real world.[20] To overcome this problem, the number scale can be reversed. This would result in dynamic compatibility; however a new problem looms. The reading of altitude in the level flight situation requires interpolation by subtracting from the value below the lubber line, rather than adding to this value as is the stereotyped convention. The problem is to devise an altimeter display which retains both the advantages of moving pointer and moving tape while eliminating both disadvantages.

Fig. 23 indicates such display configuration. Analogy to the real world is established by the linear vertical scale on the fixed background. A master index is imprinted on a transparent tape having opaque areas to offer two windows which are bisected by the master index arrows. Quick reading of this index with respect to the background scale offers qualitative information on the total altitude capability of the aircraft. Closer observation reveals that the indicated altitude is between 20 and 30 thousand feet; or, more precisely, just above 26 thousand feet.

A separate tape travels vertically to the right and under the master index; this tape may be seen through the right rectangular window. If still more precise quantitative reading is desired, a vernier value may be seen adjacent to the master index arrow. This tape may be of any desired length thus providing any degree of precision. Reading of altitude is accomplished from left to right so that the most significant figures are read first (in a pictorial manner) and the vernier figures are read second (with high precision) according to convention. The vernier numbers are arranged in the standard order of as-

[19] P. M. Fitts, "Engineering psychology and equipment design," in "Handbook of Experimental Psychology," S. S. Stevens, Ed., John Wiley and Sons, Inc., New York, N. Y., pp. 1301–1306; 1951.

[20] P. M. Fitts, "The influence of response coding on performance in motor tasks," in B. McMillan, *et al.*, "Current Trends in Information Theory," University of Pittsburgh Press, Pittsburgh, Pa., pp. 47–75; 1953.

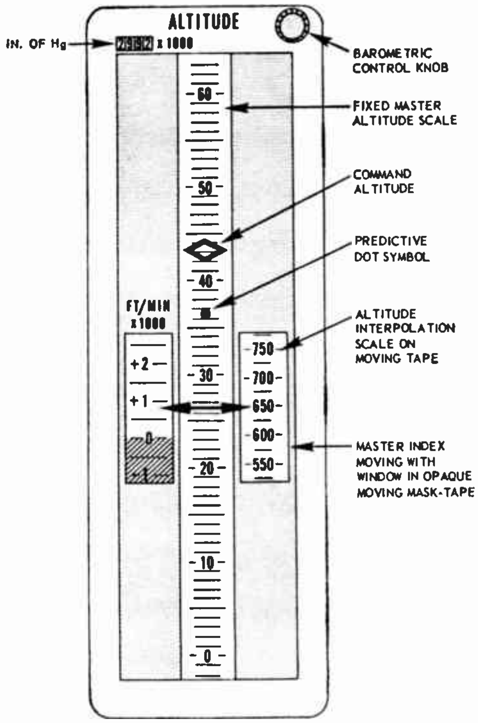


Fig. 23—Rate of climb, planning and predictive altimeter.

cending magnitude permitting the usual interpolation, but they do travel down as altitude increases. This disadvantage may be overcome through the intended use of an optical illusion. It has been shown in psychophysics that the frame of reference can dominate the field of view.[21] If the rectangular frame is of proper size, the numbers should not offer a dominant incompatibility since this frame of reference moves compatibly with altitude change. Only cursory experiments have been performed to determine proper window size over a range of rate-of-climb and, although the findings thus far are inconclusive, it is felt that such design is possible. It is important to note that this compatible altimeter requires only simple mechanical components. Further, it inhibits pilot errors and the increased time delay found in reading circular dial altimeters of various types.[22,23]

The rate-of-climb numeric value is found in the window to the left of the master index. The proper value is always colinear with the index (and the vernier altitude value). Any degree of precision is possible by the free choice of tape length. Pitch is observed on the central "window" display. Reading to the right, the resultant rate-of-climb is found to be followed by the altitude, displayed in accord with number reading convention.

Both target and predictive altitude indexes are provided, each separately moving over the linear background gross altitude scale. The second of these provides an estimate of the altitude at some particular time in the future. Experiment was accomplished to determine the optimal predictive time interval for a simple linear predictor used for level flight under imposed atmospheric turbulence.[24] Fig. 24 outlines the test arrangement. The "pilot" was provided separate indications of his present altitude and predicted altitude. His task was to maintain level flight. Fig. 25 is a photograph of the test set up. The "pilot" is provided with a one dimensional "control stick" and a CRT linear altitude scale. His galvanic skin resistance is taken with the aid of two electrodes placed on the left hand to measure the amount of autonomic nervous response. This measure is indicative of the amount of mental effort expended. Four naive subjects and four test pilots were utilized, the result was a 50-second average error score as shown in Fig. 26.

The nonpilot error was considerably greater than pilot error under zero seconds prediction (a single altitude index). Pilots appear to have taken advantage of the added information when two "bugs" appeared (the predictive index first representing the altitude six seconds into the future). At a predictive interval of 12 seconds, the average error scores were identical and, in fact, they remained in close proximity until a prediction interval of about 42 seconds. Thereafter, both error scores increased, as would be expected in view of the lower degree of relevance to immediate control decisions. The important fact demonstrated is that, even through the use of simple linear deterministic prediction, a distinct benefit can be attained.

Use of a predictive altitude display has provided an equivalent increase in level flight control capability of thousands of hours of flying experience, and it has accomplished this at almost insignificant cost.

The Velocity Instrument

Forward motion may be qualitatively observed from the central "window" display. This value may be refined into numbers by reference to the display directly to the left. Fig. 27 shows a configuration which is a simple modification of the significantly improved velocity display offered by the Hughes Aircraft Company.[25] A single moving member represents the aircraft and may be read either with respect to indicated airspeed (when it appears below the crossmember of the instrument

[21] C. H. Graham, "Visual perception," in "Handbook of Experimental Psychology," S. S. Stevens, Ed., John Wiley and Sons, Inc., New York, N. Y., pp. 895–901; 1951.

[22] C. W. Simon, C. K. Slocum, C. O. Hopkins, and S. N. Roscoe, "Altimetry Studies: I. Experimental Comparison of Three Pictorial and Three Symbolic Displays of Altitude, Vertical Speed and Altitude Command," Hughes Aircraft Co., Culver City, Calif., Memo. No. 425; January 1, 1956.

[23] C. W. Simon and S. N. Roscoe, "Altimetry Studies: II. A Comparison of Integrated versus Separated, Linear versus Circular, and Spatial versus Numeric Displays," Hughes Aircraft Co., Culver City, Calif., Tech. Memo. No. 435; May 1, 1956.

[24] L. J. Fogel and M. Dwonczyk, "Anticipatory display design through the use of an analog computer," 1958 WESCON Convention Record, pt. 4, pp. 67–88.

[25] "Psychological Aspects of Cockpit Design—A Symposium Report," M. L. Ritchie and C. A. Baker, Eds., Wright Air Dev. Center, Wright-Patterson AFB, Dayton, Ohio, pp. 57–117; April, 1957.

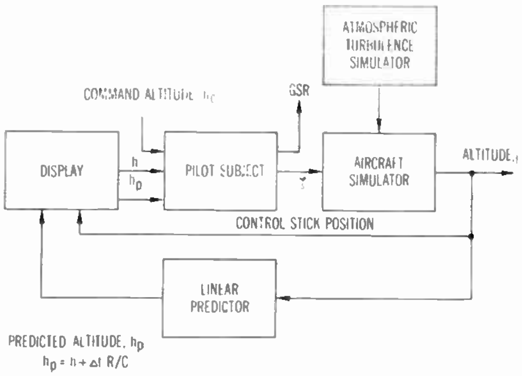


Fig. 24—Simulated man-machine system.

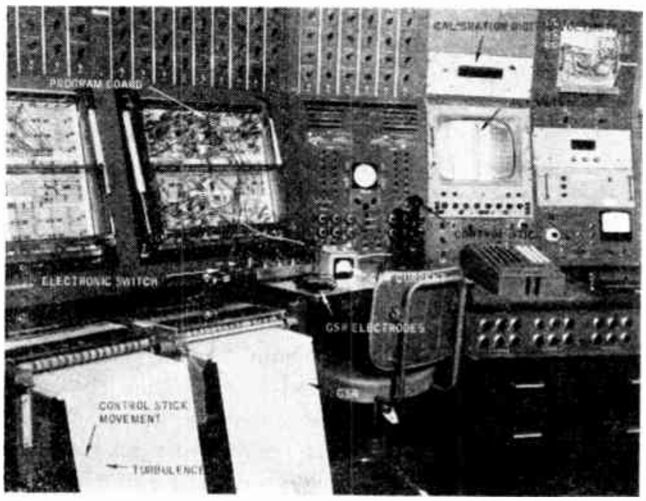


Fig. 25—Layout of the experiment.

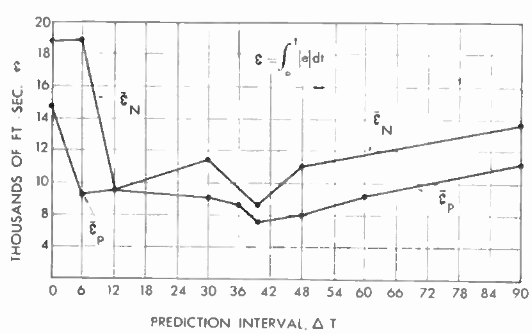


Fig. 26—Error score for pilots and nonpilots.

frame) or in terms of Mach number. Qualitative angle-of-attack may be read by reference to the silhouette marker which offers a point corresponding to max L/D ratio. This silhouette is vertically displaced by the air data computer.

Note that this display offers *either* indicated airspeed or Mach number permitting the single moving pointer. It is contended that no significant loss of information results and that simplicity is gained. The cross-frame member is supplied to emphasize the change of dimension. The scales are colinear to facilitate transfer. No significant fabrication difficulties should be introduced by this proposed configuration.

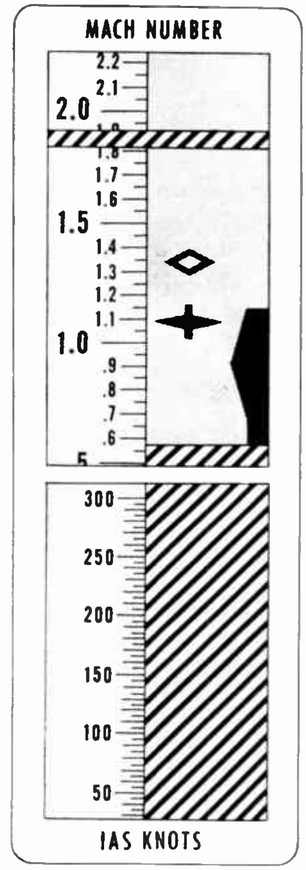


Fig. 27—Proposed velocity indicator.

The Time Display

Velocity is distance per unit time. Time is shown in an adjacent display. Reference to a clock is usually made in order to determine the time interval between the present and some fixed time. The mental process of analog to digital conversion can be eliminated by presenting numerics corresponding to present time and the remaining time to some ETA (estimated time of arrival) which may be inserted by a rotary knob. See Fig. 28. Additional difference time intervals may also be shown. A second hand dial is provided to assist the piloting of such maneuvers as a 30-second turn. Such observations normally remain as an analog positional difference representing the desired passage of time.

Engine Instruments

In today's conventional engine instruments, prechosen tolerance values are marked by colored adhesive tape or other inscription. Instruments such as these are first read to determine if the parameter is within the proper range of value. If not, more careful reading is required to interpret the pointer position in terms of a numeric. The instruments shown in Fig. 29 accomplish these initial decisions for the pilot.

After reaching cruise conditions, the pilot depresses the push-to-light switch which greatly dims the engine instruments. Each display is then automatically illumi-

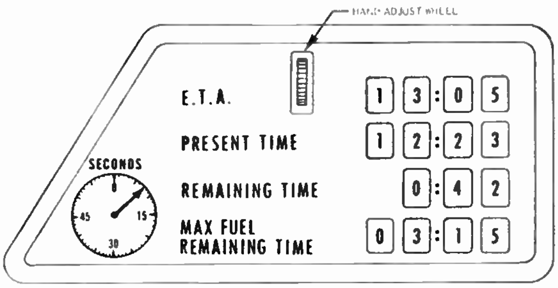


Fig. 28—Proposed time display.

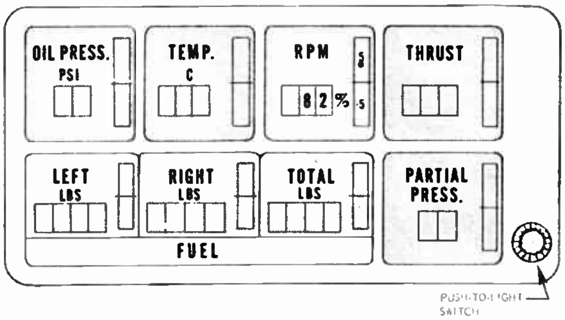


Fig. 29—Engine instrumentation.

nated, should its particular parameter fall out of the proper range. In such a case, the instrument directly presents the numeric value for the variable together with a quantitative moving dot reference for the derivative. Should the parameter fall into a dangerous range, the display illumination turns red. The actuating circuitry is obvious, the tolerance voltage levels being functions taken from the air data computer using *a priori* knowledge of engine characteristics.

Note that the individual engine parameters are arranged in a logical pattern. Reading to the left of the velocity display; velocity is caused by thrust, pressure ratio, engine rpm, temperature, oil pressure, and remaining fuel. This leads logically to the left-hand throttle display. This proposed set of engine instruments relieves the pilot of unnecessary mental effort and provides increased reliability in the task of monitoring for low probability events.

The Approach and Landing Display

Resolution and other limiting factors may prohibit the use of the "window" display for final approach and landing. It is intended, therefore, that this central display be utilized until final approach has been initiated at terrain altitude of approximately 2000 feet. The "window" display remains at this lowest altitude during the remaining portion of the flight while the pilot transfers his primary attention to the CRT directly above the "window."

Four parameters are utilized for successful approach at a particular indicated airspeed; they are:

horizon tilt, indicating aircraft bank angle;
horizon line vertical position, indicating aircraft pitch angle;
runway-end angle(s), indicating azimuth of approach; and,
normal projection of the distance from the forward ward edge of the runway to the horizon, describing the angle of aircraft descent.[26,27]

One technique for generating this required display information might be to utilize passive electromagnetic reflectors or active beacons implanted at the borders of the runway. Airborne radar could scan ahead of the aircraft and present return information on the CRT together with a horizon trace generated from the gyro. A perspective view would result as shown in Fig. 30. Approach is flown in the usual manner while keeping the bezel-inscribed cross marker on the intended touchdown point. Interpretation of this display is simple; witness, for example, the ease of interpreting the incorrect approach conditions which are shown in Fig. 31.

Flare-out is initiated by reference to the radar portrayed ellipse which consists of reflectors or beacons buried in the runway surface. The viewed length of the major axis (horizontal) is a direct function of the aircraft forward velocity; the length of the minor axis (vertical) is a direct function of the height of the aircraft above the runway. The pilot adjusts his control so as to make the ellipse become a circle within the annular tolerance ring which is part of the bezel-mark. When this condition is satisfied, he can flare-out in the knowledge that he is within the proper tolerance on both position along the runway as well as height above the runway. If the ellipse never becomes sufficiently circular, prior to its major axis' exceeding the outer edge of the ring, then he must "go around." This display technique forms a modified pictorial go or no-go gauge. Taxi instructions could be received from additional radar coded beacons implanted at significant points. It is questioned whether present radar capability can accomplish these requirements; however, the value of the display itself might furnish sufficient impetus to initiate its development.

Critical Maneuver Display

This same CRT could be called upon to aid in accomplishing certain critical maneuvers. For example, the pilot, should he be concerned with minimum time, minimum fuel expenditure, or some compromise climb, may select a mode as shown in Fig. 32. Altitude forms the ordinate, while distance ahead or future time forms the abscissa. The predicted aircraft trajectory function is shown as a dotted trace together with the computed minimum time and minimum fuel trajectory functions. The pilot operates his controls so as to place his projection upon a chosen path; in this way his accuracy of control is continually apparent. So long as his own projec-

[26] S. N. Roscoe and A. C. Williams, Jr., "Pilot Performance in Instrument Flight as a Function of the Extent and Distribution of Visible Horizon," USN, Special Devices Center, Port Washington, N. Y., Tech. Rept. SDC 71-16-3; June, 1949.
[27] S. N. Roscoe, "Flight by periscope," *University of Illinois Bull.*, vol. 48, pp. 1-46; March, 1951.

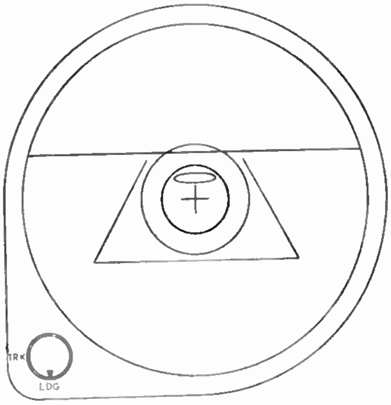


Fig. 30—Landing approach and flare-out CRT display.

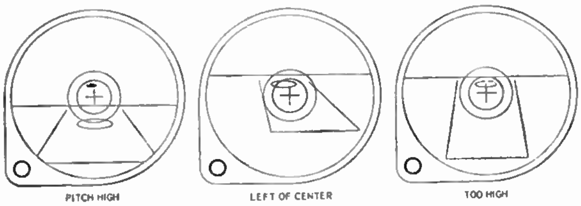


Fig. 31—Poor approaches.

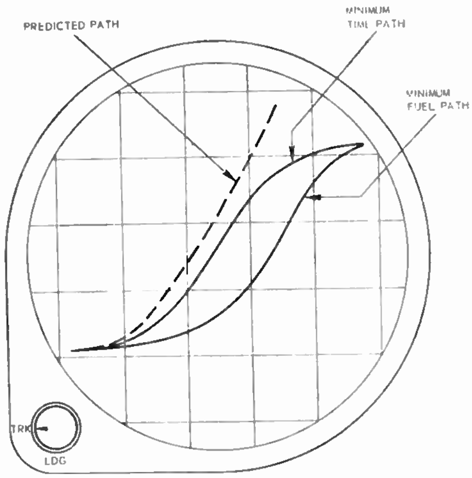


Fig. 32—Optimal airpath tracking CRT display.

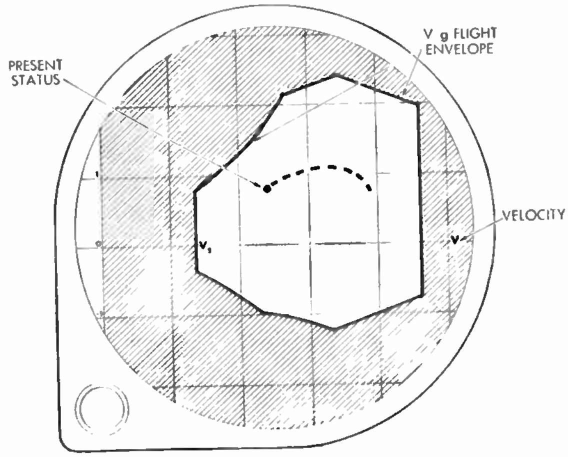


Fig. 33—Flight envelope CRT display.

tion is maintained at an intended target point, he may choose to minimize error continually, or he may trade error over different portions of the abscissa. This latter flexibility offers the pilot an additional degree of freedom which is only available through the simultaneous display of the temporal dimensions.

Pilots may hesitate to put modern aircraft through their maximum maneuverability in view of the lack of immediately available safety status information. Fig. 33 shows the aircraft predictor function presented on the CRT under a side-illuminated *V-g* diagram descriptive of the aircraft's maximum capabilities. Such a display allows the pilot to press the aircraft near to its limits with knowledge of safe flight. Generation of this display is considered to be within the present state-of-the-art. Additional reference may be made to other maneuver information displays provided to the right of the altimeter in close conjunction with the related right-hand attitude control stick.

Compatible Controls and Surround

Controls furnish feedback information to the pilot, both visually and tactually through the "feel" system. Compatibility may be maximized if there is correspondence between the control stick grip attitude and the intended attitude of the aircraft. Present control systems may be improved in this regard by placing the center of rotation just below the control "stick" handle. This would allow longitudinal angular displacement to correspond to the pitch rate and transverse angular displacement to the roll rate. It would seem desirable to have the angular "stick" position directly correspond to the attitude; however this would present undue difficulty for all-attitude control. A more feasible solution appears if the control "stick" position is made to correspond to the desired *change* of attitude. To increase pitch, by +30° for example, the control "stick" is pulled back to that angle from the neutral position. The aircraft immediately responds with increasing pitch; however, as this occurs and the error between the present pitch and the intended pitch diminishes, an artificial force is placed upon the control stick, gradually returning it to its neutral position. This force is programmed, as some direct function of the error, so that a released "stick" displacement will result in asymptotic approach to the intended attitude in a most efficient manner.

The pilot must find it possible to place new commands on the aircraft prior to its completion of response to previous commands. Should the pilot desire an additional increment in pitch before the aircraft has completed the previously commanded +30° change, he can withhold the stick from its tendency to return to the neutral position. When this occurs, the artificial force which has been counterbalanced, is converted into a voltage command corresponding to a pitch rate. If the stick is halted at 5° displacement, then the corresponding increase might be, say, 5° per second. In this manner the pilot can effect separate control of pitch or pitch rate from the temporal pattern of stick displacement. Similar action may be accomplished with respect to roll. It is recognized that it might prove more efficient to increase the ratio between "stick" displacement and attitude displacement, but this must be left for a future experiment. It should prove easy for the pilot to learn the "feel" of stick position corresponding to his desired attitude command. Yaw control can simply be made correspondent to the angular rotation of control grip about its own vertical axis. It might prove to be of value to display the command attitude on slow responding vehicles, such as submarines, and in this way achieve a degree of anticipatory display.

The controls should also be compatible with the nature of the mission. An autopilot is provided to relieve the pilot during long time intervals of cruise control, but this device may prove to be of even greater benefit during critical mission phases such as strike, separation, or re-entry. Even if the human operator were immediately aware of the failure of his autopilot, it would take him too long to assimilate present status information and acquire adequate control. Alternative to this, it is proposed that the pilot be furnished with controls which are normally void. He is alerted for critical phases of the mission and required to "fly" with the aid of a complete set of displays. It should not be possible for him to perceive that he is not in control of the vehicle so long as the autopilot functions properly and errors remain small. If, however, the autopilot fails, he is automatically made a functional component in control of the system. Acquisition requires no delay and furnishes an optimal use of the human operator in a standby capacity.

The entire cockpit has an effect upon the pilot and should be designed so as to minimize the amount of perceived information which conveys no intelligence regarding aircraft status or the mission. A static surround allows human adaptation to obliterate new sensations. The degree of comfort corresponds to the sensed signal-to-noise ratio of the physical surround. The entire cockpit design should maintain a meaningful context and in this way minimize the contributed "noise."

Conclusion

The operational requirements of current high-performance aircraft often exceed the limitations of the pilot in a number of regards. The distinct advantage attained through use of the human operator rests with his ability to make decisions while in a position to receive and assimilate first-hand data and accomplish direct control. It is, therefore, imperative that information be provided to the pilot in such a manner as to take the fullest advantage of his decision-making capability. The large percentage of accidents which are attributable to pilot error, that is, 41 per cent in multi-engine jet and 45 per cent in multi-engine nonjet,[4] demands that some positive approach be taken to each of the individual problems which have been recognized as primary causal factors contributing to aircraft accidents.

"During the past two years marked change . . . has occurred and pilots' 'vertigo' has been implicated as a primary causal factor in 4 per cent of major accidents and 14 per cent of fatal accidents."[28] The rapid increase in jet inventory together with their increased performance may be expected to raise these figures even further. The cost of accidents in both human life and dollar value underscores the importance of this problem. It is therefore suggested that the kinalog mode of attitude display be fully explored in an effort to achieve an affirmative approach toward the prevention of vertigo.

Increased velocity and maneuverability can only be safely used through reference to anticipatory displays. It would prove most valuable to obtain the required data and formulation necessary for the design of optimal anticipatory displays under given conditions of vehicle spectral response and the individual spectra of signal and noise.

The complexity of cockpit instrumentation must be reduced. This may be accomplished through the inclusion of only relevant information, integrated in a most natural manner. The proposed modified pictorial "window" display instrument furnishes data in such a manner, utilizing the real world as the common denominator. Flight control decisions are optimized even under stressful situations. This primary instrument permits the incorporation of both kinalog and anticipatory display and should lead to successful flight control with a minimum of training.

Additional quantitative instruments are also described, each being designed to fulfill the informational requirements for a piloting problem. For example, the proposed altimeter provides a consistent pictorial-numeric display with compatible read-out of high precision data. All the flight control instruments of the set were conceived under the constraint that they be immediately feasible within the present state-of-the-art. Instrumentation concepts remain worthless until they are converted into flying hardware.

Present-day rocket vehicles are limited to a single type of mission. If these carry man into space, it is primarily as protected cargo. On the other hand, future vehicles will offer performance capability allowing a wide diversity of missions so that once again it becomes necessary to call upon the human operator to act as a flight control agent. It is only through an immediate program of intensive research that the fundamental problems can be overcome and the human operator permitted to act as that singular device which provides fullest benefit from his newly acquired freedom to fly.

[28] J. A. Nuttall and W. G. Sanford, "Spatial Disorientation in Operational Flying," Directorate of Flight Safety Research, USAF, Publication M-27-56, p. 24; September 12, 1956.

The Origin of the Professional Group on Medical Electronics*

L. H. MONTGOMERY†

Summary—The IRE Professional Group on Medical Electronics was organized in 1952. It was formed to provide a means of bringing together the relatively few electronic engineers who were interested in applying their knowledge to the problems of measurement and treatment in modern medical research and practice. The Group has grown from a few scattered individuals to an organization of some 2000, with local chapters in several of the large medical centers. As is the case with other IRE Professional Groups, it has its own publication and holds meetings for presentation of scientific papers. It also cooperates with other professional societies and organizations in the presentation of meetings and in the advancement of the professional status of its members.

During World War II, the OSRD was instrumental in bringing together a few of us who were pioneering in the field of medical electronics as a profession. Since these research contracts were terminated at the close of the war, those of us who continued our interest in this fascinating branch of engineering were left without a means of communication. This condition continued with our papers being published in all types of publications from *Review of Scientific Instruments* to the Sunday supplements. At the 1951 IRE National Convention, a symposium was held on dc amplifiers with Dr. Ernest Weber as chairman. Following the formal discussion of the panel, there was considerable discussion from the floor regarding the application of dc amplifiers to the recording of bioelectric potentials. At the close of the session, while talking with the IRE Executive Secretary, George Bailey, I suggested the organization of a Medical Electronics Group. He referred me to Larry Cumming, who furnished the details on setting up a Professional Group in the IRE. Upon my return to Nashville, I wrote to some fifteen of my acquaintances who had been associated with a medical electronics project in the past (see Fig. 1). There were one or two who doubted that there would be a sufficient number of engineers with medical interests, or doctors with sufficient engineering interest to form such a Professional Group, but the response of the others was quite enthusiastic. Cumming pointed out that for recognition it was necessary to have the petition signed by a minimum of twenty-five IRE members.

Practically all of the original fifteen responded with suggestions of several additional prospective members. However, since there were only one or two in each city, this eliminated the possibility of obtaining the necessary signatures at a single meeting. Since the largest number of prospects was at the Princeton, N. J. Laboratories of RCA, I naturally made my first attempt at organizing the group there. Dr. V. K. Zworykin was good enough to set up a luncheon meeting which proved very successful, and, in appreciation of his help and encouragement, Dr. Zworykin was given the honor of being the first to sign the petition. From Princeton, I went to Washington, D. C. and the Bureau of Standards, where several prominent physicists were added to our group of petitioners. With this success, I returned to Nashville and began circulating the petition by mail. The petition was submitted to the IRE about the first of February, 1952.

The call for papers had gone out (IRE PROCEEDINGS, September, 1951, p. 1100) for the 1952 National Convention and it was suggested that each Professional Group organize a technical session. Even though the PGME was not officially recognized as such, since the petition was still being circulated, William Doherty, the program chairman, agreed to give us a session provided we could produce a sufficient number of papers to justify it. This challenge was mailed to our list of tentative members, which by this time had grown to over a hundred. Again the response was enthusiastic, presenting a series of the most difficult decisions: we had enough excellent papers for at least two sessions but only time for one. It was with the deepest regret that I was forced to return well over half the papers. This then was our first technical session. I am sure that the quality of papers presented at these early meetings did much to establish the high standards of our Group. It may be interesting to review the list of titles and authors on the 1952 program, for most of the authors are still active in PGME work:

1) "New Electronic Techniques for Spectrophotometry," C. C. Yang, University of Pennsylvania, Philadelphia, Pa.
2) "Application of Microwaves in Physical Medicine," J. F. Herrick, Mayo Foundation, Rochester, Minn.
3) "Design Problems in the Absolute Oximeter," R. H. Taplin, Canadian Marconi Co., Montreal, Can.
4) "Television Microscopy in the Ultraviolet," V. K. Zworykin, L. E. Flory, and R. E. Shrader, RCA, Princeton Laboratory, Princeton, N. J.
5) "Recording Multi-Axial Projection of Vectorcardiograms," B. P. McKay, W. E. Romans, D. A. Brody, and R. C. Little, University of Tennessee, Knoxville, Tenn.
6) "Continuous Integrating Counting Rate System for Radioactivity," M. Berman and S. Vacirca, Sloan-Kettering Institute, New York, N. Y.

The approval of the petition was tentatively granted April 7, 1952, pending minor revisions to make it conform to the "Model Constitution." The Model Constitution for Professional Groups was being revised at the time.

* Original manuscript received by the IRE, September 1, 1959.
† Vanderbilt Medical School, Nashville, Tenn.

VANDERBILT UNIVERSITY
NASHVILLE 5, TENNESSEE

SCHOOL OF MEDICINE
DEPARTMENT OF ANATOMY

April 4, 1951

Dear Sir:

For sometime I have felt the need for an organization to bring together those of us who are working in the field of Medical Electronics. This was further confirmed by the type questions asked and discussed in the Instrumentation Symposium at the recent IRE Convention. It appears that there is quite a number of IRE members who are working in the field of Medical Electronics, and whose interests do not exactly fit one of the present professional group classifications.

It is my belief that with your assistance and the help of others who are interested in clinical or research work we can form a professional group on Medical Electonics which would be very useful to those interested in Instrumentation and Research associated with the medical profession.

Of course some one must take the initiative to form such a group. In this I must ask for your assistance. If you will send me the names of IRE members of your acquaintance whom you believe will be interested in signing the petition for the fromation of this group and perhaps serving on its initial administrative committee, I shall be glad to contact them as soon as possible.

Thanking you in advance for your cooperation in this matter, I remain

Sincerely yours,

L H Montgomery

L. H. Montgomery

Fig. 1—The first step in the formation of The Professional Group on Medical Electronics.

A meeting of the petitioners was called at the close of the technical session and the changes in the constitution were approved. The following officers were named to the Administrative Committee: L. H. Montgomery, chairman; W. C. Ballard, vice-chairman; J. F. Herrick, secretary-treasurer; the other members of the first Administrative Committee were Britton Chance, Harry Grundfest, and V. K. Zworykin. During 1952, the membership rose from 126 members to 539. Fifteen pages of *News Letters* were published, containing, among other items, a bibliography of papers published by the members of PGME in various journals (vol. 2, no. 1, February, 1953).

Although PGME could not officially take part in the program, it did aid the AIEE Group on Electrical Techniques in Medicine and Biology in procuring papers for the 1952 meeting. Arrangements were also made for full co-sponsorship of a joint conference in 1953 with S. R. Gilford as chairman. During the latter part of 1952, Wilson Greatbatch of the Buffalo-Niagara Section and A. J. Morris of the San Francisco Section applied for local PGME chapters. The race to be the first local chapter ended in a tie. Both were approved April 7, 1953.

When the time arrived for the 1953 IRE National Convention, Dr. Herrick organized the PGME session. It was a pleasure to act as chairman with such a group of speakers as those present in the Starlight Room of the Waldorf Astoria.

The program consisted of the following:

1) "Electric Photography," K. S. Lion, Massachusetts Institute of Technology, Cambridge, Mass.
2) "Concerning the Use of High Energy Particles and Quanta in the Determination of the Structure of Living Organisms," R. J. Moon, University of Chicago, Chicago, Ill.
3) "Possible Medical and Industrial Application of Linear Electron Accelerators," W. C. Barber, A. L. Eldredge, and E. L. Ginzton, Stanford University, Stanford, Calif.
4) "Capacity and Conductivity of Body Tissues at Ultra High Frequencies," H. P. Schwan and Kam Li, University of Pennsylvania, Philadelphia, Pa.
5) "The Problem of the Application of Electronics to Medicine," R. S. Schwab, General Hospital, Boston, Mass.
6) "Progress Report on Electronic Mapping of the Electrical Activity of the Heart," Stanford Goldman, D. W. Spence, M. Rizika, and S. Lidovitch, Syracuse University, Syracuse, N. Y.

A meeting of the PGME Administrative Committee was called during the convention and the following were elected for the 1953–1954 term: L. H. Montgomery, chairman; Britton Chance, vice-chairman; J. F. Herrick, secretary-treasurer, with S. R. Gilford, Wilson Greatbatch, and A. J. Morris as Administrative Committee members.

The revision of the Constitution was completed and approved on May 5, 1953. An assessment of one dollar was made to cover part of the publication cost of the *News Letter* and the CONVENTION RECORD. It was explained in the *News Letter* that the future assessments would be revised as our publication schedule required. The now familiar purple and white cover was designed and the standard format of the TRANSACTIONS was established in 1953. Dr. Herrick very kindly accepted the position of Editor for the PGME TRANSACTIONS, and has made it the outstanding publication in the medical electronics field. The membership steadily increased and passed the 1000 mark during the latter part of 1953.

As previously planned, the PGME joined with the AIEE and the ISA in sponsoring a conference on "Electronic Instrumentation and Nucleonics in Medicine" at the New Yorker Hotel, November 19–20, 1953. This was its first PGME program with another professional society. The first session was on diagnostic devices and the papers were:

1) "Multi-Channel Electromyography Instrumentation and Application," J. F. Davis, McGill University, Montreal, Can.
2) "A Stimulus Monitor and Its Use in Electrophysiology," H. C. Becker and S. Peacock, Jr., Tulane University Medical School, New Orleans, La.
3) "Application of Positron Emitt ng Isotopes to the Location of Brain Tumors," G. L. Brownell, Massachusetts General Hospital, Boston, Mass.
4) "The Application of Ultra-Sonic Mechanical Waves to the Visualization of Normal Soft Tissue Structure and Disease Processes Including Cancer," D. H. Mowry, V. A. Hospital, Ft. Logan, Colo.

The theme of the afternoon session dealt with X ray and these were the papers:

5) "Isodose Plotting Instrument for X-Rays," John S. Laughlin, Memorial Center, New York, N. Y.
6) "The X-Ray Microscope," S. P. Newberry, General Electric Co.
7) "Instrumentaion Aspects of Single Plane and Stereoscopic X-Ray Motion Pictures," S. A. Weinberg, University of Rochester Medical School, Rochester, N. Y.
8) "Cinefluorography Using Image Brightness Intensification," W. S. Lusby and E. M. Hayes, Westinghouse Electric Corporation.

The following morning session was devoted to laboratory measurements:

9) "A Method for the Exact Determination of Volume Concentration of Nonconduction Particles in Conducting Solvent," H. P. Schwan and T. P. Bothwell, Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.
10) "An Instrument for Rapid Dependable Determination of Freezing Point Depression," R. L. Bowman, National Institute of Health, Bethesda, Md.
11) "Blood Serum Analysis by a Quantitative Paper Electrophoresis Apparatus," E. L. Durrum, Army Medical Center, and S. R. Gilford, National Bureau of Standards, Washington, D. C.
12) "An Electrical Method for Determining Action Spectra of CO_2-Inhibited Respiration," L. R. Castor and B. Chance, Johnson Foundation for Medical Physics, Philadelphia, Pa.
13) "An Infrared Analyzer for Continuous Respiratory CO_2 Analysis," M. D. Liston, Liston Becker Co., Stamford, Conn.

The development of standard diagnostic and therapeutic instruments in our leading hospitals has resulted from many of these research reports.

The Connecticut Valley Section applied for a local chapter and, through the activities of J. T. Filgate, was granted a charter on March 2, 1954.

During the 1954 National Convention, the PGME sponsored, in addition to an excellent technical session, "The Spotlight Session" of the Convention. It was a symposium on "Engineering Based on Biological Design." The moderator was Dr. W. R. G. Baker and the panel consisted of Dr. Leonard C. Mead, Department of Psychology, Tufts College, Medford, Mass., who spoke on "Human Engineering"; Dr. Norbert Weiner, Department of Mathematics at M.I.T., Cambridge, Mass., who spoke on "Information Theory"; Dr. S. S. Stevens of Harvard University, Cambridge, Mass., whose topic was "Biological Transducers"; and Dr. Otto H. Schmitt of the University of Minnesota, Minneapolis, who treated "Biological Servomechanisms and Control Circuitry." The symposium was presented in the main ballroom of the Waldorf-Astoria Hotel, New York, N. Y., to a standing-room-only audience. The technical session consisted of the following papers:

1) "Visualization of the Distribution of Gamma Emitters In-Vivo by Means of the Gamma Ray Pinhole Camera and Image Amplifier," R. K. Mortimer, H. O. Anger, and C. A. Tobias, University of California, Berkeley.
2) "Expansion Chamber for Measurement of Red Cell Permeation by Water," A. K. Solomon and C. V. Paganelli, Harvard Medical School, Boston, Mass.
3) "Color and Enhanced Contrast X-Ray Images," R. S. Mackay, University of California, Berkeley and San Francisco.
4) "Measurements of Slow Neutron Depth Doses," E. Stickley, Brookhaven National Laboratory, Upton, N. Y.
5) "Use of Charged Particles to Measure Skin Thickness and Other Surface Properties," F. Hutchison, Yale University, New Haven, Conn.

At the meeting of the Administrative Committee, the following were elected to serve during 1954–1955: J. F. Herrick, chairman, and Otto Schmitt, vice-chairman; Wilson Greatbatch was appointed secretary-treasurer. Lee B. Lusted was elected to the Administrative Committee with myself and A. J. Morris. It was suggested that we increase the membership on the Administrative Committee but the committee decided to defer this move until 1955. C. P. Hedges organized a group in Los Angeles for a local chapter, and Mathew Conrad, with Herman Schwan, started working on a local chapter in Philadelphia, which was granted a charter about a year later.

The joint conference of the PGME, the AIEE, and the ISA for 1954 was held in Chicago on November 10, 11, and 12. I acted as program chairman with the able assistance of E. Dale Trout. It was through his efforts that two very interesting field trips were arranged to the Argonne Cancer Research Hospital and the Argonne National Laboratory.

November 10, Morning, Session I, "Circulation and Cardiology."

1) "Normalization of Vector EKG Axes," O. Schmitt, University of Minnesota, Minneapolis, Minn.
2) "Computer Technics in Rapid Determination of Certain EKG Functions," S. Briller, Bellevue Hospital, New York, N. Y.
3) "Synthesis of Electrocardiograph Leads," R. McFee, Bell Telephone Laboratories, Whippany, N. J.; F. Johnson, University of Michigan Medical School, Ann Arbor, Mich.
4) "Determination of the Electrical Center of Ventricular Depolarization in the Human Heart," E. Frank, Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.
5) "Heart Cell Potential," E. E. Suckling, New York State University.

November 10, Afternoon, Session II, "Electrical Properties of Biological Materials."

1) "The Electrical Impedance of a Human Body Segment as Related to Impedance Plethysomography," P. Albro, Hastings Instrument Company.
2) "Electro-Ionics of Nerve Action," K. S. Cole, Naval Medical Research Institute, Bethesda, Md.
3) "The Application of the Electric Impedance Method in the Study of the Normal and Hemolyzed Red Blood Cell," H. Fricke, Walter B. James Laboratory for Biophysics, Cold Spring Harbor, N. Y.
4) "Capacity of Erythrocyte Membrane Measured over a Wide Frequency Range," Herman P. Schwan, E. L. Carstensen, and K. Li, Moore School of Electrical Engineering, University of Pennsylvania.

Evening Guest Speaker: T. E. Allibone, Associated Electrical Industries of Great Britain.

1) "Spectral Phonocardiography," G. N. Webb, Johns Hopkins Hospital, Baltimore, Md.
2) "New Developments in the Electronic Fluoroscope of the Television Type," R. Sturm and R. Morgan, Johns Hopkins Hospital, Baltimore, Md.
3) "Effect of X-Rays on Erythrocytes," O. Bluh, University of British Columbia, Vancouver, Can.
4) "Xerography," J. Roach, Albany Medical School, Albany, N. Y.
5) "Progress in X-Ray Movies," E. L. Webb, Westinghouse Electric Corporation, Baltimore, Md.

At the meeting of the Administrative Committee held on March 24, 1955, Dr. Herrick, the chairman, complimented the local chapters on the programs they had held and stated that they were doing a notable job of bringing together the engineers and doctors on a local level. The Buffalo-Niagara, San Francisco, Connecticut Valley, and Los Angeles Chapters were all represented and reported close liaison with the local medical societies.

Dr. Zworykin was elected chairman and Lee Lusted was chosen vice-chairman. Wilson Greatbatch was appointed secretary-treasurer, and additional members were elected to the Administrative Committee to expand it to fifteen members as planned previously. Dr. Zworykin announced the plans for a Medical Electronics Center at the Rockefeller Institute for Medical Research, its purpose being the promotion of cooperation between the medical and engineering professions. The first conference sponsored by this organization took place on June 20, 1955, and was on the subject of electroencephalography. There have been many other subjects covered in the intervening years. These meetings have been of inestimable value to the participants by providing them with the proper environment for the free exchange of ideas.

In retrospect, I am deeply grateful for the help which was so freely given by those who aided in the formation of the Professional Group. It was largely through their efforts that we have been able to attain our goal which, as originally defined, was "to aid its members in the application of electronic engineering techniques to the problems of the medical profession and to provide a forum for the presentation of significant developments in this field." The revised constitution points out these objectives quite clearly in Article I, Sections 2 and 3.

"Its objects shall be scientific, literary, and educational in character. The Group shall strive for the advancement of the theory and practice of radio engineering and of the allied arts and sciences, and the maintenance of a high professional standing among its members, all in consonance with the Constitution and by-laws of the IRE and with special attention to such aims within the field of interest of the Group as are hereinafter defined."

"The Group shall aid in promoting close cooperation and exchange of technical information among its members and to this end shall hold meetings for the presentation of papers and their discussion, and through its committees shall study and provide for the needs of its members."

Article III, Section 1: "The fields of interest of the Group is the study of biological and medical systems. The Group will provide a forum for the presentation of research and development problems in biology and medicine which might be aided in solution by use of electronic engineering principles and devices, and conversely, the presentation of new developments in electronic engineering which might find wide, or special, application to biological and medical research. The Group will provide means for the personal exchange of information in this area of medical and biological research and for establishing rapport between the workers in these fields."

Since 1955 is still within the memory of even the most mildly interested, I believe this would be a good place to stop and leave the later years to some future author. I have been very happy to be associated with such a group as the founders of the PGME and hope that its members will continue to seek to apply knowledge of electronic engineering to the advancement of science and the relief of human suffering. This is undoubtedly one of the most worthy humanitarian objectives. It is a field of tremendous potentialities and should call forth our most enthusiastic efforts.

Appendix

PROFESSIONAL GROUP ON MEDICAL ELECTRONICS

PGME Officers and Administrative Committee Members (1952–1959)

1951 Organizer			
L. H. Montgomery			
1952			
Chairman	—L. H. Montgomery	B. Chance	
Vice-Chairman	—C. M. Ballard	J. F. Herrick	
Secretary-Treasurer	—H. Grundfest	V. K. Zworykin	
1953			
Chairman	—L. H. Montgomery	S. R. Gilford	
Vice-Chairman	—B. Chance	W. Greatbatch	
Secretary-Treasurer	—J. F. Herrick	A. J. Morris	
1954			
Chairman	—J. F. Herrick	L. B. Lusted	
Vice-Chairman	—O. H. Schmitt	L. H. Montgomery	
Secretary-Treasurer	—W. Greatbatch	A. J. Morris	
1955			
Chairman	—V. K. Zworykin	L. E. Flory	A. J. Morris
Vice-Chairman	—L. B. Lusted	J. F. Herrick	O. H. Schmitt
Secretary-Treasurer	—W. Greatbatch	J. K. Hilliard	H. Schwan
		U. Liddel	N. Smith
		L. H. Montgomery	A. R. Tunturi
1956			
Chairman	—V. K. Zworykin	C. Berkley	U. Liddel
Vice-Chairman	—L. B. Lusted	M. Conrad	L. H. Montgomery
Secretary-Treasurer	—W. Greatbatch	L. E. Flory	O. H. Schmitt
		J. F. Herrick	H. Schwan
Advisory Member	—O. Glasser	J. K. Hilliard	N. Smith
		E. Levinthal	W. E. Tolles
1957			
Chairman	—L. B. Lusted	C. Berkley	E. C. Levinthal
Vice-Chairman	—U. Liddel	M. Conrad	L. H. Montgomery
Secretary-Treasurer	—W. E. Tolles	L. E. Flory	J. W. Moore
		W. Greatbatch	O. H. Schmitt
Advisory Members	—G. E. Burch	J. F. Herrick	H. Schwan
	L. M. Hellman	J. K. Hilliard	V. K. Zworykin
	C. F. Kay		
	J. P. Sampson		
	D. Stetten		
	S. F. Thomas		

PGME Officers and Administrative Committee Members (*cont'd*)

1958

Chairman	—U. Liddel	M. Conrad	L. B. Lusted
Vice-Chairman	—W. E. Tolles	L. B. Flory	J. W. Moore
Secretary-Treasurer	—C. Berkley	W. Greatbatch	O. H. Schmitt
		J. K. Hilliard	H. H. Zinsser
Advisory Members	—G. E. Burch		V. K. Zworykin
	D. Holaday		
	O. Paul		
	J. P. Sampson		
	D. Stetten		

1959

Chairman	—W. E. Tolles	R. Bowman	E. P. MacNichol
Vice-Chairman	—L. E. Flory	S. Gilford	J. W. Moore
Secretary-Treasurer	—G. N. Webb	J. Hervey	O. H. Schmitt
		Urner Liddel	H. H. Zinsser
Advisory Members	—G. E. Burch	L. B. Lusted	V. K. Zworykin
	D. Holaday		
	O. Paul		
	J. P. Sampson		
	D. Stetten		

PGME Membership (1952–1959)

Year	*Total Membership*	*Affiliates*
1952	539	
1953	1000	
1954	1181	
1955	1389	
1956	1542	
1957	1661	47
1958	1711	104
1959 (July)	1834	143

Societies Approved for Affiliation with the PGME

Names of Societies	Date Approved By IRE Executive and PGE Administrative Committees
American Association for the Advancement of Science	February 6, 1957
American Association of Clinical Chemists	February 6, 1957
American Chemical Society	February 6, 1957
American Federation for Clinical Research	February 6, 1957
American Medical Association (including all state & county medical societies)	February 6, 1957
American Nuclear Society	February 6, 1957
American Physical Society	February 6, 1957
American Physiological Society	February 6, 1957
American Society for Artificial Internal Organs	February 6, 1957
American Society of Biological Chemists	February 6, 1957
American Society for Clinical Investigation	February 6, 1957
American Society of Clinical Pathologists	February 6, 1957
Electrochemical Society	February 6, 1957
Electron Microscope Society of America	February 6, 1957
Institute of the Aeronautical Sciences	February 6, 1957
New York Academy of Sciences	February 6, 1957
Society for Experimental Biology and Medicine	February 6, 1957
American Institute of Biological Sciences	May 14, 1957
American Institute of Physics and its member societies:	May 14, 1957
Acoustical Society of America	
American Association of Physics Teachers	
American Physical Society	
Optical Society of America	
Society of Rheology	
Federation of American Societies for Experimental Biology and its member societies:	May 14, 1957
American Physiological Society	
American Society of Biological Chemists	
American Society for Pharmacology & Experimental Therapeutics	
American Society for Experimental Pathology	
American Institute of Nutrition	
American Association of Immunologists	

Societies Approved for Affiliation with the PGME (*cont'd*)

Names of Societies	Date Approved By IRE Executive and PGE Administrative Committees
The Biophysical Society	June 21, 1957
American Society of Mechanical Engineers	January 6, 1958
Instrument Society of America	April 7, 1958
American Dental Association	June 26, 1958
Danish Biological Society	December 8, 1958
Svenska Teknologforeningen	December 8, 1958
American Psychological Association	February 10, 1959
Aero Space Medical Association	August 17, 1959
American Academy of Orthopedic Surgeons	August 17, 1959
American Astronautical Society	August 17, 1959
Human Factors Society	August 17, 1959
Institute of Physics of Great Britain	August 17, 1959
Verband Deutscher Physikalischer Gesellschaften	August 17, 1959

PGME Chapters

Chapter	*Date Approved*	*Organizer*	*Chairman*
Buffalo-Niagara	April 7, 1953	W. Greatbatch	W. Greatbatch
San Francisco	April 7, 1953	A. J. Morris	G. K. Turner
Connecticut Valley	March 2, 1954 (inactivated, November, 1956)	J. T. Filgate	
Los Angeles	November 9, 1954	C. P. Hedges	M. C. Biedebach
Philadelphia	June 7, 1955	M. Conrad	J. M. Reid
Washington, D. C.,	June 7, 1955	J. M. Carter	F. W. Noble
N. Y., Northern N. J., L. I., and Princeton	March 1, 1956	C. Berkley	L. E. Flory
Montreal	April 10, 1956	C. Pinsky	J. T. Davis
Boston	September 25, 1956	S. Aronow	A. M. Grass
Chicago	December 12, 1956	L. Leopold	T. Fields
Omaha-Lincoln	April 7, 1958	H. G. Beenken	J. N. Barmore
San Diego	April 7, 1958	G. J. Mealey	G. J. Mealey
Houston	June 26, 1958	H. E. Childers	H. E. Childers
Columbus	November 17, 1958	B. G. Austen	B. G. Austen
Baltimore	March 30, 1959	G. N. Webb	G. N. Webb

Instrumentation in Biomedical Research*

PAUL E. KLOPSTEG†

***Summary*—This paper is, in effect, an essay looking toward the improvement of instruments used in biomedical research, their availability, and the knowledge and skills necessary to use them to the best advantage. The discussion suggests that it might be desirable to establish a center or centers where teaching, research, and engineering can be done on instruments; where scientists or technicians can learn to use them; and where instruments can be rented as needed, with or without technical assistance. One will not find in this article the names of biomedical instruments or quantitative information about their current use. Recommendations are made on the following subjects: information clearing house, instrument and technician pool, instruction, professional status, student recruitment, and hierarchy of facilities.**

* Reprint of a survey prepared for the Biology Council, Div. of Biology and Agriculture, Natl. Acad. Sciences, Natl. Res. Council, Washington 25, D. C. Additional reprints are available only from the Natl. Acad. Sciences, 2101 Constitution Ave., N. W., Washington 25, D. C., at $0.75 per copy. Please specify Publ. 472.

† Consultant to the Biology Council, Division of Biology and Agriculture, National Academy of Sciences, Washington 25, D. C.

Introduction

RESEARCH in biology and medicine, no less than in other areas of science, is critically dependent on experimental methods and devices providing for observation, measurement and control, and the processing and reduction of data. "Instrumentation," for the purposes of this report, may be used as a catchword for a very broad effort—essentially that of capitalizing on the methodology and conceptual approach of the more exact sciences (mathematics, physics and chemistry), including instrument production and distribution, and procedures for testing. The primary aim of this survey is to see how instrumentation, as thus defined, and "instrumentology," the study and application of physical principles and mathematical operations to designs for the experimental study of specific problems, might logically be expanded and adapted to the service of biological research. A secondary consideration is the enlistment into biomedicine of those who are already trained in physics and engineering.

Biology is a fragmented science; its subdivisions are dependent on the diverse experimental techniques and devices of the physical sciences to varying degree. It is peopled by a group whose training is neither in engineering nor in primary physical sciences. It seems, therefore, best to dissect the problem, sorting out its several aspects, and fixing eventually on a few basic tenets. Quite likely, in view of the complicated nature of the question, its solution will be approached not by a single route but by several, each adapted to specific, recognizable situations. There now exist units at universities, colleges and other research institutions where effective work in devising tools for biology has been, and is being, done. Their success lends incentive to the search for a larger answer to the question posed and strengthens the belief that substantial solutions to the problems will be found.

Four classes of biologists are involved: 1) the scientist who seizes upon a particular technique, developing first-rate competence in it, and utilizing it in a very primary sense; 2) the scientist who utilizes a given technique at a less sophisticated level, or but temporarily, and who might be considered to have only secondary competence in it; 3) the individual who turns the instrumentation segment of his research program over to another; and 4) the individual who foregoes the use of instruments altogether. Any survey, to be effective, must be comprehensive enough to include the more efficient use of the instruments we now have, the development of better instruments, the training of biologists in more skilled and resourceful use of instruments, and the establishment for the biological technician and technical specialist of a respected and adequately rewarded station in the scientific community.

The Need

As use and complexity of instruments in biomedical research increase, and as research becomes increasingly a team operation, our present highly informal procedures become less and less appropriate. Such unsystematic practices are both ineffective and inefficient, resulting in serious losses of time and money, in distortion of research patterns, in use of instruments ill-adapted to a particular project, in failure to realize more than a fraction of the total potential of modern developments in instrument science. Despite the fact that there are a number of persons in the biological community preeminently fitted by temperament and aptitude to become expert specialists in the use of instruments in support of research, and despite the enormously increased demand for this type of individual, it remains a most difficult task to establish for these workers a recognized station of respectability, with the deserved professional status.

There is all too little evidence that biologists as a whole feel, with genuine urgency, that something needs to be done. Many are quite happy with very minimum standards, which leads them to attempt all sorts of short-cut solutions to their instrumentation needs. Others, dissatisfied in varying degree with matters as they are, still fear moves which they view as attempting to replace traditional university training and as emphasizing method over problem. One of the grave handicaps encountered by Dr. Klopsteg in his travels about the country, collecting on-the-spot opinions, was the wide difference in viewpoint, especially in relation to the status and role of biophysics. An attempt to reconcile fully these opposing views is very likely to fail—it is better strategy to state them as they are and leave any decision to the individual reader. But in reaching such a decision, he must keep in mind that among those interviewed must have been some who were too single-tracked to step outside their specialized channel of interest and, furthermore, that even though immediate needs and the demands of the future may differ, any practical measures must be designed to fit both.

Divergence in view among participants in the survey was largely on the question how best the desired ends might be obtained—majority opinion was decidedly convinced of the urgency of the situation and the desirability of doing something tangible to alleviate it. It was pretty well agreed that those who do recognize the need and who pioneer in planning for the future must risk being wrong and move ahead on something which will alert others to the realities of the problem.

Concerning Existing Facilities and Equipment

A most obvious specific need is to make available, to the maximum degree, that equipment which is now in existence, whether it be rare and expensive items or conventional instruments, perhaps through the establishment of rental pools and a clearing house through which items only temporarily needed at a given institution could be relocated. Such services might be extended to provide, also, specialized technical personnel for the operation of equipment on loan.

There is need for extensive, continuing outside assistance to the facilities now in being—those grassroots elements now doing a large part of the work in biomedical instrumentation and destined in the future to act also as feeders to any more comprehensive center. Very substantial support could be given existing facilities, at all levels, before there is appreciable danger of going beyond the point of optimum returns.

Improvement of Techniques and Instruments

Basic to any consideration of biomedical instrumentation is the need for general improvement in the machines and their use. This recognizes, among other things, need for types of instruments not now available and knowledge of how to design instruments to fill these needs; some way must be found for getting top skills concentrated on development of new instruments. An instrument, once developed by the physicist for his own purposes, must ordinarily be engineered to a more dependable design before it can readily be adopted by other physicists or nonphysicists in their research.

It is inherent in the nature of biological research that needed skills and techniques cannot be transferred directly from physics and chemistry. A given instrument is useful only in the context of the particular problem, and most are not designed or conceived of with biological problems in view. There are perhaps impressive numbers of instruments around for which the techniques or applications to biology, however potentially important, are not obvious. Among the significant needs, then, are those ***not*** recognized—the realization that nature can be probed effectively by certain physical methods. It is this ability to divine a need in biology or medicine in terms of instruments that must be provided in the education of this new group we sometimes call biophysicists.

Training and Personnel Utilization

There are at least three aspects to the problem of efficient utilization of available and potential manpower as it relates to instrumentation in biology and medicine: 1) the need to inform or train senior scientists concerning particular instrumentation problems, with minimum expenditure of their time; 2) the desirability of reducing the interruptions now inflicted on research biologists by others seeking to "borrow" their skills and advice; and 3) the opportunity to replace expensive technicians at many points with less costly automation. There is a time element to consider: when a given machine is new and rare, one must be most concerned with the efficient use of equipment; when abundant, with the efficient use of manpower. In the latter case, the solution may call for the establishment of technician pools and for groups equipped to do specific instrumentation jobs on a consultant and contract basis. Finally, there is a need for the erection of formal courses and curricular sequences in educational institutions to train new personnel at various levels of competence and for various kinds of work in instrumentation.

It is not sufficient to supply biology and medicine with a given number of qualified persons; the diversity of personnel needs must be recognized and dealt with. First, there must be technicians, sufficiently skilled to keep standard equipment in running order, service it, and operate it. Secondly, there must be instrumentation specialists familiar with the more sophisticated instruments, their theory, construction and methodology, and capable of adapting their use to specific biomedical problems, including improvements of design and operation. These would form a new class of "biological engineers." Finally, there must be the creative instrumentologists, thoroughly acquainted with the principles of the physical sciences and capable of inventing new and better tools for the study of organic systems.

Professional Status and Recruiting

To speak of the service to biology rendered by those skilled in the research methods of the physical sciences is not to imply for them a subordinate position in the scientific community.

Quite the contrary, ways must be found to guarantee proper professional status for the instrument specialist, if the program is to succeed. A model for the relationship which should exist between biomedical investigator and instrumentation specialist may be seen in that between writer and librarian. It should not be necessary, as it now is, to make a case for those in responsible positions in instrumentation facilities by circuitous talk or to employ devious stratagems to secure for them academic or staff appointments.

Biologists deplore the fact that they do not get what they feel is their share of competent persons coming into science. Despite this shortage, and despite the very considerable body of individuals potentially fitted for these very jobs, research organizations use highly unsystematic procedures in securing and selecting their technicians and instrument personnel. If there is one paramount need, therefore, it is to supply biomedical laboratories with manpower having the requisite training and ability.

To accomplish this, biomedicine must keep its recruiting program as active as those of its competitors, thus patterning its appeal to reach all levels of capability in order to enlist the aid of those who are primarily concerned with teaching and training. There must be emphasis on channeling promising students toward the methods of physical science.

Information Exchange

In the final area of need it is imperative: 1) that information on the present fund of ideas and experience be disseminated and that availability of instruments be made known; 2) that emerging needs and inadequacies be made known to those who develop and manufacture equipment; 3) that some device for informal exchange of opinion be set up; and 4) that some way of cataloging all available equipment and skills be devised. Preparation of an authoritative handbook of principles and methods would be a desirable dividend from an active instrumentation center.

Basic Considerations

In the foregoing paragraphs, we have attempted to spell out, without extensive comment, recognizable needs in biomedical instrumentation. In addition, there are a number of issues which have been subject to differing interpretations and, often, opposing points of view; these must be taken into account before a valid set of conclusions can be formulated. It was to such issues that the conferees devoted a large proportion of their time and effort; space limitations admit only of essentially undocumented summmaries here. No special significance attaches to the order in which the following items are listed.

The Individual and the Team

Although it should always be possible, particularly at the conceptual level, for a scientist to pursue his research successfully as an individual, with such help or advice as he may seek, there seems little doubt but that, in most instances, the attack on a large and difficult problem by a cooperating group of scientists is likely to be more effective. In resolving the question of the proper relation between technician and investigator and between investigator and instrument, however, some argue that new techniques characteristically come from individual laboratories and that in a separation of research scientist and technician there lies the danger that the service worker will become sterile. Others feel that the obligation of the competent investigator to attend to every technical detail himself has been unduly stressed, though they at the same time disavow any intention of supplanting the human brain with machines.

Informal Assistance and Interference

The individual possessing an instrument is in some measure custodian of a resource, particularly at the university level, and in that measure obligated to make the machine accessible, to inform his colleagues of its value and limitations, and to educate others as broadly as his work will tolerate. But it seems evident that, if instrumentation is to expand, this kind of service must not be left indefinitely to the favors and good graces of willing colleagues—the demand for voluntary courtesies would before long reach such an absurd volume that the system would break down completely. It is just because the established research laboratory and university department must train the next generation for the long haul that they cannot tolerate continual interruptions by outside visitors.

Dispersal vs Centralization; Diversity

One must attempt to balance the relative merits of large and small facilities, of dispersed and centralized patterns of installations, judging which is the more likely to survive and prosper, and recognizing that diversity may in the end be the wiser choice.

Wholly aside from financial considerations, treated elsewhere, the concept of a single large national center has been criticized by some as too ambitious, too risky, too difficult to assemble and staff, and, by its nature, remote in space and spirit from the preponderant majority of biologists who might hope to profit from it. Those who doubt the wisdom of the comprehensive facility fail to see how all the many ramifications of biomedical instrumentation could be encompassed in one establishment without at the same time forming an unworkably enormous organization.

Equally competent scientists, of the opposite persuasion, feel that the local university instrumentation unit already exists on virtually every campus that could support one, and that means simply do not exist at the remaining institutions to undertake such a venture at "threshold" size. They feel that, in putting the facility in special reference to activities on the campus, unique problems arise as the program becomes parochial, ingrown, noncompetitive, as it encounters established customs and seems to threaten to upset existing organizational and fiscal patterns. In essence, they are convinced that if a center is to be large enough to be dramatically new and different it will be too large for university sponsorship; it it is small enough to be part of the university it will not achieve the comprehensiveness that many concerned with biomedical instrumentation feel must be reached.

The answer, developed later in this report, seems to be a hierarchy of facilities from the individual to the regional (perhaps eventually national) center, with maximum emphasis on cooperation and coordination among them.

Educational Implications

From the outset, the educational aspect of instrumentation was recognized to include instruction of professional research personnel, of technicians, and of a new class of technical specialists or biological engineers. It will range all the way from formal courses in established methods to subtle changes in the general education of biologists, such that they are in tune with the increasing emphasis on instrumentation in research and fully able to capitalize on its contributions. Involved, somewhere, is the relation of research biologists to instrument operation, selection and development, and the balance between substituting technicians for scientists, on the one hand, and liberating scientists for creative activities, on the other.

The university must be urged to set up needed courses, to employ and recognize the personnel required to handle instruction, and to see that the teaching is as carefully organized as are the more orthodox items on the curriculum. There cannot be a course for everyone and for every need—some workable compromise is called for. And there will always be problems which cannot await university courses. Here the scientist must work out his own salvation; the more competent he is, the less will he require outside help.

Grant Policies and Fellowships

The general trend in fund-granting policies is toward greater leniency, toward allowing costs of visits to other laboratories for instructional purposes, and toward permitting procurement of instruments to be included on the research grant application budget. It will probaby continue to be difficult for an unproven investigator to get the more expensive items in this way but, as appropriations become more liberal, inclusion of funds for equipment on grant applications will have less and less effect on their final disposition. The parent institution, normally the actual recipient of the funds, should be encouraged to assume responsibility for the adequate use of the instruments purchased therewith.

Each new crop of fellowship recipients, by their almost random migration to new posts, creates a valuable cross-fertilization. There are perhaps two shortcomings in the fellowship programs as a means to acquire *ad hoc* instrumentation skills: 1) they do not, except in a few cases, provide for the mature investigator; and 2) tenure is ordinarily for one year, which is for the older scientist too long a time to spend in learning a new technique. Perhaps greater over-all flexibility would be desirable.

Continuity of Support and Staff

There can be no doubt that the success of any plan will hinge upon the enthusiasm of the devoted individuals comprising the staff, inspired to work not just as service personnel, who will see that the job gets done. It is this devotion that will permit adjustment to the host of intangible factors certain to plague the enterprise, and will insure continuity during the early, critical years. Whatever the source of funds, the granting agencies *must* assume responsibility for continuity. The new venture should be carried on for enough years to rid it of operational defects, after which it may be expected to continue on its own momentum, with support from a variety of sources.

Personnel Problems

Planning cannot be in a vacuum; it must deal with practicalities. The purely technical problems cannot be separated from those of human relations which, in the minds of many, constitute the one biggest difficulty confronting the instrumentation program.

Personnel relations involve not only the interplay between biologist and physical scientist, but among all of the diverse individuals—scientists, instrument makers, designers of equipment and techniques, students—who will be associated in an instrumentation facility. To be assured of success, a plan must be suited equally to the conformist and to the individualist. As the venture depends for its successful outcome on cooperation, its organization and operation must be so skillfully drawn that points of friction are foreseen and avoided, that causes for grievance or complaint are minimized. Perhaps the greatest factor in the success of a center will be the adroit management of a director skilled in the appraisal of character as well as ability, and wise in the leadership of a diverse group.

Nothing so clearly points up the issues at stake as the controversy, to put it bluntly, over the status of biophysics. This group is trying to win recognition as a separate discipline and its members are understandably concerned about any move possibly reducing their role to one of subservience to biomedicine. Psychologically, it would appear easier to develop the requisite cooperation between a biologist and a man who regards himself as appointed to the instrument field than between the biologist and the man who is primarily a physicist or chemist. This points up the possible advantages of beginning with the formation of a new institution, or of attaching the center to an organization already sympathetic to this sort of thing; indeed, the National Laboratories may be, uniquely, the place where the idea would be accepted. Finally, what of the students who attend? They will fairly begin to ask where they may go, without becoming second-class citizens, as they move out into a new job.

Communications and Publication

One deterrent to progress in biomedical instrumentation has been lack of adequate intercommunication among those concerned. Instances may be cited where several laboratories, recognizing the same problem, have independently undertaken to develop the apparatus with which to handle it, each without knowledge of what the others were doing. In other cases, research workers have set about devising an instrument which, in a well-developed design, has already been commercially available. Hence, an information clearing house is indicated, which might well make available current technical information about commercial instruments, apparatus and other research aids. It might serve as the means for channeling, in both directions, information about needs and requirements, and about ways of meeting them, between users and producers of scientific instruments.

Assuming efficient dissemination of information on established instruments and methods, and close cooperation on development of new techniques, it is still necessary to get those interested in instrumentation together for informal interchange of opinion on new techniques, data reduction, application of statistics, applications to biology, etc. Such informal gatherings must be supplemented by more systematized occupational guidance and instruction. It could be that the informal gathering is a device having special promise for instrumentation. If gathered about instruments and methods, irrespective of discipline, people might cooperate in a very productive way. The avowed theme ought always to be centered on problems in instrumentation, in preference to research results, *per se*.

Viewed in one light, there are now plenty of journals in which the results of research of biophysical nature can be published. Nevertheless, its dispersal through a multiplicity of journals renders retrieval of the information more difficult and tends to retard professional recognition. The literature of science is weak in critical reviews of instrumentation; journals are not happy to receive papers on method unless in conjunction with research results. Reviews by competent men on instruments, showing the lacunae of understanding, the limitations, range of use, etc., would be a positive step which could be fairly easily instituted. There is no shortage of published material, but rather of organized information on methods and techniques, articles which tie history with the present and include unsolved problems.

Service Functions

It is doubtful if solutions worked out for the large universities and research centers can be translated directly to the innumerable small institutions, where good research could go on and which are a natural resource of manpower for the biomedical field. To what extent can commercial laboratories be expected to satisfy the more humble needs of these institutions? It is no effort to find examples of such service laboratories, profit and nonprofit, now in being, which provide for "outpatient" tests and measurements, equipment rental, or external contract operations.

At the moment, rental is customary only on larger and more expensive items, or in cases where the manufacturer is trying to introduce a new machine. There could be a central service for routine instruments if the demand were great enough. What would be needed is not so much a physical pool as a pooling system or procedure. But just because equipment is thus made available is no assurance that people will know how to use it. Rental and training make a neat package, and one suspects that the technicians and the instrument will in the long run go together.

Many hold certain reservations about widespread adoption of rental; for the less expensive items, it may only occasionally be a long-range economy, and both instrument dealer and manufacturer can be expected to worry about the effect on sales volume. Speed, in any case, and accessibility to the worker in the poorly financed laboratory are two telling arguments in favor of a rental system. And in a situation where so very large a proportion of research is supported by grants in aid, the policies of granting agencies must be sympathetic to inclusion of rental items on the project budget estimates.

Rental would permit exploratory use of an instrument, the utility of which had not yet been ascertained. If successful, it would then often be purchased. At present, it has to remain untried for the most part.

Conclusions and Recommendations

Implicit in the results of the survey, and in the several conference discussions of those results, are the conclusions that: 1) there is need for action in the area of biomedical instrumentation—that matters cannot be left to *laissez-faire* treatment; 2) even so, there is no royal road or master formula; one can but lay down general guiding principles; 3) it is probably best to start one, or perhaps a few, centers under a variety of formulas; and 4) the movement will lead to a new class of "biological engineers" at the professional level, requiring adaptations in the educational program and attention to their acceptance and status. As it now stands, the survey gives only limited support in the immediate future to the big national center envisoned at the outset.

In seeking a solution, one may take either the perfectionists' approach, waiting until everything is optimum, or that of the individual who starts with courage along a chosen line, subject to evolutionary improvement with the aid of the biological community. One can wait for an overwhelming demand or anticipate the need and try to prepare for it. Whatever the eventual plan, its success will depend on the enthusiasm of the devoted individuals who see that it is carried out.

A possible series of moves to improve the situation might consist of: 1) recommendations on what can be done at the present time to make more effective use of existing facilities; 2) recognition of education as the only promising long-term solution; 3) identification of and encouragement for facilities and programs now in being; and 4) demonstration that perhaps the best catalyst for science as a whole is to bring into existence, under optimal circumstances, a center or centers where interested persons can make maximal contributions.

Where a number of institutions in one area engage in biomedical research, a cooperative unit may best serve their interests; this unit may be comprehensive or may restrict its activity to specified functions. What is done will depend on available facilities, ease of financing, and on the competence and motivation of the participating scientists. Obviously, an initially modest facility may, by its contributions, merit expansion into a large, comprehensive center.

Six specific recommendations were formulated in an effort to crystallize opinion; these are presented here without any implications of priority in time or relative importance:

1) *Information Clearing House:* That steps be taken to establish an information clearing house through which new developments in biomedical instrumentation may be more effectively disseminated, needs may be expressed and needs met, and that this be done either independently or in association with a contemplated instrumentation center. The information center so established would undertake to collate published literature and government research reports as a library service function.
2) *Instrument and Technician Pool:* That a study be made to determine the optimum scope and organization of a pool of instruments to serve biological research, a companion roster and pool of technicians, and a service facility available for performance of specific tasks on an external contract basis.

3) *Instruction:* That provision be made for instruction in the use of standard instruments and for the formal training of technical specialists and technicians so urgently needed.
4) *Professional Status:* That the deliberate design, development and use of instrumentation for research requires close personal association between the biologist and the "engineer"; this association is favored by day-to-day contact in academic surroundings. The senior positions require scholars with a high level of professional training and a primary interest in the work for its own sake, *i.e.*, men eligible for academic or staff appointment.
5) *Student Recruitment:* That students with special aptitudes in instrumentation and the laboratory arts or engineering be identified early and encouraged to seek training in biomedical and biophysical research on the basis that it holds promise of a rewarding career.
6) *Hierarchy of Facilities:* That instrumentation facilities may assume all degrees of complexity, from technical assistant to complete multi-institutional or national center, capable of further development, and that such a hierarchical structure should be fostered.

Sample Model

It can be convincingly argued that the best and quickest way of working out the many troublesome details is actually to establish an institute of instrumentation and to solve each difficulty as it arises. In short, the important thing may be to get started, however modest the initial step. At what point in the hierarchy this venture should be undertaken is a matter for debate, but selection of a particular level should not be inferred as prejudicing the adoption of action programs at other levels. And it remains a possibility that all of the suggested solutions are desirable; that what is most needed is to provide cooperation and coordination. Consideration should be given to the wisdom of forming a group to facilitate this coordination, and to the returns which might be expected from a possible network of consultants. Or, such a coordinating agency, pulling together the new unconnected elements of biomedical instrumentation, might be precursor to an eventual instrumentation center.

Action is needed. Needs, when recognized, are normally met. Full implementation might most easily be achieved by breaking the problem into segments: letting the manufacturers provide periodic training courses; setting about the immediate establishment of an information transfer center; helping individual facilities incorporate more and more of the features shown by the survey to be advisable; on a larger scale, establishing new units to be operated as conjoint enterprises of several institutions properly disposed; and, as the most inclusive solution, eventually forming one or more comprehensive national institutes.

Those who are engaged in planning instrumentation research facilities for biomedicine, in expanding existing laboratories and soundly operating them, might profit substantially from reports of experiences in various universities about the country. Such reports have been presented in conferences sponsored by the Biophysics and Biophysical Chemistry Study Section of the NIH, and are available there. We are here particularly concerned with formulating a sample blueprint for that type of instrumentation facility judged most promising by the participants in the present survey—a *regional center* developing from the joint activities of a group of interested universities and research laboratories.

Cooperative Mechanisms

There is a new and healthy trend in U. S. academic circles toward cooperative enterprise—witness the parceling out of special functions by the Southern Regional Education Board, cooperation such as that among medical schools of the mountain states; the activities of Associated Universities, Inc., responsible for administration of the National Laboratory at Brookhaven; and the Midwestern Universities Research Association. In these instances, for the first time, universities have been willing to recognize their individual limitations and make compacts for integrative action.

Certain advantages of this machinery for an instrumentation effort were immediately apparent to the committee: it fits into the sought-after regional concept; it avoids the special difficulties inherent in the local campus scene; and it makes possible a breadth of coverage beyond that available to existing facilities without the risks involved in setting up a very large, independent center. Whenever a number of institutions pool their efforts the program acquires two distinct aspects: 1) a corporation of some sort, of which the individual universities are members; and 2) provision for actual use of the instrumentality by members of the corporation and/or others. It seems wisest to go slowly in any plan to concentrate a very large accumulation of physical equipment and buildings in a new center, and to rely as much as possible on a "paper" organization with coordinating machinery for encouraging the development of special facilities scattered judiciously throughout the cooperating group, each member providing an appropriate fragment toward the harmonious pattern of the whole.

Desiderata

Whether the regional center here advocated exists in the abstract as the effective combination of separate elements in cooperating institutions, or as a physically distinct facility, more or less self-sustaining, certain specific items must be taken into account:

1) Assurance that there are able, interested scientists involved who are strongly aware of the need for

physical and personal aids for the conduct of quantitative biomedical research, and who would use them if they were made available.

2) The kind of physical and intellectual environment commonly associated with research in the physical sciences.

3) An atmosphere conducive to collaboration among scientists from various fields, for example, assistance to a biologist in acquiring a technique which he intends to employ; of teamwork among a group, each of whom brings his specialized knowledge to bear on a common problem; of free, informal discussion; of formal advisory aid from staff experts in the planning, design and construction of the experimental tools of research and assistance in their proper use.

4) Formal courses as well as informal training in the methodology of the physical sciences, including both the philosophical and practical aspects, with emphasis on biomedical problems.

5) Temporary in-residence association, with the center, of scientists from other institutions, as well as permanent faculty appointments to the scientific staff of qualified members of any of the science faculties of the cooperating institutions.

The desirable solid foundation for the establishment and operation of a center would include: a) means for the procurement and handling of biological materials, to maintain their supply and suitability for research studies; b) classrooms and laboratories designed and equipped for teaching and exploring the methodology of the physical sciences in biomedicine; c) associated shops to render the services essential to the research techniques and for the practical training of technicians; d) a supply of standard commercial instruments and such special ones as will meet the requirements, with technicians for their maintenance and, where indicated, their operation; and e) an adequate library to support the research and teaching program.

A model center, established to meet the desiderata just enumerated, would have, in the aggregate, a staff competent to handle, for example: mechanics, heat and acoustics; optics, with emission and absorption spectroscopy, microscopy and photography; electricity and magnetism, with electrical measurements, electronics and control devices; radiation, with X rays, X-ray and electron diffraction, electron microscopy, radioactive sources and tracers; hydrogen ion and oxidation-reduction potentials; polarography, chromatography and molecular spectroscopy, etc. It would be equipped with commercially available instruments and apparatus; new devices would develop out of specialized research requirements. The shops would be staffed by technicians qualified to teach and train, and one or more competent reference librarians would be employed. Obviously the "paper" phase, or counterpart, of the regional center would not assemble staff and equipment in a single spot, but would otherwise fit the pattern outlined.

Administration and Support

At this point, it will be necessary to find someone to take charge of the action phase, for even if the particular solution offered in this summary is not accepted, the basic considerations will stand. A likely move, once the survey is completed, is to call an organizational meeting of persons invited from universities, industry, granting agencies, instrument societies, etc., to decide what may wisely be undertaken. It will be up to the individuals at each institution to assess their own place in the total scheme, what they can be expected to do, given a reasonable amount of guidance.

The method of choice in financing will depend on the particular idiom adopted, but cannot fairly be considered an insurmountable obstacle. Money is available in large amounts for the right kind of program, and individuals who are convinced of the need for action must somehow be made aware of this fact. Partial support, at most, can be expected from industry. There is no question, on the other hand, of strong federal support for central or regional centers and perhaps even for the construction of new space for smaller laboratories now inadequately housed. Certainly, it seems a proper government function to establish and finance an information clearing house and library facility.

Within a wide limit of comprehensiveness, a meritorious plan, carefully worked out and presented by scientists of recognized competence, should encounter no special difficulties in obtaining funds from a number of government agencies. Later, possibly, funds from private sources may become increasingly available and there is always the likelihood of some support from the institutions comprising the managing corporation which administers the regional center.

Acknowledgment

Permission to reprint this report was obtained from the National Academy of Sciences—National Research Council, upon the approval of its Division of Biology and Agriculture. Regarding the origin and endorsement of the report, Dr. Paul Weiss, then Chairman of the Biology Council that sponsored it, wrote at follows: "The Biology Council has adopted the report substantially as it was submitted in summary by Dr. Paul Klopsteg [who surveyed the subject alone] with the endorsement of the Drafting Committee (Philip H. Abelson, Arnold O. Beckman, H. Stanley Bennett, Henry C. Meadow, Francis O. Schmitt, and Paul Weiss)." After further editing by Dr. R. B. Stevens, the report was endorsed by the Biology Council, of which he was Executive Secretary.

On the Role of the Engineer in Biomedical Instrumentation*

JOHN P. HERVEY†

Summary—**The reprinting in this issue of the PROCEEDINGS of the NAS-NRC study of "Instrumentation in Biomedicine" gives occasion for a brief review of the reasons why effective cooperation of engineers with biomedical scientists requires close association on a full-time basis, preferably in an academic or quasi-academic organization. Similarly, the proposed national or regional "institutes for instrumentation" should have academic purposes and attributes. For these institutes to be fully effective, however, there will still be a need for engineers or physical scientists *within* each biomedical organization. While most engineers will become specialists to some degree, they should do so only after receiving the broadest possible training in the physical sciences.**

NEWCOMERS to the field of biomedical instrumentation should perhaps be reminded that it has been nearly forty years since the thermionic vacuum tube was first used as an amplifier in a biology laboratory. The impressive advances made in physiology, to name just one science, in this span of time bear eloquent witness to the advantages which accrue from the use of the most refined instrumentation (of whatever sort) that can be brought to bear on a research problem. For most of this period, however, the use of physical methods in instrumentation has been restricted to those biologists who already have had, or were willing to acquire, a considerable degree of competence in the physical sciences. In only a few university departments or endowed institutes were there engineers whose function it was to develop suitable equipment. Through the years, it became increasingly evident that the biological sciences could use much more physical instrumentation than was available to them; but because suitable apparatus was rarely available from industry, new ways had to be sought to overcome the lack.

The recognition of this need led to the investigation by the Biology Council of the National Academy of Sciences, which is summarized in a report entitled "Instrumentation in Biomedical Research" and reprinted elsewhere in this issue.[1] This report warrants careful study by all members of the PGME. While it is clear from even a casual reading that there was not unanimous agreement on some of the major proposals, such as the establishment of a central or national institute for instrumentation, nevertheless, there is much for the engineer to learn from the exposition of the differing views, and it is important for him to know how a group of distinguished biologists (using the word in a broad sense) evaluates the needs.

Certainly, the limitations of staffing and financing instrumentation development under the usual academic pattern have been clearly shown; yet, it is also obvious that the biologist must have access to a diversity of professional talent in the physical sciences if biology is to prosper. The newcomer might well ask why this need cannot be met within the traditional academic framework of cooperative arrangements between biomedical departments and engineering schools, or at the institutional level, by cooperative arrangements between universities. A few such arrangements at the departmental level exist, but more have never progressed beyond the initial exploratory phase, presumably because one party or the other felt that the resulting benefits would flow chiefly in one direction. Also, there is a need for continuity of personnel and interest, which implies participation by faculty members having tenure; and with the current increase of students in the physical sciences, it seems improbable that these individuals could find the time, even if they were so inclined. Another avenue, which is in use to a certain extent, is that of development contracts with industrial concerns or sponsored-research organizations. However, the high costs of such work appear to limit its usefulness to the development of high-speed data-processing machines (where the cost can be spread among a large number of individual users), or the development of equipment which will be manufactured in relatively large numbers for routine diagnosis or therapy.

This brings us to a brief consideration of the new institutes for instrumentation proposed by the Biology Council. As the report notes, the disadvantages of a single national institute, as opposed to several regional ones, are numerous and substantial. (Other objections could be added, such as the hazard that a national center, by reason of its geographical isolation, could become as noncompetitive and parochial as some of the university endeavors which the report criticizes.) It seems probable then that it is the regional institute with which we should concern ourselves here. At the risk of seeming to be merely repetitious, let us emphasize again the parts of the Council's "desiderata" which are of immediate relevance to engineers. They say that this organization should have:

* Original manuscript received by the IRE, September 2, 1959.
† Rockefeller Institute, Woods Hole, Mass., and New York, N. Y.
[1] P. E. Klopsteg, "Instrumentation in biomedical research," this issue, p. 1999.

"2) The kind of physical and intellectual environment commonly associated with research in the physical sciences.

"3) An atmosphere conducive to collaboration among scientists from various fields, for example, assistance to a biologist in acquiring a technique which he intends to employ; of teamwork among a group, each of whom brings his specialized knowledge to bear on a common problem; of free, informal discussion; of formal advisory aid from staff experts in the planning, design and construction of the experimental tools of research and assistance in their proper use.

"4) Formal courses as well as informal training in the methodology of the physical sciences, including both the philosophical and practical aspects, with emphasis on bio-medical problems.

"5) Temporary in-residence association with the center, of scientists from other institutions, as well as permanent faculty appointments to the scientific staff of qualified members of any of the science faculties of the cooperating institutions."

Obviously, what is proposed is a serious minded, academically-oriented organization whose special attributes of full-time dedication to instrumentation for biomedical research would overcome the special disadvantages of the cooperative arrangements discussed above. Wisely, the Council suggests that such a center may grow from a "paper" organization, based on those special instrumentation facilities which are, or will be, in existence among the cooperating institutions.

Even were the institutes for instrumentation now in being, there would remain a need in each biomedical group for persons professionally trained in the physical sciences to be in residence on an everyday basis. It is just not feasible, as has been proposed by one prominent engineer, to learn the rudiments of biology or medicine while golfing with biologists or physicians! There are many reasons for this, but to mention the more important ones will suffice here.

As a rule, biomedical people and those in the physical sciences lack a common language in the operational sense, and only by day-to-day association can this be acquired to the point where effective interchange of ideas becomes possible. Furthermore, only the physical scientists who have this close association in a research atmosphere can come to appreciate the functional subtleties and structural complexities of the living organism. They are the people who will form the structure of the bridge between the biomedical group and the institute for instrumentation when established.

Obviously, all these are large changes, and it is as yet too early to know how and to what extent they will be achieved. But it is only by making a large effort to overcome the present deficiencies that the biomedical sciences can reach the stage where no well-conceived research proposal is frustrated simply for lack of professional personnel to instrument it.

For the young engineer who wonders whether he should enter this field now, one must emphasize once more the importance of the broadest possible training based on as much mathematics, physics and physical chemistry as he can acquire. It seems very doubtful that he should attempt to orient his study in the direction of biology during his undergraduate years; it is far better for him at this stage to achieve the broadest engineering education available to him. If, after a year of part-time teaching and graduate study, or perhaps a year in industry, he feels that his interests and satisfaction lie in instrumentation, he should be in a position to enter the field and grow with it, becoming, let us hope, what the report calls a "creative instrumentologist."

Medical Electronics Center—Interdisciplinary Coordination*

V. K. ZWORYKIN†

Summary—The Medical Electronics Center of the Rockefeller Institute seeks to advance the application of electronic techniques in the life sciences and medicine through conferences on specialized topics, by the publication of the *Bibliography of Medical Electronics* and by acting as a clearing house of ideas and facilities in the field of medical electronics. However, a much greater effort is needed to realize the potentialities of electronics in medicine. An important step forward, overcoming some of the major obstacles to progress, would be the creation of one or several Medical Engineering Research Institutes, having the specific object of facilitating the development, initial production, and clinical testing of apparatus for medical research and practice, and serving at the same time as a training center in the boundary field of medical electronics.

* Original manuscript received by the IRE, September 1, 1959.
† RCA Laboratories, Princeton, N. J. Also affiliated with the Rockefeller Institute, New York, N. Y.

Nowhere is the problem of utilizing the advances of knowledge more acute than in the field of application of the latest technology to problems in biology and medicine. The physician and biological investigator have carried their art almost as far as can be expected within their unaided sensory limitations. In order to extend the sensitivity frequency, spectral responses, and bandwidth of the human sciences, resort must be had to modern, and therefore, electronic instrumentation.

A number of texts have appeared recently that attempt to cover the field of physical techniques and biological research.[1,2] Such texts, while they serve a useful purpose, are obsolete almost from their publication date, particularly since relatively poor international communication inhibits the proper interchange of information.

This paper describes an effort to overcome some of these difficulties in the form of a Medical Electronics Center, which was organized in 1955 at the Rockefeller Institute for Medical Research (now the Rockefeller Institute).

In an attempt to improve interdisciplinary communication, the Center holds conferences, usually on an invitational basis, on problems which have, as yet, not been attacked by electronic methods, or where the application is so new that the problems themselves are not well understood. Here, electronic engineers, biophysicists, biologists, medical doctors, computer programmers, etc., have an opportunity to learn about each other's language and problems.

Conferences have been held on the subjects of electroencephalography and its amplifier requirements, the detection of micro amounts of protein, artificial pacemakers, and cardiac prosthesis. At a full-scale conference, the problems of the design and use of artificial internal organs and their feedback controls were elucidated. The proceedings of this conference were published by the IRE earlier this year.

Another project at the Medical Electronics Center is the *Bibliography on Medical Electronics*, the first edition of which has already been published, in 1958, by the Professional Group on Medical Electronics and distributed to all members of the Professional Group.

The problem of finding information concerning the applications of electronic techniques to medicine is probably more difficult than any other literature-searching problem. Suitable references may appear in almost any medical or electronic publication. The literature searcher is familiar only with electronics or medicine. In the task of preparing this *Bibliography*, we have been fortunate enough to have the assistance of a great many of the members of the PGME, and, due to the international contacts of the Rockefeller Institute, an increasing number of the contributors to the *Bibliography* are from other countries. Literature from all sources will be welcomed by the *Bibliography* staff at the Rockefeller Institute. A first supplement to the original *Bibliography* was issued in June, 1959.

In its activities, the Medical Electronics Center also acts as a clearing house, attempting to coordinate the activities of, and introduce to each other, the biological and medical investigators with problems and those in the physical and electrical sciences who have knowledge of techniques.

With its limited staff, however, the Medical Electronics Center is perforce limited to the role of an interdisciplinary catalyst.

Although many specialized institutes exist in the United States and elsewhere, devoted exclusively either to biological investigations or to the physical sciences, there is no research institute with the objective of furthering the application of engineering knowledge to medicine. The result is that these problems are, as yet, attacked in a relatively sporadic and uneconomical fashion.

Many hospitals now have small groups devoted to health physics and electronic instrument repair. Conversely, most commercial electronic organizations have one or two engineers working on the potentially large medical electronic market. These activities are, in general, small, understaffed, and uneconomical.

The problem of instrumentation in the bio-sciences and its proper organization has been considered of sufficient importance so that it has been the subject of a special National Academy of Science Report, "Instrumentation in Bio-Medical Research."[3] This report has proposed, among other things, that a model instrumentation committee be established which would "have in the aggregate, a staff competent to handle, for example: mechanics, heat, and acoustics; optics, with emission and absorption spectroscopy, microscopy, and photography; electricity and magnetism, with electrical measurements, electronics and control devices; radiation, with X rays, X-ray and electron diffraction, electron microscopy, radioactive sources and tracers; hydrogen ion and oxidation-reduction potentials; polarography, chromatography, and molecular spectroscopy, etc. It would be equipped with commercially available instruments and apparatus; new devices would develop out of specialized research requirements. The shops would be staffed by technicians qualified to teach and train, and one or more competent reference librarians would be employed."

Nowhere does such a facility exist today, although in its capacity as a clearing house, the Medical Electronics Center has available, usually on a voluntary basis, the consultative services of many of the large commercial electronic organizations.

[1] G. Oster and A. W. Pollister, "Physical Techniques in Biological Research," Academic Press, New York, N. Y.; 1955–1956.

[2] J. H. Lawrence and C. A. Tobias, "Advances in Biological and Medical Physics," Academic Press, New York, N. Y.; 1953–1959.

[3] P. E. Klopsteg, "Instrumentation in biomedical research," this issue, p. 1999.

Even with the existence of the proposed consultation, a number of serious bottlenecks exist in the widespread practical implementation of new instrumentation discovery. The individual scientist may assemble equipment having, for his particular purpose, some commercially available components held together by old rubber bands, adhesive tape, spit, and chewing gum. Such instrumentation is generally capable of being operated properly only in the original investigator's laboratory. Nevertheless, the spectacular results (from a medical viewpoint) obtained with such aggregations, frequently leads to their duplication, *in toto*, by many other investigators, and subsequently gives all electronic instrumentation an unsavory reputation for unreliability.

On the other hand, even where equipment has been properly designed by reputable, competent electronic engineering organizations, usually with the development paid for by government or medical subvention, the equipment never reaches the state of the large-scale manufacture it deserves because of the lack of suitably trained management, sales, and production personnel in the commercial organization, leading to a reluctance to enter the field of medical electronics.

In order to attack both the problems of the initial instrumentation engineering and also the initial distribution to a sufficient number of investigators, I have proposed the creation of endowed, medical engineering research institutes which would operate in the following manner.

The functions of such institutes would be to expedite the translation of electronic, electrical, and mechanical engineering advances into practical tools and to encourage their use by the medical profession. They would provide support and work opportunities for the inventor whose lack of technical assistance and physical facilities results in today's loss of valuable ideas bearing upon medical needs. They would serve as clearing houses for engineering information relating to medicine and as a point of liaison between instrument manufacturers and hospitals or medical groups. This would help to overcome any undue conservatism among doctors in regard to new medical procedures by ensuring the distribution of small numbers of new devices to hospitals where doctors could become acquainted with their advantages by first-hand experience.

A precedent for the establishment of the Medical Engineering Research Institutes exists today in the successful Institute of Experimental Surgical Apparatus functioning in the Soviet Union. The Russian institute employs large groups of mechanical and electrical engineers and medical specialists and is provided with a special hospital, laboratory, and pilot plant. Recent exhibits of its work, at the Brussels World's Fair and elsewhere, have given impressive evidence of the substantial and useful results that can be achieved through a concerted effort in applied medical engineering.

Proposals, essentially similar in nature and with provision for education in medical engineering, have also been made almost simultaneously in many parts of the country, notably at Drexel Institute of Technology, Philadelphia, Pa. (D. LeCroissette), Columbia University, New York, N. Y. (H. Zinsser), and Johns Hopkins University, Baltimore, Md. (S. Talbot).

The establishment of a Medical Engineering Research Institute in this country could be undertaken on a relatively modest scale, with a total endowment of perhaps half a million dollars. Initially, it would require a small permanent staff which would guide and advise part-time workers and scientists or engineers on sabbatical leave from their regular positions. Apparatus could be constructed on order by instrument manufacturers and tested and demonstrated at hospitals affiliated with the proposed Institute. The Institute would acquire its own laboratories and shops only after it had demonstrated the need for them by the results of its initial operations on a smaller scale. It is reasonable to expect that the Institute would become entirely self-supporting within a relatively short period.

The creation of such an endowed Medical Engineering Research Institute could overcome a major deterrent which exists today in the absence of an adequate financial return for individual companies in all but a few specialized areas of medical engineering. Discussion of this proposal already has aroused an enthusiastic response among various companies and professional organizations. Several companies now engaged in activities that would fall within the province of the proposed Institute have pointed out that this work is proceeding at the present time on an extremely limited scale, simply because it does not fit in with their regular activities. In these cases, it has been suggested that the companies would assign such work to the Institute and share the development costs during the pioneering period.

A further advantage of such institutes would be to encourage invention and promote the exchange of ideas and experience, generating an ever greater flow of new devices, systems, and techniques. Thus, the institutes would fill a definite and alarming gap in the pattern of medical practice in the United States today. In this basic sense, they would be eminently worthy candidates for funds assigned to philanthropic purposes.

Correspondence

Semiconductor Varactors Using Surface Space-Charge Layers*

We propose here a family of semiconductive devices characterized by a voltage-sensitive capacitance that resides in the space-charge region at the surface of a semiconductor bounded by an insulating layer. One attractive application for such devices is to parametric amplification, for which they have certain advantages over p-n diodes. Another potential application is to distributed parametric amplification. An unusual feature of some of the structures is that large concentrations of excess hole-electron pairs can be produced in the semiconductor.

The existence of an electrical capacitance at the surface of a semiconductor, and the ability to modulate it by means of an electric field applied normal to the surface, are well known.[1] At low frequencies, the capacity arises partly from the space-charge region and partly from surface states. Since the surface state trapping times are believed (at least for germanium) to be $\sim 10^{-8}$ seconds,[2,3] the effect of surface states on the capacitance may be ignored at waveguide frequencies. Fig. 1[4] shows the space-charge capacitance as a function of the surface barrier height and of the dimensionless parameter (p_0/n_i) characterizing the body resistivity.

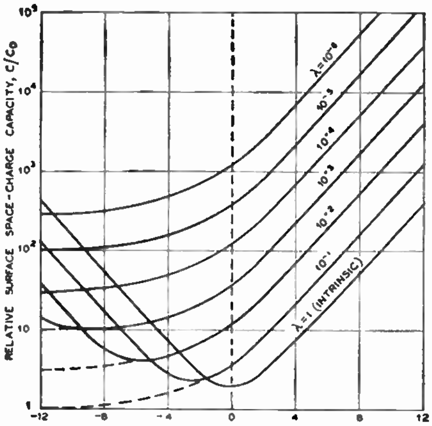
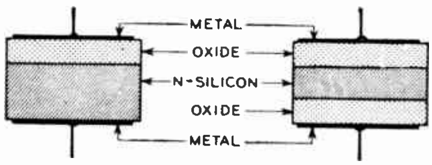


Fig. 1—Relative surface space-charge capacity (C/C_0) vs barrier height $(Y \equiv eV/kT)$. The running parameter $(\lambda \equiv p_0/n_i)$ defines the body resistivity. C_0 is $\epsilon\epsilon_0/l$, where l is the intrinsic Debye length; C_0 is equal to 8.6×10^{-9} far-cm^{-2} for germanium, 3.9×10^{-10} far-cm^{-2} for silicon. For p-type samples, replace Y by $(-Y)$ and λ by $(1/\lambda)$.



(a) (b)

Fig. 2—(a) An ON varactor. (b) An ONO varactor.

Consider the ON structure shown in Fig. 2(a), where N stands for a layer of semiconducting material (n-type silicon, for example), and O is an insulating dielectric (SiO_2, for example). The capacitance between the metal film and the ohmic contact to the semiconductor consists of the capacitances of the surface space-charge region and of the oxide layer in series. If the oxide layer is sufficiently thin, the capacitance of the device is substantially that of the space-charge region; most of the voltage applied between the terminals of the device appears across the space-charge region and serves to modulate its capacitance. With the oxide *positive* with respect to the n-type semiconductor, an *accumulation* layer results (the righthand branches of the curves in Fig. 1), and the capacitance, C, goes up as exp $(eV/2kT)$; with the oxide *negative*, one gets first a *depletion* layer [for which C varies as $(-V)^{-1/2}$], followed, at sufficiently high voltage, by an *inversion* layer [for which C varies as exp $(-eV/2kT)$]. The state of the surface under zero applied voltage depends on the numbers and energies of the surface states, and can be controlled to a considerable extent by control of the conditions under which the oxide layer is formed.[5]

Consider first a modulation scheme in which the negative swing of voltage is insufficient to take the surface into the inversion layer region. Since only majority carriers are required, problems of generation of minority carriers will not occur, and the frequency response of the device is limited only by the rate at which the surface capacitance can be charged by current flow through the bulk of the semiconductor. The lower limit to the capacitance will now be given by the minimum in one of the curves of Fig. 1; the upper limit is set by the capacitance of the oxide layer. Substitution of typical values indicates that the voltage sensitivity of the device can easily be made to extend over several decades of C before the latter limit is reached.

For use in a parametric amplifier the performance of a device of this class is described by two parameters: the cut-off frequency, $(1/2\pi RC)$, where R is taken to include *all* losses due to in-phase currents, and the proportionate change in capacity for a significant change in voltage (say kT/e). In both respects the proposed devices compare favorably with p-n diodes. Losses from the in-phase diffusion conductance of a p-n diode (a necessary concomitant, in the forward direction, of a nonlinear diffusion capacitance) are absent; and the proportionate change in capacity is still high, because of the exponential dependence of C upon V in the "accumulation" region.

The exact behavior at large reverse voltage of an ON structure having a nonconducting oxide depends on the lifetime τ, which is here considered as including both body and surface generation. If $\tau \ll (1/f)$, where f is frequency, the extra holes required for the inversion layer are generated rapidly, and the inversion-layer capacitance is exhibited almost instantaneously. If $\tau > (1/f)$:

1) for the first few cycles after application of the microwave voltage, the device will behave as though the depletion layer capacitance continued out to large negative values of V (dotted lines in Fig. 1);
2) however, if the microwave voltage persists, extra holes will accumulate, and inversion layer capacitance will be exhibited.

(Note that these two capacitances have opposite responses to a change in V.) However, if the oxide has sufficient conductance, holes cannot accumulate to form an inversion layer, and only the depletion layer capacitance will be exhibited. The required conductance is small, comparable to that of a reverse-biased p-n junction.

An ONO structure is shown in Fig. 2(b). It is similar to a p-n-p diode in that C varies at twice the pump frequency. If the two surfaces are identical, a small applied voltage decreases one space-charge capacitance by the same amount that the other is increased, and the net effect is nil. At higher voltages, however, the depletion capacitance, being smaller, outweighs the accumulation capacitance, and the net effect is no longer nil. For still larger voltages, if sufficient time is allowed for the necessary carriers to be generated and if the oxides are nonconducting, an accumulation layer will exist on one side and an inversion layer on the other, the effects of which will *add*. An ONO, having no metal-semiconductor contact, is free of contact resistance problems met in p-n diodes.

The simplicity of the ONO structure, and the fact that no electrical connection need be made to the semiconductor, encourage the consideration of laminates, *e.g.*, ONONO . . . , or tape-like composites of oxidized silicon particles embedded in a dielectric medium. Laminar structures could be used to fabricate microwave devices whose active regions are no longer small in extent in comparison with a wavelength, thereby simplifying construction, increasing power handling, and at the same time opening up the possibility of making extended structures with distributed gain properties.

The carrier-generation effect may prove to be of value quite apart from parametric amplifying considerations. If a large voltage is applied to the terminals of an ONO or OIO

* Received by the IRE, July 31, 1959.

[1] W. L. Brown, W. H. Brattain, C. G. B. Garrett, and H. C. Montgomery, "Semiconductor Surface Physics," University of Pennsylvania Press, Philadelphia, Pa., p. 111; 1957.

[2] H. C. Montgomery, "Field effect in germanium at high frequencies," *Phys. Rev.*, vol. 106, pp. 441–445; May 1, 1957.

[3] C. G. B. Garrett, "High frequency relaxation processes in the field effect experiment," *Phys. Rev.*, vol. 107, pp. 478–487; July 15, 1957.

[4] C. G. B. Garrett and W. H. Brattain, "Physical theory of semiconductor surfaces," *Phys. Rev.*, vol. 99, pp. 376–387; July 15, 1955.

[5] M. M. Atalla, E. Tannenbaum, and E. J. Scheibner, "Stabilization of silicon surfaces by thermally grown oxides," *Bell. Sys. Tech. J.*, vol. 38, pp. 749–784; May, 1959.

(I denotes "intrinsic"), a layer of holes and a layer of electrons, largely produced by thermal generation, will appear at opposite faces of the semiconductor. When the voltage is removed, these carriers flow into the body of the semiconductor. The maximum concentration, P, of electron-hole pairs, averaged over the volume of the semiconductor, before recombination begins is readily shown to be:

$$P = \frac{\epsilon\epsilon_0 E}{et}$$

where $\epsilon\epsilon_0$ and E denote dielectric constant and dielectric strength of the dielectric layers; e, the electronic charge; and t, the thickness of the semiconductor. If an alternating voltage be applied, of frequency $f > (1/\tau)$, there will be a steady-state concentration of electron-hole pairs which is somewhat smaller than P. Excess carrier concentrations $> 10^{16}$ cm^{-3} appear readily attainable. This technique may be of value for the study of trapping in semiconductors and for devices depending on conductivity modulation.

W. G. Pfann
C. G. B. Garrett
Bell Telephone, Labs., Inc.
Murray Hill, N. J.

WWV Standard Frequency Transmissions*

Since October 9, 1957, the National Bureau of Standards radio stations WWV and WWVH have been maintained as constant as possible with respect to atomic frequency standards maintained and operated by the Boulder Laboratories, National Bureau of Standards. On October 9, 1957, the USA Frequency Standard was 1.4 parts in 10^9 high with respect to the frequency derived from the UT 2 second (provisional value) as determined by the U. S. Naval Observatory. The atomic frequency standards remain constant and are known to be constant to 1 part in 10^9 or better. The broadcast frequency can be further corrected with respect to the USA Frequency Standard, as indicated in the table; values are given as parts in 10^{10}. This correction is *not* with respect to the current value of frequency based on UT 2. A minus sign indicates that the broadcast frequency was low.

The WWV and WWVH time signals are synchronized; however, they may gradually depart from UT 2 (mean solar time corrected for polar variation and annual fluctuation in the rotation of the earth). Corrections are determined and published by the U. S. Naval Observatory.

WWV and WWVH time signals are maintained in close agreement with UT 2 by making step adjustments on time of precisely plus or minus twenty milliseconds on Wednesdays at 1900 UT when necessary; retarding time adjustments were made on August 5 and 26, 1959.

WWV Frequency†

1959		#1	#2	#3
August	1	−29	−32	
	2	−30	−31	
	3	−30	−31	−26
	4	−30	−31	−25
	5	−30	−31	−27
	6	−31	−31	−26
	7	−31	−30	−25
	8	−31	−30	
	9	−31	−30	
	10	−31	−30	−25
	11	−31	−30	−26
	12	−30	−31	−25
	13	−30	−31	−26
	14	−29	−31	−27
	15	−29	−31	
	16	−29	−31	
	17	−28	−31	−26
	18	−28	−30	−26
	19	−28	−30	−26
	20	−27	−30	−27
	21	−27	−30	−26
	22	−28	−30	
	23	−28	−30	
	24	−29	−30	−26
	25	−29	−30	−25
	26	−29	−30	−26
	27	−29	−30	−26
	28	−29	−30	−26
	29	−28	−30	
	30	−28	−30	
	31	−28	−30	−26
September	1	−27	−30	−26
	2		−30	−26
	3		−29	−25
	4		−29	−25
	5		−30	
	6		−30	
	7		−29	
	8		−29	−25
	9		−29	−26
	10		−29	−25
	11		−29	−25
	12		−29	
	13		−29	
	14		−30	−24
	15		−29	−25
	16		−28	−25
	17		−29	−26
	18		−34	−31

† WWVH frequency is synchronized with that of WWV.

Column #1 Vs NBS‡ atomic standards, Boulder, Colo., 30-day moving average seconds pulses at 15 mc.

Column #2 Vs atomichron at WWV, measuring time one hour at 2.5 mc.

Column #3 Vs atomichron at the U. S. Naval Research Laboratory, Washington, D. C., measuring time 56 minutes at 2.5 mc.

‡ Method of averaging is such that an adjustment of frequency of the control oscillator appears on the day it is made. No change or adjustment in the control oscillator was made during August, 1959; a frequency adjustment of minus 5 parts in 10^{10} was made on September 17, 1959.

National Bureau of Standards
Boulder, Colo.

* Received by the IRE, September 28, 1959.

Electrical Characteristics of Some Gallium Phosphide Devices*

Several interesting effects have been observed during the course of an investigation of the electrical behavior of gallium phosphide crystals.

Thin platelets of GaP prepared in the manner described by Wolf, Keck, and Broder[1] were etched in mixtures of hydrochloric and nitric acid. Large area contacts were then made to one surface of the platelets by ultrasonic agitation of the platelets in liquid indium-gallium. Contacts to the top surface of the platelets were made with tungsten probes.

Most platelets were polycrystalline and appeared to contain multiple internal junctions. Such platelets electroluminesce when sufficient voltage is applied across their surfaces to cause current flow irrespective of the polarity of the voltage applied at each surface contact.[2] The voltage-current characteristics are similar to that of two diodes placed back to back (*n-p-n* or *p-n-p* structures) resulting in low leakage current flow until sufficient voltage is applied of the correct polarity to break down one or the other of the diodes.

The experiments to be discussed were, however, performed using single crystal platelets which were obtained by selection or by cleavage of larger platelets. When contacts were applied to several of the single crystal platelets, single diode forward and reverse characteristic curves were obtained, no electroluminescence was obtained in the forward bias condition even at currents as high as 200 ma, and the variation of capacitance with voltage across the surfaces was similar to that of a unijunction device. The diodes electroluminesced in the region of reverse voltage breakdown in a manner similar to that observed for other semiconductors; in the case of a cleaved platelet containing one internal junction, white light appeared as small dots on the surface and as white lines along the edges; in other cases, white light appeared directly under the top contact (tungsten probe).

Many of these diodes had reverse leakage currents of less than 0.03 μa (exact values were not determined because "pickup" interfered with the measurement) and breakdown voltages between 5 and 10 volts.

Fig. 1(a) shows a typical forward characteristic curve for the GaP diodes while Fig. 1(b) is the forward characteristic curve of a germanium point contact diode on the same scale. The voltage at the point of inflection of the GaP characteristic curve is 1.2 volts; this compares with inflection voltages of 0.55 volt and 0.2 volt for silicon and germanium, respectively. The bandgaps of the three materials are GaP—2.25 ev, Si—1.1 ev, and Ge—0.7 ev. The slope of the curve above the point of inflection corresponds to a 90 ohm resistance for the GaP diode and a 50 ohm resistance for the Ge diode. The forward resistance of the GaP diodes could be decreased by "forming" in similar fashion to point contact transistor "forming." A "formed" large area diode yielded a forward current of 0.5 a at 5 volts, showing a limiting resistance of 8 ohms. This diode had a capacitance of 40 $\mu\mu$f at 2 v reverse bias and a leakage current of less than 1 μa.

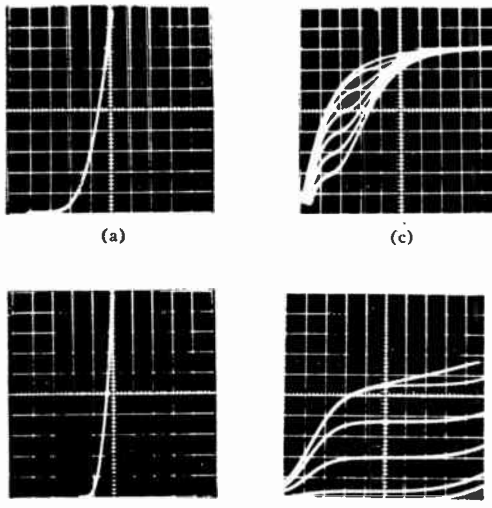
Preliminary temperature tests made on a few diodes showed that the reverse leakage currents increase to about 1 μa at 250°C while one diode showed rectification at 800°C.

GaP diodes of this type are potentially useful for high temperature applications in missile and satellite instrumentation, and, in addition, appear promising for rapid recovery switching networks. These diodes are visibly electroluminescent in the region of reverse breakdown voltage, and the turn-

* Received by the IRE, July 16, 1959.

[1] G. A. Wolf, P. H. Kack, and J. D. Broder, "Preparation and properties of the III–IV compounds," *Phys. Rev.*, vol. 94, p. 753; May, 1954.

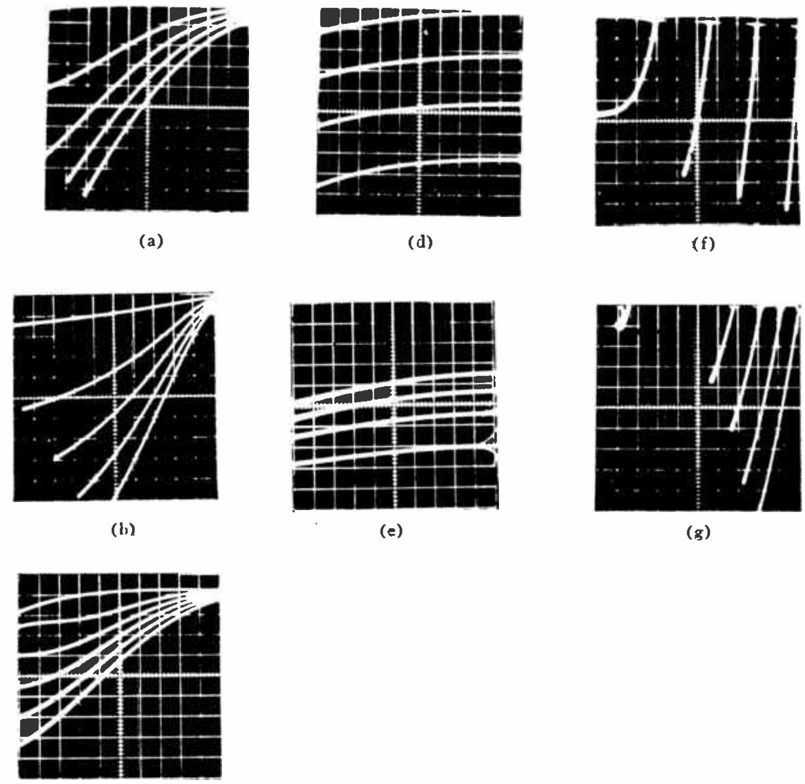
[2] G. A. Wolf, R. A. Hebert, and J. D. Broder, "Electroluminescence of gallium phosphide," *Phys. Rev.*, vol. 100, p. 1144; November, 1955.



(a) (c)

(b) (d)

Fig. 1.



(a) (d) (f)

(b) (e) (g)

(c)

Fig. 2.

on and turn-off time of the electroluminescence is being investigated for use in electronic control systems.

Placing another point on the platelet surface in the vicinity of the first point and applying forward bias current steps from a constant current supply at one point while sweeping the other point over a range of reverse voltages produced a display of output characteristics curves as shown in Fig. 1(c) (a Tektronix Transistor Curve Tracer was used for these displays). The curves appear similar to those of a transistor having a very high saturation resistance and an alpha of approximately one. However, the voltage saturation occurs, as can be seen, in the breakdown region of the reverse biased output diode. Curves of this type could be obtained on many of the crystals tested even when the points were spaced more than 100 mils apart. What appears to be a saturation region actually occurs for values of collector voltage at which the collector "shorts" to the emitter. The "shorting" of the two widely spaced terminals may be related to the electroluminescent properties of GaP since the release of large numbers of carriers would tend to decrease the resistance to current flow between the "collector" and "emitter" probes. A display of input voltage vs output voltage ("emitter" vs "collector" voltage or voltage feedback) showed that at values of output collector voltage corresponding to the saturation region, the feedback ratio is 1, showing a low resistance or "short" between "collector" and "emitter." This phenomenon, which occurs across the entire platelet at specific values of controlling voltages, has several possibilities for practical application.

An interesting display of output curves is obtained by biasing the output diode in the forward direction and the input diode in the reverse direction from a constant current stepping supply such as that of the Tektronix Curve Tracer. The input voltage steps are normally well above the breakdown voltage of the reverse biased input diode so that the input diode electroluminesces for each step of input current causing a short between input and output terminals. The constant current steps then appear as current steps in the output diode forward characteristic, giving the effect of a transistor output curve display as seen in Fig. 1(d).

Upon investigation of a large number of crystals, a small percentage of them were found to be capable of producing output characteristic curves as shown in Fig. 2(a). These output curves are very similar to the output curves of a point contact Ge transistor as shown in Fig. 2(b). As seen in Fig. 2(a), the collector current increases as collector voltage is raised from *zero volts*, and the collector current depends upon the emitter current steps in a manner similar to that of the point contact Ge transistor. An experimental analysis of the conditions necessary to obtain the characteristics shown in Fig. 2(a) on GaP crystals showed that for the action to occur, the points had to be spaced extremely closely, in the order of 1 mil or less spacing, and furthermore these curves could be obtained only upon selectively probing and locating active minute regions of single crystal platelets. In order to eliminate the possibilities of nonlinear surface leakage giving rise to the display shown in Fig. 2(a), measurements were made of the leakage currents for different polarities between the two closely spaced probes. The results indicated that surface leakage currents were at least an order of magnitude below the values of collector current obtained in Fig. 2(a). Fig. 2(c) shows the output characteristic curves at relatively high voltages under the condition where this action is found. As the collector voltages is raised from zero, more and more emitter current is drawn to the collector until the region of collector diode reverse breakdown voltage is reached. In this region, a "short" occurs between emitter and collector giving rise to the "saturation region" of the curves as seen in Fig. 2(c).

Fig. 2(d) shows the feedback characteristic, V_e vs V_c, of the device; Fig. 2(e) shows the V_e vs V_c feedback curves of a Ge point contact transistor. The similarity of Figs. 2(d) and 2(e) is notable. Fig. 2(f) and Fig. 2(g) show marked similarity of the I_c vs V_e curves of the GaP and Ge devices. Although not shown, graphs were made of the input diode voltage-current curves for various values of V_c and these were found to be very similar to those of the Ge point contact transistor.

The similarity of the electrical characteristics of the Ge point contact transistor and those of the GaP device shows the existence of active electronic mechanisms in the GaP device.

The author wishes to acknowledge the help of J. Broder and Charles McAfee, in preparing the material for these experiments.

J. Mandelkorn
U. S. Army Signal Res. and Dev. Lab.
Fort Monmouth, N. J.

Multiple-Element Hall-Effect Sensor*

The development of semiconductor materials with high-mobility carriers has resulted in an increased application of Hall-effect to magnetic field measurements.

The Hall-potential in a sample of a rectangular sheet of semiconductor, as shown in Fig. 1, is given by

$$V_H = \frac{R_H IB}{t} \text{ volts} \qquad (1)$$

where R_H is the Hall-coefficient[1] in meters^3/coulomb, I is the current in amperes, B is the magnetic flux density in webers/meter^2, and t is the thickness of the specimen, in meters, in the direction of the magnetic field. This expression is correct, within 5 per cent, for length-to-width ratios greater than 2.5.[2,3]

The sensitivity of such a specimen in the measurement of magnetic fields is then

$$S = \frac{V_H}{B} = \frac{R_H I}{t}. \qquad (2)$$

For a given material (R_H given), the sensitivity increases with the current and decreases with the thickness of the specimen. On the other hand, the current through the sample is limited by heat dissipation from the material.

Setting as a limiting parameter the power per unit area, p, of the thin rectangular sheet, depicted in Fig. 1, of a semiconductor of conductivity σ, the current through the sample is

$$I = \left[(2plw)_\sigma \frac{wt}{l}\right]^{1/2} = (2p\sigma t)^{1/2} w \qquad (3)$$

and the sensitivity becomes

$$S = R_H \left(\frac{2p\sigma}{t}\right)^{1/2} w. \qquad (4)$$

Thus, in order to increase the sensitivity of the device, its dimension in the direction of the Hall-potential should be enlarged. The minimum thickness is determined by the required mechanical strength.

The increase in source resistance of the Hall-potential is of secondary importance since the resistivity of most of the recently-used intermetallic semiconductors for Hall-effect sensors is very low, of the order of 10^{-5} ohm-meter.[4]

Hall-effect sensors have been designed[5] in which the length-to-width ratio has been reduced without the associated reduction in Hall-potential. Similar to a battery composed of cells, several Hall-effect sensors, each with $l/w=3$, have been connected in series, while the biasing current was fed in parallel (see Fig. 2).

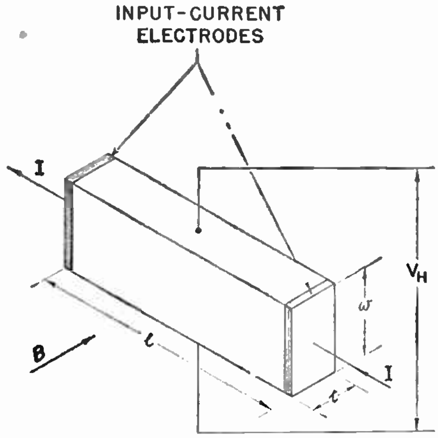


Fig. 1—Schematic view of a Hall-effect sensor.

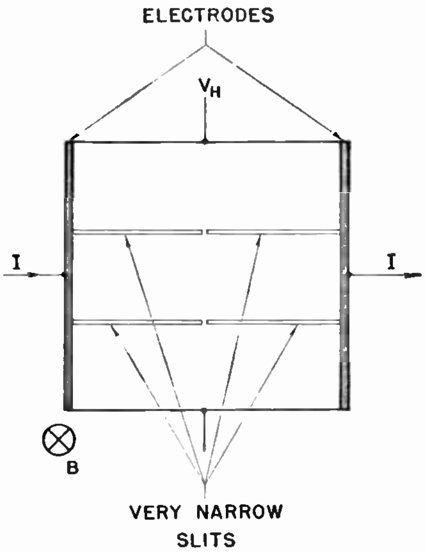


Fig. 2—Multiple Hall-effect sensor configuration.

Although more efficient, this type of sensor would require an increase in the area of the specimen. Thus, such a device becomes less effective in measuring magnetic field intensities at a point. The following is a design of a multiple-element Hall-effect sensor without any increase in its area. Fig. 3 represents a schematic view of such a sensor composed of four elements. In the case described, the connections of the Hall-potential terminals are made in a manner which requires that the currents in successive elements be reversed. When a single source for all the currents into each element is used, any asymmetry in the positioning of the Hall-potential terminals or in the structure of each element may cause a current flow through these terminals and affect the output. Resistors are inserted in series with each element to assure uniform voltage distribution. Construction of uniform elements with properly positioned terminals may eliminate the need for the series resistors and the associated loss of power dissipated in them. In the case of magnetostatic fields, the use of alternating currents, as shown, facilitates measurement.

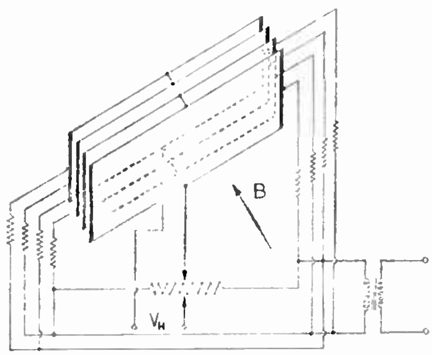


Fig. 3—Schematic view of a four-element Hall-effect sensor.

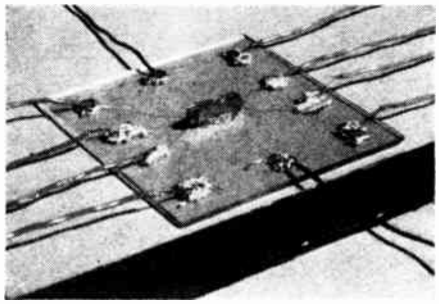


Fig. 4—Four-element Hall-effect sensor mounted on a glass slide.

Fig. 3 represents the experimental arrangement used in measuring static fields. In the experiment performed, each element was made of a sheet of indium antimonide $\frac{3}{8}$ inch$\times\frac{1}{8}$ inch$\times$0.01 inch separated by 0.005 inches, resulting in an over-all thickness of 0.055 inch. The series resistors were 3.9 ohms each. The potentiometer at the output terminal was used to balance out any IR drop due to possible asymmetry in the position of the Hall-potential terminals. The measured field was 8 oersteds and the current in each element was 100 ma at 500 cps. The output for one element was 60 μvolts. Successive additions of elements rendered outputs of 120, 180, and 240 μvolts, respectively. Thus, addition of elements in series resulted in a direct increase in output.

Fig. 4 shows a photograph of the multipole probe.

It should be pointed out that a similar arrangement can be used to measure alternating magnetic fields in which the biasing current is supplied by a dc source such as batteries. However, in measuring alternating fields it is essential that any stray pickup be entirely eliminated by proper arrangement of the associated wiring. Otherwise, any reading of the Hall-potential could not be separated from the stray pickup since respective harmonics of each signal are 90° out of phase with each other.

M. Epstein
H. M. Sachs
L. J. Greenstein
Armour Res. Foundation
Illinois Inst. of Technology
Chicago 16, Ill.

* Received by the IRE, May 1, 1959; revised manuscript received, June 18, 1959.

[1] W. Shockley, "Electrons and Holes in Semiconductors," D. Van Nostrand Co., Inc., New York, N. Y., pp. 211–217; 1950.

[2] H. J. Lippmann and F. Kuhrt, "Der Geometrieeinfluss auf den Hall-Effekt bei rechteckigen Halbleiterplatten," *Z. Naturforsch.*, vol. 13a, pp. 474–483; June, 1958.

[3] I. Isenberg, B. R. Russell, and R. F. Greene, "Improved method for measuring Hall coefficient," *Rev. Sci. Instr.*, vol. 19, pp. 685–688; October, 1948.

[4] E. W. Saker, F. A. Cunnell, and J. T. Edmond, "Indium antimonide as a fluxmeter material," *Brit. J. Appl. Phys.*, vol. 6, pp. 217–220; June, 1955.

[5] M. Epstein and R. B. Schulz, "Magnetic-field pickup for low-frequency radio-interference measuring sets," presented at IRE Natl. Convention, New York, N. Y., March, 1959; to be published in 1959 IRE National Convention Record.

Experimental *X*-Band Preamplifier Tubes with 4.5 DB Noise Figure*

Two years ago it was discovered experimentally that a first-order reduction in the noise figure of beam-type microwave amplifiers can be effected by establishing a unique type of electron flow in the immediate vicinity of the cathode.[1] Guns utilizing this type of flow have thus far led to noise figures less than 3 db at *S* band.[2-4] An experimental investigation of the frequency dependence of the new noise reduction technique has been undertaken in order to provide further insight into the fundamental nature of the mechanism and so that something can be said of the ultimate potential of beam-type amplifiers when comparing them with other type of microwave preamplifiers (*e.g.*, parametric amplifiers).

This paper reports some initial results of this investigation in which reproducible noise figures of 4.5 db and less have been attained at *X* band. These results came about as a direct application of the concepts arrived at in the original experiments at *S* band. Although they cannot be explained fully within the framework of present noise theories, calculations in this and other laboratories are beginning to provide some insight into the basic mechanism of this type of low-noise electron gun and to offer the hope of still further noise reduction using these general techniques.

The type of electron flow used in these experiments, which has been described elsewhere,[2] involves injection of an expanding beam from the edge of the cathode into a specially shaped electric field configuration; the space charge, confined to a thin annulus by a large magnetic field, depresses the potential and allows the beam to be accelerated gradually through an extended region in which the thermal velocity spread is comparable to the mean velocity.

The design of an electron gun to achieve the above type of electron flow involves the following principal features. First, the electrode and field configurations are arranged so that appreciable emission occurs only in a narrow region at the cathode edge, where the electric field has a large component normal to the direction of flow (*i.e.*, the direction of the magnetic confining field) and consequently tends to expand the beam. Second, the off-cathode potential gradient at the cathode edge must be sufficiently large to cause injection of relatively high current density (probably almost temperature-limited) into the divergent field region. "Linear edge current" (current per unit length of available cathode edge for a beam of about 0.001-inch radial thickness) is a useful parameter which measures the extent to which these empirical conditions are satisfied; it also can be used as a basis for scaling this type of low-noise gun.

That the above features are important in minimizing beam noisiness is demonstrated in the evolution of the low-noise gun used in the *X*-band experiments. Fig. 1 shows the steps in this development and, as will be discussed, how the final performance was realized as a result of optimizing the potential distribution at the cathode edge.

The particular type of tube used in these experiments was a backward-wave amplifier. It was scaled directly from the *S*-band amplifier previously described[5] by maintaining the linear edge current constant at nominally 0.22 ma per centimeter and using the same helix diameter relative to wavelength. The resulting beam diameter is 0.120 inch.

Three types of cathode geometries were utilized, namely, a flat button cathode, a "stepped cathode," and an annular cathode. The simplest is the conventional button cathode with electrodes (not shown) arranged to produce the kind of potential contour shown in Fig. 1(a). Although the first tubes using this geometry had noise figures as low as 7 db, it was evident that the electric field at the cathode was not a sufficiently strong function of radius to produce a sharply defined edge emission. That is, it was not possible to increase the edge emission to saturation without simultaneously drawing a relatively large space-charge-limited (and noisy) current from the cathode face.

Fig. 1(b) illustrates the "stepped cathode" geometry. This new type of cathode was developed primarily as the result of an electrolytic tank investigation in an effort to achieve independent control of cathode edge and face emission. It is evident that the cathode post tilts the equipotential contours so as to increase the field at the edge and, at the same time, reduce the field away from the edge. The cathode coating is restricted to the lower step.

Several tubes were constructed using the stepped cathode and produced minimum noise figures of 6 db. However, it was apparent that the gun geometry had not been optimized, *i.e.*, the edge field had not been increased sufficiently to essentially eliminate space-charge-limited current away from the edge for a given total beam current, and the potential profile for the edge electrons was not ideal.[2] It is believed that the stepped cathode holds great promise particularly for forward-wave amplifiers where the RF performance is not such a critical function of current and beam location as it is in the backward-wave case.

Finally, an annular ring cathode was constructed with a simplified "Christmas tree" structure,[5] similar to that used in the *S*-band experiments, which supports electrodes inside the hollow beam. As indicated in Fig. 1(c), this results in tight control of the field at the cathode edge. The potential gradient at that point is more symmetrical about the edge and can be increased while minimizing the face current. Although a second beam now originates from the inner edge of the cathode, this beam interacts negligibly with the RF fields of the helix which decay exponentially away from the helix wires.

Several tubes utilizing the "Christmas tree" gun produced minimum noise figures of 4.5 db at 8900 mc with gains in excess of 20 db. The lowest measured noise figure (two tubes) was 4.3 db, as measured by several different methods.

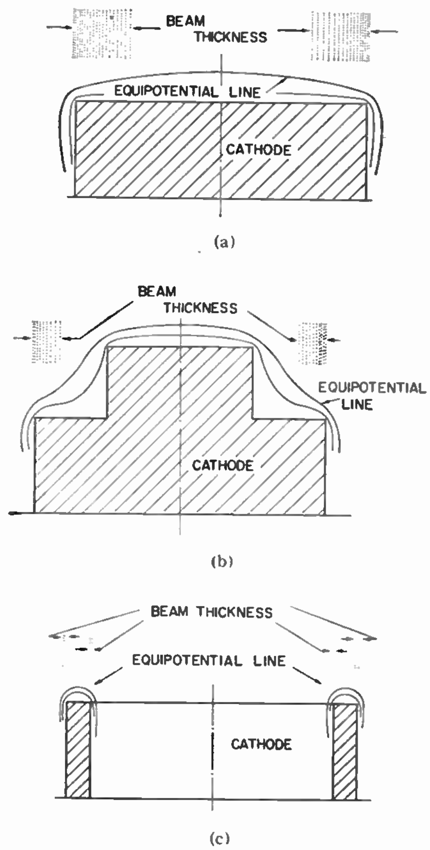


Fig. 1.

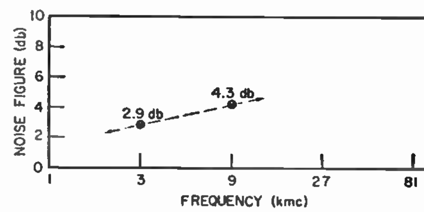


Fig. 2.

It should be noted that it is probable that these results can be improved since the many parameters known to be important such as gun-circuit separation, circuit loss, etc., have not been optimized. It is also emphasized that the same general techniques should produce similar results in the broadband foward-wave amplifiers.

These results, when coupled with the previous results at *S* band,[2] present an interesting basis for speculation concerning attainable noise performance of traveling-wave tubes at still higher frequencies. This is illustrated in Fig. 2. Although a straight-line extrapolation of existing measurements might be optimistic, it is clear that pushing these results toward the millimeter-wave region would constitute an interesting and possibly very fruitful area of research. Although problems of circuit loss and available cathode edge became increasingly difficult as fre-

* Received by the IRE, April 13, 1959.

[1] M. R. Currie and D. C. Forster, "New Results on Noise in Electron Beams," presented at the Fifteenth Annual Electron Tube Conf., Berkeley, Calif.; 1957.

[2] M. R. Currie and D. C. Forster, "New mechanism of noise reduction in electron beams," *J. Appl. Phys.*, vol. 30, pp. 94–103; January, 1959.

[3] M. R. Currie, "A new type of low-noise electron gun for microwave tubes," Proc. IRE, vol. 46, pp. 911–912; May, 1958.

[4] G. E. St. John and M. Coulton, "*S*-band traveling-wave tubes with noise figure below 4 db," Proc. IRE, vol. 46, p. 911; May, 1958.

[5] M. R. Currie and D. C. Forster, "Low noise tunable preamplifiers for microwave receivers," Proc. IRE, vol. 46, pp. 570–579; March, 1958.

quency goes up, calculations[6] indicate that beam noisiness as a function of frequency can be compensated for using the general techniques described here.

In summary, we feel that it would be unfortunate indeed if, in the rush to parametric devices and masers, exploration of the fundamental noise limitations of slow space-charge wave amplifiers were neglected. Certainly, the present results will be improved upon. This, together with the flexible RF characteristics of slow-wave amplifiers, will insure that they continue to occupy a prominent position in the array of low-noise microwave devices.

J. E. Nevins
M. R. Currie
Electron Dynamics Dept.
Hughes Res. Labs.
Culver City, Calif.

[6] W. M. Mueller, Hughes Res. Labs., Culver City, Calif., private communication.

Traveling-Wave Tube Equations Including the Effects of Parametric Pumping*

The purpose of this letter is to disclose a very general set of traveling-wave tube equations. These equations were derived in order to investigate the process of parametric amplification of space-charge waves in a confined flow TWT electron beam of finite cross section.

The derivation of the equations is based on the usual TWT model of (11), (41), and (44) of Louisell and Quate.[1] The symbols used in these three equations retain the same meaning here.

Let ω_1 be the applied signal frequency and ω be the frequency of the pump wave applied by a cw local oscillator. Then, due to the nonlinear nature of the electron beam, an idler at frequency $\omega_2=\omega-\omega_1$ will be generated as well as frequencies $\Omega_{lm}=l\omega_1+m\omega_2 \equiv -\Omega_{-l,-m}$ where l, m are zero, positive, or negative integers. The characteristic impedance of the line is frequency dependent, and at frequency Ω_{lm}, it may be written as

$$K_{lm}=\left\{\sqrt{L_1/C_1}\right\}_{lm} \overset{\Delta}{=} 4C_{lm}{}^3V_0/I_0 \equiv K_{-l,-m},$$

where we define $4C_{lm}{}^3$ as the ratio of the helix impedance at frequency Ω_{lm} in the absence of the beam to the dc beam impedance.

Next we define a parameter

$$\alpha_{lm}=(\omega_q)_{lm}/\Omega_{lm}=R(\gamma_{lm}b)\omega_p/\Omega_{lm}$$
$$\equiv \alpha_{-l,-m} \overset{\Delta}{=} C_{lm}\left\{\sqrt{4QC}\right\}_{lm},$$

where R is the plasma reduction factor,[2] $\gamma_{lm}=\Omega_{lm}/u_0$, ω_p is the plasma frequency of an infinite beam, b is the beam radius, and $4QC$ is Pierce's space-charge parameter.[3]

The phase velocity of the fast space-charge wave on the beam in the absence of the helix at frequency Ω_{lm} is $u_0/(1-\alpha_{lm})$ where α_{lm} is given above. The phase velocity of a wave on the helix in the absence of the beam at frequency Ω_{lm} is given by

$$\left\{\sqrt{L_1C_1}\right\}_{lm}{}^{-1}=u_0/(1-\xi_{lm}),$$

where ξ_{lm} measures the helix phase velocity deviation from the dc beam velocity at frequency Ω_{lm}, and is $(-Cb)_{lm}$ in Pierce's notation.[3]

and its harmonics. We assume that the amplitudes of these waves are small compared with the dc. The nondiagonal terms for which $l \neq m$ are the ac signal and idler terms. These are, in turn, assumed small compared with the pump and pump harmonic terms.

Substituting the expansions similar to (1) into (11), (41) and (44) of Louisell and Quate,[1] dropping products of pump and pump harmonic terms compared with the dc as well as products of ac terms, and further making the assumption that all the α_{lm}'s are small compared with unity, the coefficients of $\exp[j\Omega_{ab}t]$ in these equations reduce to

$$\left\{\frac{j}{\gamma_{ab}}\frac{d}{dz}+\xi_{ab}-\alpha_{11}'\right\}\Phi_{ab}=-\kappa_{ab}(\sigma_{ab}-\rho_{ab}) \qquad (2)$$

$$\left\{\frac{j}{\gamma_{ab}}\frac{d}{dz}+2-(\alpha_{11}'+\xi_{ab})\right\}\Psi_{ab}=\kappa_{ab}(\sigma_{ab}+\rho_{ab}) \qquad (3)$$

$$\left\{\frac{j}{\gamma_{ab}}\frac{d}{dz}-(\alpha_{11}'\pm\alpha_{ab})\right\}\left\{\begin{matrix}\sigma_{ab}\\ \rho_{ab}\end{matrix}\right\}$$
$$=\kappa_{ab}(\Phi_{ab}+\Psi_{ab})+\frac{1}{2}\sum_{l=-\infty}^{+\infty}{}'\,\alpha_{ll}'u_{ll}\sqrt{\frac{\alpha_{a-l,b-l}}{\alpha_{ab}}}\left\{1\mp\frac{\alpha_{ab}}{\alpha_{ll}'}\pm\frac{\alpha_{ab}}{\alpha_{a-l,b-l}}\right\}\sigma_{a-l,b-l}$$
$$+\frac{1}{2}\sum_{l=-\infty}^{+\infty}{}'\,\alpha_{ll}'u_{ll}\sqrt{\frac{\alpha_{a-l,b-l}}{\alpha_{ab}}}\left\{1\mp\frac{\alpha_{ab}}{\alpha_{ll}'}\mp\frac{\alpha_{ab}}{\alpha_{a-l,b-l}}\right\}\rho_{a-l,b-l} \qquad (4)$$

We now expand the helix voltage and current and the beam velocity and current in a double Fourier series of the form

$$\frac{V_c}{V_0}=\sum_{l=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}V_{lm}(z)\cdot\exp\left\{j\Omega_{lm}[t-(1-\alpha_{11}')z/u_0]\right\}, \qquad (1)$$

with similar expansions for I_c/I_0, v/u_0 and i/I_0. A common wave has been factored out at each frequency whose phase velocity is given by $v=u_0/(1-\alpha_{11}')$ where $\alpha_{11}' \ll 1$. This wave will be taken as any normal mode on the system at the pump frequency. For example, if there is no coupling between the helix and beam at the pump frequency, ω, then $\alpha'_{11}=\pm\alpha_{11}=\pm\omega_{q11}$. That is, v may be the phase velocity of either normal space-charge mode on the beam at the pump frequency, ω. If coupling is allowed between the beam and helix at ω, then v may represent the phase velocity of any one of the three normal modes of the TWT. It is assumed, however, that the pump wave propagates in only one mode, and that only one mode of each pump harmonic frequency wave is excited. Factoring out this pump wave assures equations with constant coefficients.

Since V_c, I_c, v and i are real, we require that the expansion coefficients in (1) satisfy $V_{lm}(z)=V^*_{-l,-m}(z)$ with similar relations for the other expansion coefficients. Thus, it is in order to use complex notation that we have included negative values of l and m. Requiring that α_{lm} and K_{lm} be real results in the relations $\alpha_{lm}=\alpha_{-l,-m}$ and $K_{lm}=K_{-l,-m}$.

Refer to the terms in the Fourier expansions. The term for which $l=m=0$ is the dc term in each case. They are normalized so that $V_{00}=I_{00}=0$, $u_{00}=1$, and $i_{00}=-1$. Other diagonal terms for which $l=m\neq 0$ corresponds to frequencies $l(\omega_1+\omega_2)=l\omega$, which identifies terms associated with the pump

where the prime on the sums indicates that the term $l=0$ is to be omitted, and $a\neq b$. Further, $\kappa_{ab}=\sqrt{C_{ab}{}^3/2\alpha_{ab}}=C_{ab}(QC)_{ab}{}^{-1/4}/2$, and $\alpha_{ll}'=-jC_{ll}\delta_{ll}$ where δ_{ll} is one of the TWT propagation constants at frequency $\Omega_{ll}=l\omega$. The mode amplitudes are defined by

$$\left\{\begin{matrix}\Phi_{ab}\\ \Psi_{ab}\end{matrix}\right\}=\frac{1}{2}\sqrt{\frac{V_0I_0}{8C_{ab}{}^3}}\left[V_{ab}\pm 4C_{ab}{}^3I_{ab}\right]$$

$$\left\{\begin{matrix}\rho_{ab}\\ \sigma_{ab}\end{matrix}\right\}=\frac{1}{2}\sqrt{\frac{V_0I_0}{2\alpha_{ab}}}\left[u_{ab}\mp\alpha_{ab}i_{ab}\right]. \qquad (5)$$

Φ_{ab}, Ψ_{ab}, σ_{ab}, and ρ_{ab}, are the amplitudes of the normal forward and backward transmission line modes, and slow and fast space-charge modes, respectively, in the absence of coupling. They are defined so that the power in these waves is the mode amplitude squared. The left-hand sides of (2)–(4) are simply the wave equations for these waves with the arbitrary pump wave factored out. The right-hand terms show how these waves are coupled. Any or all frequencies can propagate on the circuit by choosing various coupling coefficients zero.

As we have said, these are very general equations. Only a segment of them can or need be considered for any one problem. In nearly all cases, for instance, (3) may be decoupled from the others, since the backward transmission line wave is ordinarily unsynchronized with the other waves to such an extent as to render the effect of coupling inconsequential. Also, for most parametric problems, the slow space-charge wave may be decoupled.

We have studied several problems using these equations. The results of our studies are too lengthy to record here, but they will appear presently in published form, as will the derivation of (2)–(4).

J. S. Cook
W. H. Louisell
Bell Telephone Labs., Inc.
Murray Hill, N. J.

* Received by the IRE, July 10, 1959.

[1] W. H. Louisell and C. F. Quate, "Parametric amplification of space-charge waves," Proc. IRE, vol. 46, pp. 707–716; April, 1958.

[2] G. M. Branch and T. G. Mihran, "Plasma reduction factors in electron beam," IRE Trans. on Electron Devices, vol. ED-2, pp. 3–11; April, 1955.

[3] J. R. Pierce, "Traveling Wave Tubes," D. Van Nostrand Co., Inc., New York, N. Y.; 1949.

Phase Relationships in Short-Slot Hybrid Couplers*

It is well known that a signal fed into one of the ports of a symmetrical short-slot hybrid coupler results in output signals which differ in phase by 90°. However, the correct lead or lag relationship between these output signals has not been commonly known and has only recently been brought to light by workers at the Bell Telephone Laboratories[1] and Naval Research Laboratories.[2] Their results are in complete agreement with a parallel investigation made at the Westinghouse Air Arm Division. Our investigation included an analysis and two experiments very similar to those described by Tompkins,[2] giving indisputable confirmation of the phase relationships predicted by the analysis.

The purpose of this letter is to present a somewhat more complete picture of the phase relationships in short-slot hybrids, in the form of the phasor diagrams shown in Fig. 1. In these diagrams, the phases of the output signals E_3 and E_4 are referred to the phase of a signal $E_3°$ which would exist at the main guide output terminal (port 3) if there were zero coupling to the auxiliary guide, *i.e.*, if the main guide had no slot in it. This reference phasor is shown dashed because it does not actually exist when there is finite coupling.

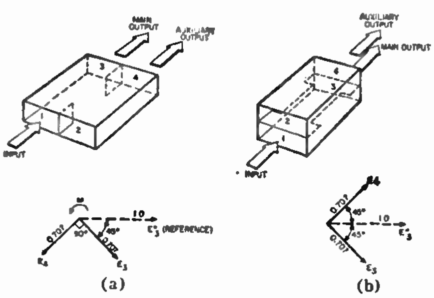


Fig. 1—Short-slot hybrid phasor diagrams; (a) sidewall hybrid, (b) topwall hybrid.

Fig. 2—Phasor diagrams for 10-db short-slot couplers; (a) sidewall coupler, (b) topwall coupler.

It may be noted that the main guide output signal E_3 lags the reference signal $E_3°$ by 45° in both the sidewall hybrid and the topwall hybrid. However, as stated by McCabe, *et al.*,[1] and Tompkins,[2] the auxiliary guide output signal E_4 *lags* the main branch signal E_3 by 90° in the sidewall hybrid, and *leads* by 90° in the topwall hybrid.

The short-slot hybrids are essentially directional couplers with 3-db coupling. The same analysis which led to the above phase relationships can be generalized for short-slot couplers having other than 3-db coupling values. For example, the phasor diagrams for 10-db short-slot couplers are as shown in Fig. 2. In this case, the main branch output signal E_3 has an amplitude of 0.948 relative to the reference signal $E_3°$ (amplitude 1.0) and a phase lag of $\alpha = 18.43°$ in both types of coupler. The auxiliary branch output E_4 still bears the same 90° phase relationships to E_3, *i.e.*, lagging in sidewall coupler and leading in the topwall, but is reduced in amplitude to 0.316.

It is apparent that as the fractional coupling approaches zero (infinite db), the main arm output E_3 approaches the reference $E_3°$ in both phase and amplitude, while the auxiliary signal E_4 diminishes to zero amplitude but always maintains its 90° phase relations with respect to E_3.

H. E. SCHRANK
C. H. GRAULING, JR.
Air Arm Div.
Westinghouse Electric Corp.
Baltimore, Md.

* Received by the IRE, June 30, 1959.

[1] B. McCabe, E. J. Schwesinger, and M. J. Traube, "Phase relationships in short-slot hybrid couplers," Proc. IRE, vol. 47, p. 1004; May, 1959.

[2] R. D. Tompkins, "Phase relationships at the output of a short-slot hybrid junction," Proc. IRE, vol. 47, p. 1156; June, 1959.

E-Type *X* Circulator*

An *X* circulator consisting of the *H*-type four-port waveguide junction containing a nonreciprocal element (a properly magnetized ferrite) was proposed and its experimental results in *X* band was discussed by the author.[1]

In this paper the author would like mainly to propose a new type of *X* circulator (Fig. 1) which is simple in structure, small in volume, light in weight, and which has good characteristics. He would also like to discuss the experimental results of the proposed circulator in *X* band.

A ferrite (or garnet) element of the proper material, shape, and size, for the characteristics of a circulator is inserted with the right impedance element at the exact place in the usual *E*-type four-port waveguide junction, in which four standard rectangular waveguides in *X* band (WR-90)

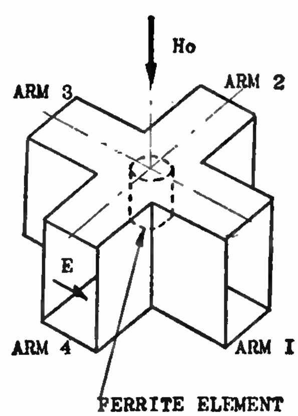


Fig. 1.

* Received by the IRE, April 6, 1959.

[1] S. Yoshida, "*X* circulator," Proc. IRE, vol. 47, p. 1150; June, 1959.

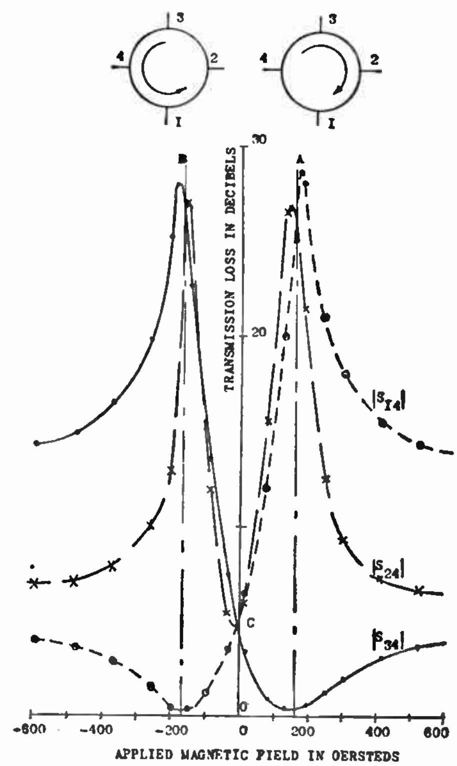


Fig. 2.

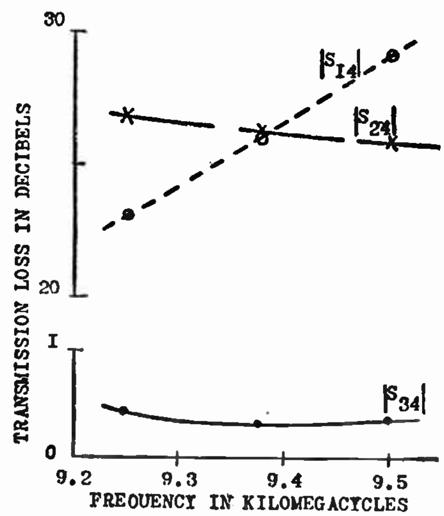


Fig. 3.

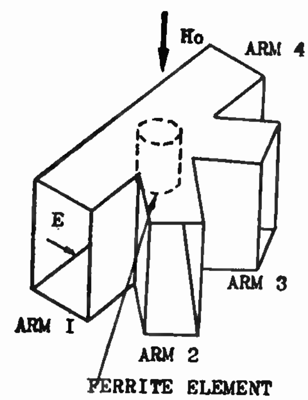


Fig. 4.

intersect at right angles at one point, and a dc external magnetic field is applied to the ferrite element transverse to the narrow sides of this waveguide junction by an external electromagnet.

When the wave of the frequency 9375 mc enters, for example, through arm 4 from a matched generator, this wave 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 *Ho* corresponding to point *A* or *B* in this figure is applied, the clockwise or counter-clockwise circulator is made up, respectively. In practice barium ferrites having about 160 oersteds are set to this junction. The frequency characteristics of the circulator corresponding to point *A* in this figure are shown in Fig. 3. In the above-mentioned example, a rod of manganese magnesium ferrite is placed on the upper narrow wall at the center of the junction as the nonreciprocal element, with the impedance element which is needed to improve the characteristics of the circulator (Fig. 1). When the wave enters from any other arm, the previously-mentioned characteristics are kept almost the same.

When the external dc magnetic field *Ho* corresponding to point *C* in Fig. 2 is applied, a *E*-type four-port matched equal power divider of nonreciprocal circuit is made up successfully in practice.

In the case of the *E*-type four-port waveguide junction composed of four standard rectangular waveguides in *X* band, two of which intersect at 180°, and three of which intersect at 60°, a circulator (Fig. 4) is made up successfully by using almost the same technique, but its characteristics are inferior to the *E*-type *X* circulator at present.

SHINICHIRO YOSHIDA
Matsuda Res. Lab.
Tokyo Shibaura Electric Co.
Kawasaki, Kanagawa-Ken, Japan.

An *E*-Type *T* Circulator*

Many types of *E*-type *T* circulators have been investigated by Fox, Miller, and Weiss.[1]

In this paper the author would like to propose an *E*-type *T* circulator consisting of the *usual* *E*-plane *T* junction containing a nonreciprocal element (a properly magnetized ferrite) shown in Fig. 1; this is simple in structure, small in volume, light in weight, and has good characteristics; he would also like to discuss the experimental results of the proposed circulator in *X* band.

A ferrite (or garnet) element of the proper material, shape, and size for the characteristics of a circulator is inserted with a correct impedance element at the right place in the *usual* *E*-plane *T* junction of rectangular waveguide in *X* band (WR-90), and a dc external magnetic field is applied to this ferrite element, almost transverse to the narrow sides of this waveguide junction by an external electromagnet.

When the wave of the frequency 9375 mc enters, for example, through arm 3 from a matched generator, it comes out from arms 1 and 2 terminated by matched detectors. Transmission losses are shown against the applied dc magnetic field in Fig. 2. When the external dc magnetic field Ho corresponding to point *A* or *B* in this figure is applied, the clockwise or counter-clockwise circulator is completed, respectively. In practice, barium ferrites having the magnetic field Ho corresponding to the point *A* are set to this junction; the frequency characteristics of the circulator are shown in Fig. 3. In this example, a rod of manganese magnesium ferrite is placed on the lower narrow wall at about the center of the junction as the nonreciprocal element, with the impedance element which is needed to improve the characteristics of the circulator. In the case of the *E*-plane *Y* junction, another type of circulator (*E*-type *Y* circulator) is made successfully by using almost the same technique, and its characteristics are slightly superior to the *E*-type *T* circulator, especially with regard to the point of the bandwidth. A microwave switch can be made by applying the magnetic field corresponding to the point *A* and *B* in Fig. 1 alternately to the ferrite element.

SHINICHIRO YOSHIDA
Matsuda Res. Lab.,
Tokyo Shibaura Electric Co.,
Kawasaki, Kanagawa-Ken, Japan.

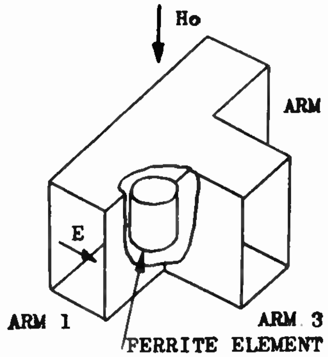


Fig. 1.

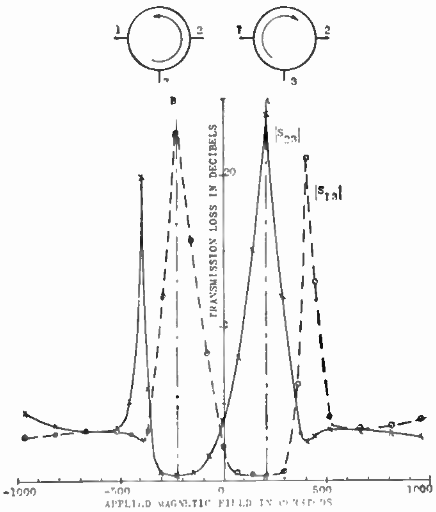


Fig. 2.

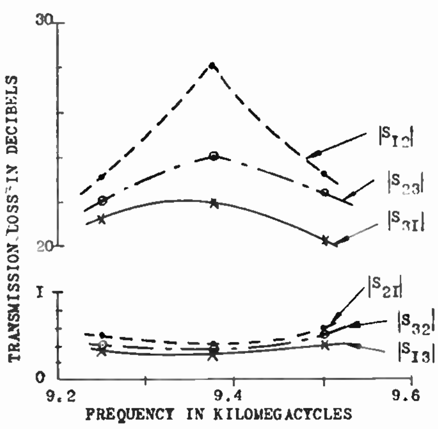


Fig. 3.

* Received by the IRE, April 28, 1959.

[1] A. G. Fox, S. E. Miller, and M. T. Weiss, "Behavior and application of ferrites in the microwave region," *Bell Sys. Tech. J.*, vol. 34, pp. 5–103; January, 1955.

Crystal Noise Effects on Zero IF Receiver*

The principle of a zero intermediate-frequency receiver[1] has been recently applied successfully to a pulsed Doppler radar at *X* band.

The development of this system was undertaken at the Research Laboratories Division, Bendix Aviation Corporation, to utilize the Doppler principle without the inherent limitations or complexity of a CW system.

The first concern in this effort was the extent of crystal mixer noise at low frequencies. Measurements of crystal noise figures were made with the results shown in Fig. 1. These data confirm the results of Greene and Lyons. It can be seen that only a slight improvement can be realized in crystal noise figure at frequencies beyond approximately 100 kc.

The method employing in making these measurements is illustrated in Fig. 2. Local oscillator noise was minimized by the use of a balanced mixer. In this application a common source was used for both the signal and local oscillator. A simulated Doppler signal was obtained by serrodyning a traveling-wave tube. In this way the local oscillator and signal remained essentially constant in frequency with respect to each other. This eliminated any noise that would result from frequency shifts of the local oscillator.

Measurements were also made with an argon noise tube and a close correlation was obtained between these results and those obtained by the single frequency method.

* Received by the IRE, April 17, 1959.

[1] J. C. Greene and J. F. Lyons, "Receivers with zero intermediate frequency," PROC. IRE, vol. 47, pp. 335–336; February, 1959.

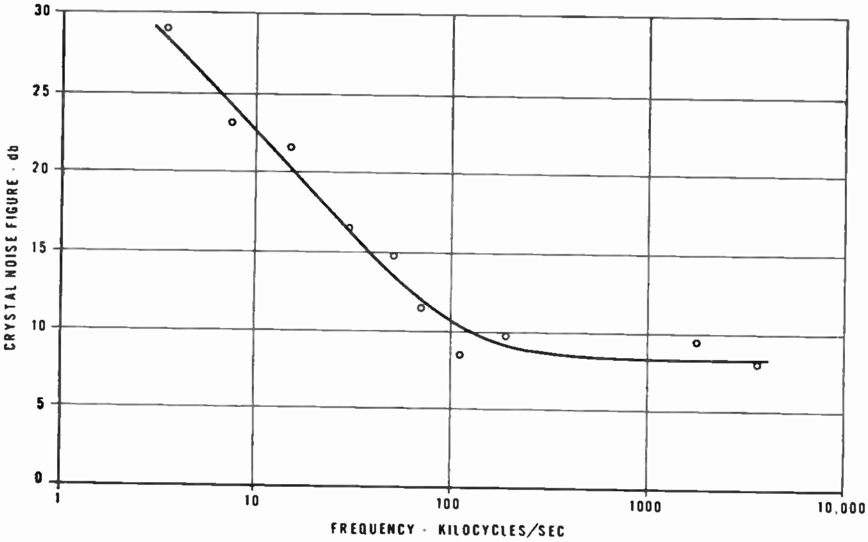


Fig. 1.

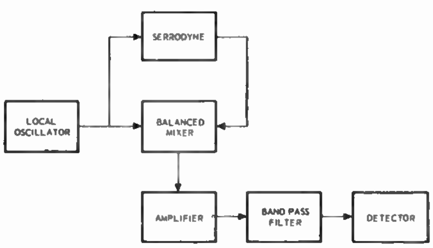


Fig. 2.

The curve shown in Fig. 1 was obtained for 1N23-E crystals, which proved best at this frequency. Several other crystals were tested and although all were noisier than the 1N23-E's, they displayed similar characteristics in that the noise became essentially constant beyond approximately 100 kc.

The data were corrected for the noise contributed by the amplifier used in the measurements, so the curve in Fig. 1 illustrates the noise figure of the crystals and local oscillator only.

In this application (pulsed Doppler) the required information is contained in each Doppler side band, so it was possible to use a narrow-band receiver in a frequency region outside the high-noise region of the crystals. In this system the receiver had a 700-cycle bandwidth centered at 120 kc. This design provided a sensitivity of -111 dbm.

As mentioned by Greene and Lyons, the use of zero IF precludes discrimination between the upper and lower side bands of the received signal, since the spectrum is folded over about zero frequency. This prevents distinguishing between opening and closing Dopplers; however, in this application this was not necessary.

Narrow-band reception also destroys pulse shape, of course, so range information cannot be obtained by utilizing pulse position in time. Range information can be obtained, however, by using other techniques.

G. B. Andrews
H. A. Bazydlo
Bendix Aviation Corp.
Research Labs. Div.
Southfield, Mich.

Satellite Communication*

The Pierce and Kompfner article[1] is no doubt only the first of many that will soon be published to demonstrate the tremendous capability of satellites for world and space communication.

Pierce and Kompfner, taking into consideration only galactic and water-vapor antenna noise, have suggested that the suitable operating frequency lies between 2000 and 8000 mc. In a more generalized article,[2] the writer has included other factors that are of equal or greater importance in determining this frequency: sunburst noise, receiver noise, bad weather (fog and rain) turbulence effects and attenuation, and the design criteria of transmitter and receiver antennas that are *area*-limited (*i.e.*, specified), *gain*-limited, or mixed (*i.e.*, one of each).

A general conclusion is reached that: for communication within or through the earth's atmosphere the optimum frequency (for least satellite transmitter power) is between 150 and 7000 mc; and for free space the frequency should be the highest possible that can be generated coherently. The 150-mc optimum figure is for a 1000°K receiver (present-day performance) and *gain*-limited antennas. The 7000-mc optimum figure is for a 1000°K receiver and *area*-limited antennas.

The Pierce and Kompfner example of a passive repeater has area-limited antennas and, substantially, a 0°K receiver, for which the author's article[2] shows the optimum frequency to be 4000 mc—exactly the mean of their two choices. The optimum frequency is not critical, both 2000 and 8000 mc being only 3 db worse than 4000 mc. The Pierce and Kompfner example of an active repeater is the same as the author's criterion of mixed antennas and a 0°K receiver. For this, the optimum frequency is 2000 mc—exactly Pierce and Kompfner's choice.

The author's article[2] agrees with Pierce and Kompfner that the active repeater in a synchronous (*i.e.*, 24-hour-orbit) satellite has the greater communication capability, but goes much further. With the synchronous satellite as a relay, space communication in general (reconnaissance, navigation, detection, etc.) should be realizable out to Mars with small solar-powered, all-semiconductor systems. Specifically, the requirements are for only 1 watt of transmitter power, one-square-meter antennas, a one-square-meter solar cell, and 30°K receiver noise for a signal-to-noise ratio of 30 db. The satellite-to-ground link would be at 3 kmc and could have an information bandwidth of 1 mc; whereas space link must be 100 kmc and be limited to a 10-cycle bandwidth.

Such an active synchronous satellite relay system is functionally illustrated by Fig. 1. It features all-printed-circuit, all-semiconductor components to achieve light weight and small size—possibly using only the surface of a balloon a few meters in diameter. The receivers and transmitters would both be parametric mixers in printed electronic steerable phased-array antennas of a multiplicity of elements. Transistor amplifiers would be supplemented by capacity-diode harmonic generators to achieve milliwatt pump power at microwave frequencies now, and millimeters later. The multiple properly phased up-converters would give the required 1 watt of transmitter power. Noncoherent radiation could be relayed to the ground by conventional semiconductor detectors, amplifiers, and modulators.

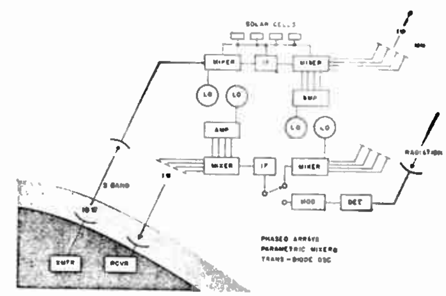


Fig. 1—A proposed all-semiconductor synchronous satellite relay.

The all-semiconductor active satellite is certainly not yet feasible, and solutions to many aspects of the system, such as Doppler shift and pointing control, have not yet been considered; but from an energy viewpoint alone the 300-mc ground link is nearly feasible today.[3,4] The areas in which development is needed are principally microwave printed steerable phased arrays, parametric mixers, and the components of the microwave-to-millimeter semiconductor oscillator-amplifiers and harmonic distorters. Fortunately, the earliest need is for the terrestrial and lunar applications, for which the present art is almost good enough; the

* Received by the IRE, April 29, 1959.

[1] J. R. Pierce and R. Kompfner, "Transoceanic communication by means of satellites," Proc. IRE, vol. 47, pp. 372–380; March, 1959.

[2] C. T. McCoy, "Space Communications," Philco Res. Rept. No. 279; November 15, 1958. Proc. IRE, to be published.

[3] A. Uhlir, Jr., "The potential of semiconductor diodes in high-frequency communications," Proc. IRE, vol. 46, pp. 1099–1115; June, 1958.

[4] C. G. Thornton and J. B. Angell, "Technology of micro-alloy diffused transistors," Proc. IRE, vol. 46, pp. 1166–1176; June, 1958.

less urgent applications of interplanetary communication can afford to wait on the necessary extension of the art to millimeter wavelengths.

C. T. McCoy
Research Division
Philco Corp.
Philadelphia, Pa.

Authors' Comment[5]

McCoy has referred to our paper[1] as no doubt only the first of many which will soon be published to demonstrate the tremendous capability of satellites for world and space communication. While it is certain to be one of many, it is not the first. Clarke proposed 24-hour orbit "stationary" communication satellites as early as 1945,[6] and one of us published concerning communication satellites in 1955.[7] It is also not noteworthy that the idea of an interplanetary relay satellite in a Trojan position was proposed as early as 1942.[8] A rather detailed analysis of the problems of interplanetary and interstellar radio communication was given as early as 1952.[9]

Concerning some technical points in McCoy's paper, our Fig. 3 shows (following Hogg) not only the effective noise temperature due to galactic and water-vapor emission, but also the effect of the oxygen in the atmosphere. It is indeed the latter that sets the limit on what can be achieved with a sensitive receiver, and we doubt whether some of the factors mentioned by McCoy, such as sunburst noise, receiver noise, bad weather (fog and rain), and turbulence effects are anywhere near as important.

A 1000°K receiver is not very good present day performance for either 150 or 7000 mc. At 6000 mc we have attained a noise temperature of under 20°K including antenna loss and noise picked up by sidelobes.

We don't understand the concepts *gain limited* and *area limited.* One can choose various antenna gains or areas as examples but these certainly don't represent limitations. The same may be said about power. Roughly, the attainable power may be expected to vary as the square of the wavelength, and one is free if he likes to assume this in optimizing a system; but it isn't a very accurate assumption at any one time. Thus, broad optima are interesting but not entirely realistic.

One reason for favoring 2000 mc over higher frequencies has to do with tube life and the possibility of transistorization in active repeaters. It also eases tracking problems.

J. R. Pierce
R. Kompfner
Bell Telephone Labs.
Murray Hill, N. J.

[5] Received by the IRE, May 21, 1959.
[6] A. C. Clarke, "Extraterrestrial relays," *Wireless World*, vol. 51, pp. 305–308; October, 1945.
[7] J. R. Pierce, "Orbital radio relays," *Jet Propulsion*, vol. 25, pp. 153–157; April, 1955.
[8] G. O. Smith, "QRM interplanetary," *Astounding Science Fiction*, vol. 30, pp. 109–128; October, 1942.
[9] J. J. Coupling, "Don't write, telegraph," *Astounding Science Fiction*, vol. 49, pp. 82–96; March, 1952.

Letter from Mr. McCoy[10]

The author welcomes the early history, but finds that the "Astounding Scientific Fiction" reference seems to be too astounding or not astounding enough to be carried by any library in Philadelphia according to Union Library Catalogue.

Generalized optimizations certainly must be used with care, but if dominant parameters are chosen with explicit description (which was not feasible in the letter) quantitative curves and conclusions should be more than interesting. The added "factors" which the author still does not consider relatively negligible augment the general conclusions:

1) that external antenna noise is an increasing sensitivity limitation with low microwave frequencies,
2) that terrestrial bases only must contend with a severe limitation from atmosphere absorption, noise, turbulence in the high microwave region, and
3) that low-noise receivers increase system sensitivity significantly only when antenna noise is relatively small.

The author is well aware of the commendable low-noise microwave research of the Bell Telephone Laboratories, and wishes to show any capability by interpolation between numerically rounded 1000°K and 0°K sensitivity limits. The former might better have been called "pre-maser" rather than the ephemeral "present-day" performance.

The antenna "gain" and "area" concepts are conventional and merely specify which parameter is explicit in the transmitter-receiver power-transfer formula. For some applications the gain is a dominant parameter (*e.g.*, a nondirectional requirement), for others it is the area (*e.g.*, a weight, or size limitation). The modifier "limited" can be replaced with "specified" if desired.

C. T. McCoy

[10] Received by the IRE, June 24, 1959.

Unusual Propagation of Satellite Signals*

From time to time letters have appeared in the Proceedings of the IRE which describe a certain unexplained phenomenon involving satellite transmissions.[1] The following is a brief, but slightly incomplete summary of further interesting observations.[2] An explanation is proposed.

When Sputnik I was announced by the Russians, the Electromagnetic Radiation Laboratory set up four 20-mcps interferometers to track it. Pen graphs recorded the intensity of the audio signal vs time, and magnetic tapes recorded the signal frequencies. The data revealed that on four consecutive days (October 18–21) at approximately 1850 hours, a rather surprising propagation phenomenon occurred.

Peculiarities of Experimental Data Pen Recordings

1) The intensity of the signal received when the satellite was between 730 miles and 2000 miles from the receiver was occasionally as great as the signal received during the direct overhead pass.

2) On October 18, the pattern showed regions of peaked intensity when the satellite was approximately 1380 miles away and again when it was 730 miles away. The effect occurred not only as Sputnik approached but also as it departed.

3) Each day the number of peaked regions reduced successively. There was only one on October 21, and on October 22, there was no evidence of the phenomenon.

4) Great irregularities were noted in the peaked regions.

Magnetic Tape Recordings

1) The irregularities in the peaks were found to be due to beat tones produced by the simultaneous reception of up to three closely spaced frequencies in the peaked regions.

2) Near the points where the high intensity suddenly cut off, two of the three frequencies approached the same value.

Partial Summary of Proposed Explanation

A mathematical analysis of the shifts in received frequency indicates that the signals in the peaked regions arrived from Sputnik (located just below the maximum of the only ionospheric layer present—the F layer) via reflections between the earth and ionosphere. The presence of several frequencies was apparently caused by a superposition of reflection modes as well as the presence of "high" and "low" waves due to electron density variation with altitude. The fact that the phenomenon is possible only when the earth reflections are from a very good conductor, such as the ocean, can explain the daily reduction in the number of peaked regions since the orbit was such that on each successive day there was less possibility for ocean reflections. The merging of two frequencies at the end of a peaked region can be explained by the known merging of high and low waves just before a ship zone is reached. The high intensity at the end of a peaked region, and its sudden cutoff is for the most part explained by the focusing effect at the edge of a skip zone and the sudden absence of reception when such a zone is entered.

Lt. Edmond M. Dewan, USAF
Electromagnetic Rad. Lab.
Air Force Cambridge Res. Center
Bedford, Mass.

* Received by the IRE, May 18, 1959.
[1] See for example, J. D. Kraus and J. S. Albus, "A note on some signal characteristics of Sputnik 1," Proc. IRE, vol. 46, pp. 610–611; March, 1958.
See also R. Parthasarathy and G. C. Reid, "Signal strength recordings of the Satellite 1958 δ 2 (Sputnik III) at College, Alaska," Proc. IRE, vol. 47, pp. 78–79; January, 1959.
[2] The complete report of this work appears in the *Proc. Congrès Internatl. sur la Propagation des Ondes Radioelectriques, Liege, Belgium*; October, 1958.

Selective Fading Effects on UHF Tropospheric Scatter Paths*

The application of UHF tropospheric propagation for wide-band transmission beyond the horizon has been widely discussed in the literature of recent years. However, there are few references to actual experimental evidence of the frequency selective nature of the signal fluctuations. It therefore might be of interest to publish some results of tests that had been done during the winter 1954/1955 on one of the early military scatter systems in the Labrador area.[1] Lacking any practical information at that time on questions such as usable bandwidth, multipath distortion, etc., the tests were intended to throw some light on the frequency selectivity of the ever-present signal fluctuations in scatter propagation. The results were by no means comprehensive and certainly, from a scientific point of view, rather qualitative. However, they served their purpose in demonstrating what can happen under normal circumstances, and where the limitations of the scatter type of transmission are likely to show up.

The tests were done on two overland links, both 150 statute miles long, and equipped with 60-foot parabolic antennas, operating around 700 mc. Since no difference was found between the two links, no further distinction will be made between them. For the purpose of the test the transmitter was frequency-modulated by a sine-wave signal of 160 kc at deviation ratios between 2 and 5, thus generating side-band spectra of equidistant lines extending to more than 1 mc on either side of the carrier frequency. At the receiving end, the IF signal (around 30 mc) of one of the receivers (bandwidth 4 mc, no diversity operation for this test) before entering the limiter stages was fed to a panoramic analyzer. It was displayed here on a CRT screen with logarithmic calibration. This display was observed and during interesting periods of time photographed by an automatic camera with a frame-repetition rate of 2 per second and an exposure time of $\frac{1}{10}$ of a second. The continuous visual observation of the spectrum showed that at this rate of exposures no significant effects would be lost. The sweep frequency was 20 per second, and the screen brightness and persistency such that successive pictures were practically independent of each other.

Samples of the photographs are shown in Figs. 1–8, selected arbitrarily from a number of complete series. The capital letters refer to the particular series of photographs, and the frame numbers indicate the sequence of frames, beginning with number 1 at the start of the series and then following at a rate of 2 per second. Each series was extended in time until a sufficient number of completely symmetrical spectrum displays was obtained, these being assumed to represent the undistorted side-band distribution under the particular FM deviation conditions. One of the symmetrical displays was then used as the arbitrary 0-db reference level for the transcription into the fading waveforms shown below. The frequency fig-

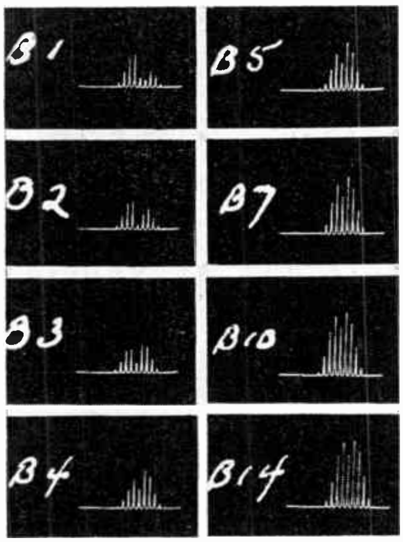


Fig. 1—Samples of spectrum photographs, *B* series.

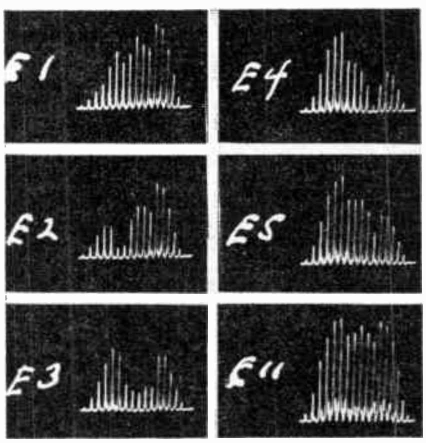


Fig. 2—Samples of spectrum photographs, *E* series.

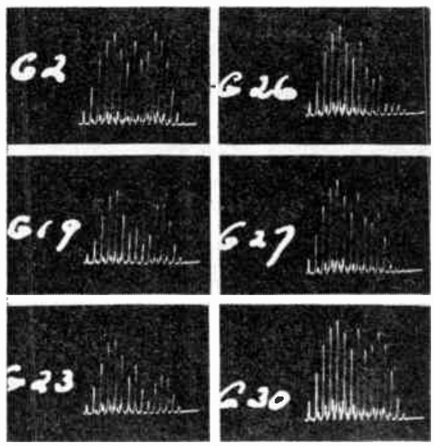


Fig. 3—Samples of spectrum photographs, *G* series.

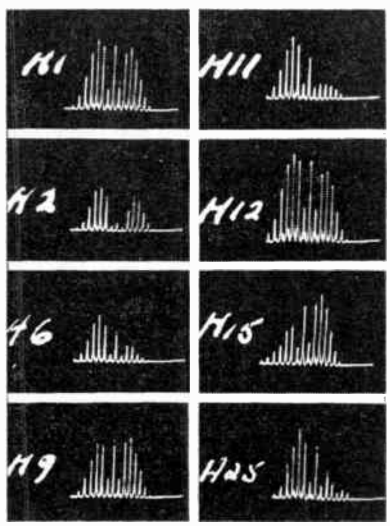


Fig. 4—Samples of spectrum photographs, *H* series.

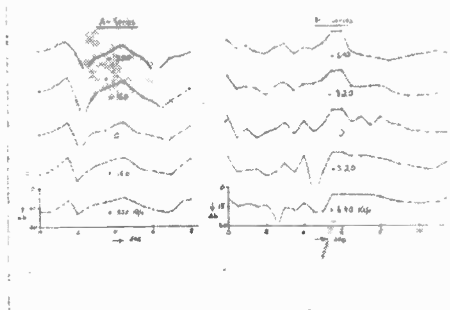


Fig. 5—Waveform of fading, *A* and *H* series.

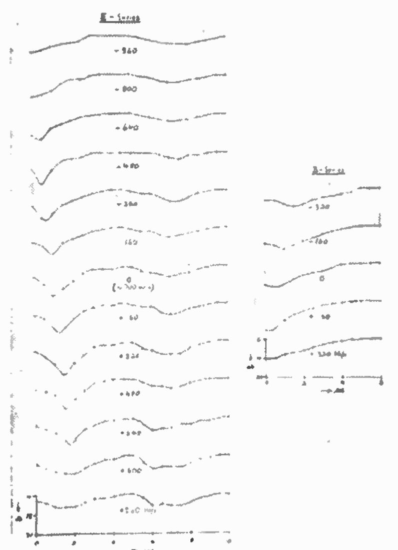


Fig. 6—Waveform of fading, *E* and *B* series.

* Received by the IRE, April 15, 1959. Permission was granted by U. S. Air Force Dept. to publish these notes.

[1] The system was installed for the U. S. Air Force by the Bell Telephone Company of Canada with the Pinetree Project Office in Ottawa, Ontario, as Design Authority. A Pinetree Project Office report, dated July, 1955 (which has since been quoted frequently), described the system and the various tests briefly. However, no further publication has taken place.

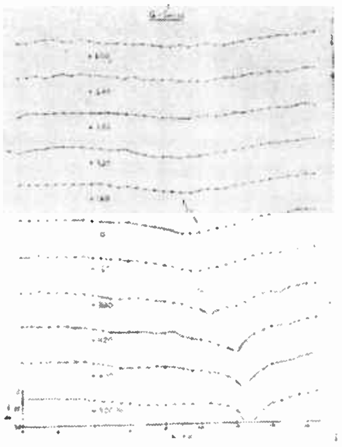


Fig. 7—Waveform of fading, G series.

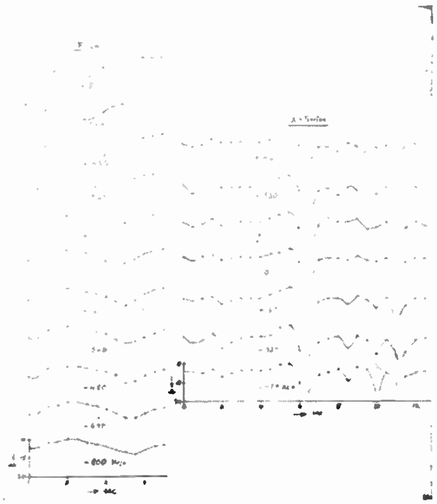


Fig. 8—Waveform of fading, F and X series.

ures on these waveforms indicate the difference from the carrier frequency.

The majority of fading effects had the characteristics of a level depression moving in either direction through the frequency band; only a few had more random character (see, for example, X series). The greatest observed level differences over the frequency band were of the order of 15–20 db per mc lasting for not more than one second.

It was apparent during the tests that the spectacular, selective effects occurred more frequently (at times, at least once every minute) when the median signal strength had dropped to a low value. They became less frequent and less spectacular with increasing median, disappearing completely when the median was very high and the fading rate very slow. However, the tests were broken off before this relationship was further corroborated.

Having demonstrated the selective nature of the signal fluctuations, a number of noise loading tests were performed over the same two links and during the same general period of time. The intention was to study the amount of nonlinear distortion (cross modulation) introduced by the propagation mechanism. The noise modulation of the FM transmitter was confined to the band 40–120 kc and cross-modulation products were measured in the receiver output (without diversity) above and below this band. However, because of the fact that selective effects were generally coupled with low signal-strength values it was not possible to separate any propagation distortion, at low FM deviation ratios, from the thermal-noise contributions. On the other hand, at high FM deviation ratios, the equipment distortion became predominant. There were only a few instances where additional noise could be attributed to propagation distortion. The general conclusion was therefore that under the given systems conditions (*i.e.*, within the modulation band up to 160 kc) no noticeable cross modulation due to scatter propagation was to be expected.

Thanks are extended to F. Gall with whom these tests were conducted during the short time interval between the completion of installation and the commencement of continuous operational use of the system.

H. J. von Baeyer
Telecommun. Div.
RCAF Headquarters
Ottawa, Ont., Can.

Tangent Method Analysis for High-Insertion-VSWR Lossless Two-Ports*

The tangent method of measuring the parameters of lossless two-ports[1] is widely known and has associated with it a variety of analysis procedures[2] applicable to the various ranges of γ and suitable for obtaining various degrees of precision. A relatively direct and accurate means of estimating very high values heretofore, however, has not been available. Such high γ values occur, for example, in the case of small coupling irises and are of practical interest since $|\gamma|$ is identically the insertion VSWR of the structure.

The usual data taken in such a case would, when plotted, look like the solid curve shown in Fig. 1, or, at "half-wave distances" away, like the dashed curve. High values of γ can be readily abstracted if the data required for the small portions of the two curves enclosed in the circle in Fig. 1 is directly measured and plotted to a large scale as in Fig. 2. As shown, the distance g is defined as the distance between two lines drawn tangent to the two sections of the curve making angles of 45° with the axes. It is assumed here that both D and S are plotted to the same scale (units of length)

* Received by the IRE, April 22, 1959. The work reported here was supported by the U. S. Air Force Cambridge Research Center under Contract AF-19(604)-2031.

[1] N. Marcuvitz, "On the representation and measurement of waveguide discontinuities," PROC. IRE, vol. 36, pp. 728–735; June, 1948.

[2] M. Wind and H. Rapaport, "Handbook of Microwave Measurements," Microwave Res. Inst., Polytechnic Inst., Brooklyn, N. Y., sec. 6, by L. B. Felsen; 1954.

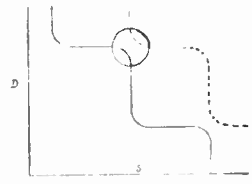


Fig. 1.

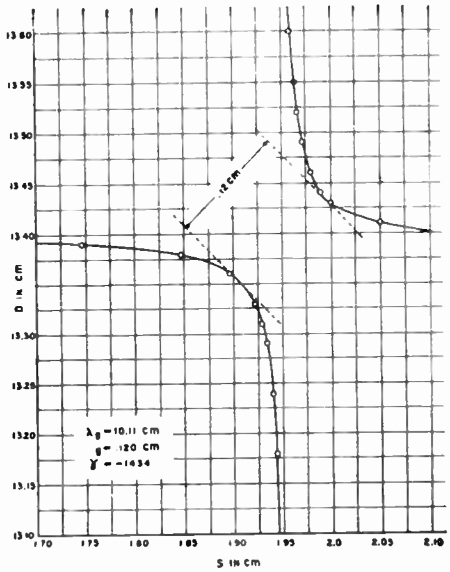


Fig. 2.

and that g is measured with this scale. γ is now given exactly by

$$\gamma = -\cot^2\left(\frac{g\pi}{\sqrt{2}\lambda_g}\right).$$

The corresponding approximate (small angle) formula

$$\gamma \simeq -2\left(\frac{\lambda_g}{g\pi}\right)^2$$

is more convenient and is in error by only 3 per cent for values of $|\gamma|$ as small as 20. This error, of course, decreases rapidly for larger values. If, as is often done, D and S are normalized to λ_g, then g will be obtained from the curve in the same normalized units and λ_g must be taken as unity in the above formulas. The present analysis is inherently more accurate than other schemes now available for the range $|\gamma| > 6$.

The curve shown in Fig. 2 was taken with good quality laboratory equipment, but without special precautions or calibrations. Cursory examination of the data shows that a reasonable maximum error on the value $\gamma = -1434$ obtained is 100. The VSWR of the same two-port terminated by a matched load when measured by the usual means was found to be about 600. The value 600 is characteristically low in view of the usual limitations imposed by crystal law and amplifier noise.

Thanks are due to John D. Mallon, Jr., Rome Air Materiel Area, for his cooperation in making these data available.

H. M. Altschuler
Microwave Res. Inst.
Polytechnic Inst. Brooklyn
Brooklyn 1, N. Y.

A Correction Necessary for the Application of the Doppler Effect to the Measurements of Distances to Satellites*

In three previous letters,[1-3] the following equation was presented, defining the minimum passing range of a satellite from the receiving antenna

$$R_m = \frac{-v^2}{\lambda(\dot{f})_{max}}. \qquad (1)$$

In this equation,

R_m = minimum range between satellite and receiving antenna.
v = satellite's relative speed with respect to receiving antenna.
λ = wavelength of the emitted signals.
$(\dot{f})_{max}$ = maximum rate of change of frequency of the received signals.

This equation, as it was derived by Schwartzman and Stahl,[3] is based on a plane geometry that is not logically applicable with sufficient accuracy to calculations for artificial satellites. As will be demonstrated subsequently, this approximation can easily introduce errors of 15 per cent in the calculation of R_m, which obviously must be taken into consideration.

In the demonstration that follows, I will consider for simplicity the circular orbit case, applicable also for satellites with little vertical component of velocity.

Fig. 1 shows the geometry relating the observation point and the satellite. From it we obtain,

$$R^2 = R_0^2 + H^2 - 2R_0H \cos D \qquad (2)$$

where

R = distance between the satellite (S) and receiving antenna (O).
H = distance from the earth center (C) to satellite (S).
R_0 = earth radius.
D = angle formed at earth center between the vertical at receiving antenna and the vertical at satellite.

Differentiating (2) with respect to time,

$$R\dot{R} = H\dot{H} - R_0\dot{H} \cos D + R_0H \sin D\dot{D}. \qquad (3)$$

For circular orbits, $\dot{H}=0$, and (3) can be written,

$$R = \frac{HR_0 \sin D\dot{D}}{\dot{R}}. \qquad (3a)$$

The frequency deviation Δf caused by the Doppler effect is

$$\Delta f = -\frac{\dot{R}}{\lambda} \qquad (4)$$

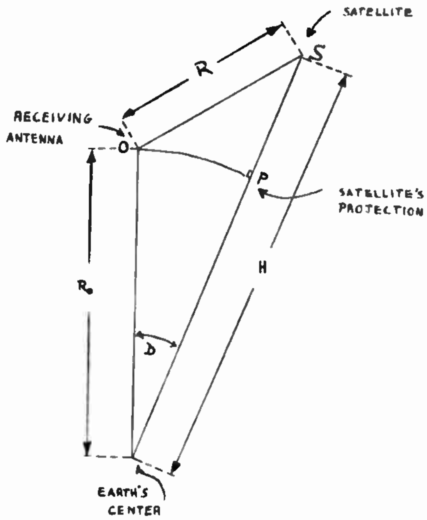


Fig. 1.

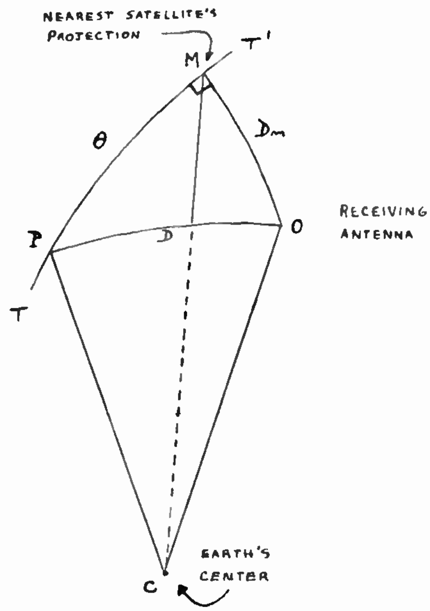


Fig. 2.

in which the higher order effects are neglected in view of the low satellite velocity relative to the electromagnetic propagation velocity.

Replacing $\dot{R}$ from (4) in (3a), we obtain

$$R = -\frac{HR_0 \sin D\dot{D}}{\lambda\Delta f}. \qquad (5)$$

Fig. 2 represents the spherical geometry that relates the observation point with the satellite's relative movement plane. The great circle arc $\overline{TT'}$ corresponds to the relative motion of the projection (P) of the satellite on the surface of the earth. The great circle arc $\overline{OP}$ from the observation point (O) to point P is the angle D previously defined; D_m is the minimum value attained by D when $\overline{OP}$ is normal to $\overline{TT'}$.

For the spherical right triangle OPM, the following trigonometric relation holds:

$$\cos D = \cos\theta \cos D_m \qquad (6)$$

where θ is the arc of the great circle $\overline{TT'}$ from P to the nearest projection at M. Differentiating (6) with respect to time, we get

$$\dot{D} = \frac{\sin\theta \cos D_m \dot{\theta}}{\sin D}. \qquad (7)$$

Replacing $\dot{D}$ from (7) in (5), we obtain

$$R = -\frac{HR_0 \cos D_m}{\lambda} \frac{\sin\theta\dot{\theta}}{\Delta f}. \qquad (8)$$

When the satellite is at the minimum range, we have

$$t = t_m$$
$$R = R_m$$
$$D = D_m$$
$$\dot{R} = 0$$
$$\Delta f = 0$$
$$\theta = 0.$$

Substituting these values in (8), we have

$$R_m = -\frac{HR_0 \cos D_m}{\lambda} \lim_{t \to t_m} \left[\frac{\sin\theta \cdot \dot{\theta}}{\Delta f}\right]$$
$$= -\frac{HR_0 \cos D_m (\dot{\theta})^2}{\lambda(\dot{\Delta f})_{max}}$$
$$= -\frac{HR_0 \cos D_m (\dot{\theta})^2}{\lambda(\dot{f})_{max}}.$$

Since $v = \dot{\theta} \cdot H$, we can write

$$R_m = -\frac{R_0}{H} \cos D_m \frac{v^2}{\lambda(\dot{f})_{max}}. \qquad (9)$$

For D_m values less than 5°,

$$R_m \approx -\frac{R_0}{H} \frac{v^2}{\lambda(\dot{f})_{max}}. \qquad (9a)$$

For a satellite with a height of 1500 km over the earth's surface, the factor R_0/H of (9) represents a reduction of 19 per cent in R_m, from the value calculated with (1).

J. M. Brito, Infante
Escuela de Electronica
Armada de Chile
Vina del Mar, Chile.

* Received by the IRE, April 27, 1959.

[1] R. R. Brown, P. E. Green, B. Howland, R. M. Lerner, R. Manasse, and G. Pettengill, "Radio observations of the Russian earth satellite," Proc. IRE, vol. 45, pp. 1552–1553; November, 1957.

[2] A. M. Petersen, "Radio and radar tracking of the Russian earth satellite," Proc. IRE, vol. 45, pp. 1553–1557; November, 1957.

[3] A. Schwartzman and P. D. Stahl, "Doppler equation for earth satellite measurements," Proc. IRE, vol. 46, pp. 915–916; May, 1958.

The Ion-Trap Results in "Exploration of the Upper Atmosphere with the Help of the Third Soviet Sputnik"*

The recent paper by Krassovsky[1] describing the ion-trap experiment carried aboard the Soviet satellite, Sputnik III, concludes on the basis of the experimental current-voltage characteristic of the collector that the effective electron temperature at an altitude of 795 km is not less than 15,000°K, corresponding to a negative vehicle potential of 6.4 volts with respect to the plasma. The computed electron temperature

* Received by the IRE, April 23, 1959.

[1] V. I. Krassovsky, "Exploration of the upper atmosphere with the help of the third Soviet Sputnik," Proc. IRE, vol. 47, pp. 289–296; February, 1959.

is high and has been questioned by other authors.[2] This note questions the interpretation of the experimental data and presents an analysis which leads to lower values for vehicle potential and electron temperature.

The experimental curve for 795 km as given in Fig. 8[1] is reproduced here (curve A of Fig. 1). The abscissa is the potential of the ion-trap grid relative to the satellite which equals the difference between the trap and satellite potentials relative to the plasma ($\phi_{TRAP}-\phi_s$). To obtain the vehicle potential the following reasoning was employed: "The point in which the collector current stops decreasing with the increase of the potential of the trap cover grid, with respect to the surrounding medium, corresponds to the retardation potential of the ions which can be determined on the basis of the well-known formula shown in Appendix II." The equation given in Appendix II is

$$e\phi_R = \frac{mV^2}{2}. \qquad (1)$$

where e is the ion charge, m the mass of the ion, V the velocity of the satellite, and ϕ_R the retarding potential. Thus with an assumed mass of 16 AMU (O+) and the known satellite velocity, the corresponding retarding potential is 5.1 volts. Experimentally the current stopped decreasing at a value of ($\phi_{TRAP}-\phi_s$) equal to 11.5 volts. Since the Russian paper assumes that at this point $\phi_{TRAP}=\phi_R$, the vehicle potential is computed to have been 6.4 volts negative with respect to the medium.

The preceding argument is based on the premise that the kinetic energy of the ions with respect to the satellite is 5.1 electron volts. This is true only if one is careful to specify that it is an *average* kinetic energy. If it is assumed that in the undisturbed atmosphere the velocity distribution is isotropic, then there are almost equal numbers of ions above and below this average kinetic energy if the satellite velocity is large compared to the mean thermal velocity of the ions.

Therefore, a retarding potential corresponding to the average kinetic energy will only stop approximately one-half the incident ions, and the point on the abscissa corresponding to a retarding potential of 5.1 volts is not at 11.5 volts but somewhere on the linear portion of the curve which extends from 5 to 9 volts.

This conclusion is verified by a derivation of the current as a function of the grid potential. If a fixed co-ordinate system is chosen with the positive x axis in the direction of the satellite velocity vector, an ion with a thermal velocity component c in this direction has a velocity component $v=c-V$ in a co-ordinate system moving with the satellite. The velocity distribution in free space if assumed to be Maxwellian is given by

$$\frac{dN}{dc} = \frac{N_0}{a\sqrt{\pi}} e^{-c^2/a^2}. \qquad (2)$$

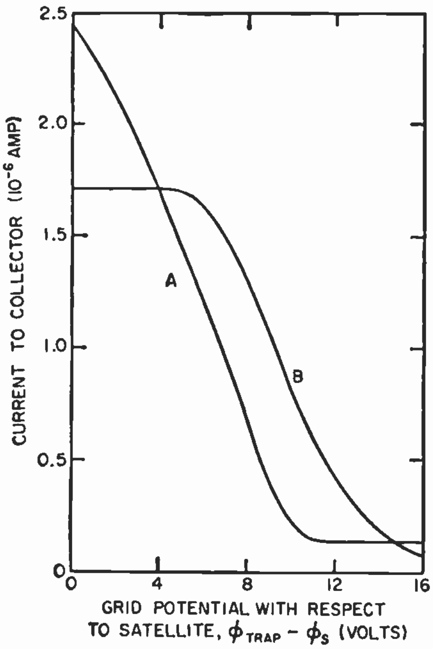


Fig. 1—Ion-trap current-voltage curves. A) Russian results. B) Theoretical curve for a similar trap with N_0, T and ϕ_s of 2.7×10^5 ions/cm³, 8800°K, and −3.9 volts.

where N_0 is the ambient ion density and a the most probable ion velocity ($\frac{1}{2}ma^2=kT_+$). The velocity distribution in the satellite co-ordinate system is obtained by replacing c by $v+V$:

$$\frac{dN}{dv} = \frac{N_0}{a\sqrt{\pi}} e^{-(v+V)^2/a^2}, \qquad (3)$$

and all ions with $v-\sqrt{2\phi e/m}$ will not reach the collector. The collector current is

$$I = \int \alpha e A v dN$$

$$= \frac{\alpha e A N_0}{a\sqrt{\pi}} \int_{-\infty}^{-\sqrt{2\phi e/m}} v e^{-(v+V)^2/a^2} dv, \qquad (4)$$

where A is the collection area and α is a grid transparency coefficient. If the current is considered as a positive quantity, then when integrated it becomes

$$I = \alpha e A N_0 V \left[\frac{1}{2} + \frac{1}{2}\operatorname{erf}(x) + \frac{a}{2V\sqrt{\pi}} e^{-x^2}\right], \qquad (5)$$

where x is given by

$$x = \frac{1}{a}\left(V - \sqrt{\frac{2\phi e}{m}}\right). \qquad (6)$$

The expression in the bracket in (5) gives explicitly the $f(\phi)$ of the equation in Appendix I of the Russian paper. This approaches unity at $\phi=0$ for large values of the ratio (V/a). If $V=0$ the current reduces to the well-known equation for a Langmuir probe.[3] When $\phi=\phi_R$ as given by (1), then $x=0$, and the current is

$$I_{\phi=\phi_R} = \alpha N_0 e A\left(\frac{V}{2} + \frac{a}{2\sqrt{\pi}}\right), \qquad (7)$$

which for large (V/a) approaches one-half the current for zero volts.

It is not easy to distinguish ϕ_R on the experimental curve when the vehicle potential is unknown. A more promising approach is to consider the slope of the curve. At ϕ_R it can be shown that $dI/d\phi$ has a maximum (negative) value.

Actually, this slope is almost constant over a considerable interval about $x=0$ and is given by

$$\left(\frac{dI}{d\phi}\right)_{\phi\cong\phi_R} \cong -\frac{\alpha N_0 e^2 A}{\sqrt{2\pi m k T_+}}. \qquad (8)$$

If the relation between I and N_0 for curve A when $\phi=0$ is used,

$$N_0 \text{ (ions/cm}^3) = 1.6 \times 10^{11} I \text{ (amps)}, \qquad (9)$$

then it is possible to solve for N_0, T_+, and ϕ. from (8), (9), and the experimental curve if a relation between T_+ and ϕ_s can be found. For instance, if the ions and electrons are in thermal equilibrium, the following relation is appropriate for a rocket and would be only slightly different for a satellite as long as only normal diffusion charging processes are operative:[4]

$$T_e = T_+ = \frac{2e|\phi_s|}{k \ln(m_+/m_-)}. \qquad (10)$$

The results for N_0, T, and ϕ_s from (8), (9), (10), and curve A are 2.7×10^5 ions/cm³, 8800°K and −3.9 volts, and have been used in (5) to compute curve B of the figure. These values, particularly of T and ϕ_s, should not be considered to be actual values, nor should curve B be expected to match curve A, not only because of the probability of the absence of thermal equilibrium, but also because of the following factors which must still be considered. The residual current, due probably to ionization between the grid and collector, was neglected here, but easily could be taken into account. Orbital motion of ions between the grid and collector would reduce the collector current,[3] and an increase of the sheath thickness for negative grid potentials would cause an increase in the effective collection area A. A plane geometry ion-trap could eliminate these last two complications and thus provide reliable measurements of these quantities, particularly if ϕ_s were determined independently.

The author would like to express his appreciation to R. E. Bourdeau and G. P. Serbu of the Space Sciences Division, and C. A. Pearse of the Naval Research Laboratory, for their helpful criticism.

E. C. Whipple, Jr.
Beltsville Space Center
National Aeronautics and
Space Administration
Washington 25, D. C.

[2] W. W. Berning, "Earth satellite observations of the ionosphere," Proc. IRE, vol. 47, pp. 280–288; February, 1959.

[3] I. Langmuir and K. Compton, "Electrical discharges in gases," *Revs. Mod. Phys.*, vol. 3, pp. 191–257; April, 1931.

[4] I. M. Imyanitov, "Measurement of electrostatic fields in the upper layers of the earth's atmosphere," *Usp. Fiz. Nauk*, vol. 63, pp. 179–201; September, 1957.

A Double Pumping Scheme Applicable to Low-Frequency Masers*

Two multiple level pumping schemes have been used successfully to improve the performance of paramagnetic masers. One is termed "push-pull" pumping[1] and the other "push-push" pumping.[2] Another possible scheme for increasing maser performance, which might be called "parallel" pumping, is suggested here.

The energy level configuration for the parallel pumping scheme assumes the form shown in Fig. 1.

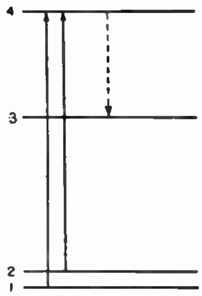


Fig. 1.

It is seen that this scheme employs pumping transitions between levels 1 and 4, and 2 and 4. The essential point is that levels 1 and 2 are sufficiently close together to be pumped by the same cavity mode and klystron mode.

Assuming that induced emission is to occur between levels 4 and 3, the difference of populations of levels 4 and 3 achieved in this case is calculated by making use of the dynamic equations of the spin system:

$$\frac{dn_i}{dt} = -n_i \sum_{j=1}^{4} w_{ij} + \sum_{j=2}^{4} n_j w_{ji} + \sum_{j=1}^{4} (n_j - n_i) W_{ji} \quad i = 1, 2, 3, 4 \quad j \neq i \tag{1}$$

and the conservation equation:

$$N = \sum_{i=1}^{4} n_i. \tag{2}$$

Here n_i is the population of the ith level; N is the total number of active spins; w_{ij} is the thermal relaxation transition probability of the $i \leftrightarrow j$ transition,

$$w_{ji} = w_{ij} \exp \frac{hf_{ij}}{kT}, \quad j > i,$$

where f_{ij} is the resonant frequency of the $i \leftrightarrow j$ transition; and W_{ij} is the induced transition probability of the $i \leftrightarrow j$ transition

$$W_{ji} = W_{ij}.$$

Assuming steady-state operation, *i.e.*,

$$\frac{dn_i}{dt} = 0 \qquad i = 1, 2, 3, 4 \tag{3}$$

and solving (1) and (2) simultaneously, we obtain, under the assumptions of small signal and large pumping:

$$(n_4 - n_3)_1 = \frac{Nh}{4kT} \frac{w_{32}f_{32} + w_{31}f_{31} - w_{43}f_{43}}{w_{34} + w_{23} + w_{13} + W_{43}} \cdot \tag{4}$$

Repeating this analysis with $W_{14}=0$, one has:

$$(n_4 - n_3)_2 = \frac{Nh}{4kT} \frac{w_{31}(w_{21}f_{32} - w_{41}f_{43}) + (w_{41} + w_{31} + w_{21})(w_{32}f_{32} - w_{43}f_{43})}{w_{31}(w_{41} + w_{21}) + (w_{41} + w_{31} + w_{21})(w_{13} + w_{32} + W_{43})} \cdot \tag{5}$$

The power emitted by the paramagnetic crystal is given by:

$$P_{43} = (n_4 - n_3) h W_{43} f_{43}. \tag{6}$$

Straightforward algebraic manipulation shows that for a given maser system

$$(P_{43})_1 > (P_{43})_2.$$

The actual improvement in performance is dependent on the relaxation times which are characteristic of particular transitions. Under the reasonable assumption of equal relaxation times an improvement by approximately a factor of two may be expected.

The parallel pumping scheme is of particular importance when low frequency operation of the ruby maser is desired. Such operation is conveniently obtained at large values of polar angle and low magnetic field. There, the energy level configuration is essentially that shown in Fig. 1.[3] For example, gain bandwidth products in excess of 40 mc at 4.2°K have been obtained at the following operating point: $\theta = 90°$, $H = 750$ gauss, $f_p = 14.0$ kmc, $f_s = 4.08$ kmc. This suggests a gain bandwidth product of at least 100 mc at 1.25°k.

J. E. King
A. Birko
G. Makhov
Willow Run Labs.
The University of Michigan
Ann Arbor, Mich.

* Received by the IRE, May 25, 1959. This research was supported by Project Michigan (administered by the U. S. Army Signal Corps).

[1] G. Makhov, C. Kikuchi, J. Lambe, and R. W. Terhune, "Stimulated Microwave Emission in Ruby," presented at Internatl. Conf. on Solid State Physics and Its Applications to Electronics and Telecommunications, Brussels, Belgium; June, 1958.

[2] J. W. Meyer, "Multiple or Regenerate Mode Pumping of Solid State Masers," Lincoln Lab., M.I.T., Lexington, Mass., Quart. Prog. Rept., p. 71; November, 1958.

[3] W. S. Chang and A. E. Siegman, "Characteristics of Ruby for Maser Applications," Stanford Electronics Lab. Tech. Rept. No. 156–2; September 30, 1958.

Low Noise Parametric Amplifier*

In their paper[1] reporting experiments with an S-band reactance amplifier, Knechtli and Weglein point out that a method of reducing the proportion of damping caused by the varactor can be achieved by reducing the coupling between the varactor and the associated resonant circuit.

An alternative procedure would be to *increase* the coupling of the source and load. The drive then has to be increased to generate a larger negative conductance, just as it has to be increased with reduced varactor coupling to maintain the conductance.

Suppose the bias and pump voltage have been set to give an optimum value of conversion capacitance, and that the significant terms in the Fourier expansion are C_v (the average) and C_c (the first conversion term). For this condition let the equivalent shunt resistance be R_v. If the remainder of the circuit contributes a shunt resistance R, we have the approximate result

$$\left(\frac{1}{R} + \frac{1}{R_v}\right) \propto C_c^2.$$

If the coupling is now reduced the result is say, $C_{c/n}$, contributed in the circuit. Hence the maximum circuit conductance for equivalent conditions will be

$$\frac{1}{n}\left(\frac{1}{R} + \frac{1}{R_v}\right) = \frac{1}{nR} + \frac{1}{nR_v}.$$

Since the damping contributed by the varactor has already been changed to $1/nR_v$, the coupling of the remainder of the circuit has to be changed so that the proportions are eventually the same.

In practice there is stray capacitance in addition to C_v and the bandwidth will be greatest when the maximum possible coupling is used

$$\left(\text{maximum ratio of } \frac{C_c}{C_v + C_{\text{strays}}}\right).$$

In addition, the cavity losses will give least effect when the maximum coupling is used, although this factor is probably unimportant except when the varactor is cooled and the cavity left at ambient temperature.

A big advantage of operating in the large C_c condition is that the conversion and gain are rendered less sensitive to pump level variations.

When bandwidth is not important, a convenient technique for reducing the varactor coupling is to connect a capacitance in series, though this makes it less easy to apply bias.

The above arguments have assumed for simplicity that both signal and idler-circuit damping are changed in the same proportions; this is certainly true for the amplifier used by Knechtli and Weglein.

In practical applications the simple circuit conditions are no longer valid because of stray parasitic reactances which may be very serious at high frequencies and restrict the bandwidth.

In closing, the author suggests the name REACTIFIER for this type of amplifier, instead of parametric amplifier, etc.

N. Houlding
Microwave Associates, Inc.
Burlington, Mass.

* Received by the IRE, May 4, 1959.

[1] R. C. Knechtli and R. D. Weglein, "Low noise parametric amplifier," Proc. IRE, vol. 47, pp. 584–585; April, 1959.

[2] H. Heffner and G. Wade, "Gain, band width, and noise characteristics of the variable parameter amplifier," *J. Appl. Phys.*, vol. 29, pp. 1321–1331; September, 1958.

Contributors

Donald W. Baker (S'58) was born in Skagway, Alaska, on April 12, 1932. He is presently a senior in electrical engineering at the University of Washington, Seattle, where he holds the position of senior electronic technician in the Department of Physiology and Biophysics. His responsibilities include the design and development of instruments for medical investigations.

D. W. Baker

From 1951 to 1955, he was a radar fire control systems technician in the USAF. His last year of enlistment was spent as a electronic technician at the Electronics Research Directorate, Air Force Cambridge Research Center. Since that time, and until 1957, he was an instrumentation technician at the Boeing Windtunnel in Seattle, Wash.

❖

Roland C. Bostrom (S'49–A'50–M'55) was born in Jamestown, N. Y., on July 7, 1923. He received the B.S. degree in electrical engineering from the University of Michigan, Ann Arbor, in 1949.

R. C. Bostrom

Since 1949 he has been employed at the Airborne Instruments Laboratory, Mineola, N. Y., where he worked on VHF instrumentation and magnetic detectors. In 1955 he joined the Department of Medical and Biological Physics, where he has been engaged in the application of computer techniques to medical problems.

Mr. Bostrom is a member of the Inter-Society Cytology Council, the New York Microscopical Society, Sigma Xi, Tau Beta Pi, Eta Kappa Nu, and Sigma Pi Sigma.

❖

Britton Chance (M'46–SM'46–F'54) was born in Wilkes-Barre, Pa., on July 24, 1913. He received the B.S. degree in 1935, and the M.S. degree in 1936, both from the University of Pennsylvania, Philadelphia, Pa. He was a Johnson Foundation fellow at Pennsylvania from 1939 to 1940, and received the Ph.D. degree in physical chemistry in 1940. He also received the Ph.D. degree in physiology in 1942, and the D.Sc. degree in 1952, both from Cambridge University.

B. Chance

From 1940 to 1941 he was acting director of the Johnson Foundation at the University of Pennsylvania; in 1941 he became an investigator for the Office of Scientific Research and Development; from 1941 and 1945 he was a research associate, group leader, and associate division head at the Radiation Laboratory, Massachusetts Institute of Technology, Cambridge. He was a Guggenheim Fellow at the Nobel Institute, Stockholm, Sweden, and at the Molteno Institute, Cambridge, England, from 1946 to 1948, and in 1948 became a scientific consultant to the attaché for research of the U. S. Navy in London, England. From 1941 through 1949 he had been an assistant professor of biophysics at the Medical School, University of Pennsylvania. In 1949 he became a professor and also director of the Johnson Foundation. In 1950 he was awarded the Presidential Certificate of Merit and in 1952 the Paul Lewis award of the American Chemical Society.

He was a consultant for the National Science Foundation from 1951 to 1956, working with the National Research Council and the American Red Cross committee on blood and blood derivatives. He was on the visiting committee of the Biology Department, M.I.T., from 1954 to 1956. He was also a Phillips lecturer at the University of Pittsburgh and a Pepper lecturer at the University of Pennsylvania, in 1956 and 1957. He has been on the visiting committee of the Bartol Research Foundation, from 1955–1959, and since 1957 has been a councilor for the Society of General Physiologists. Since 1959 he has been on the President's Scientific Advisory Committee.

His areas of interest have included automatic ship steering; photoelectric control units; radar timing and computing devices; sensitive spectrophotometers; enzyme-substrate compounds and reaction mechanisms of catalases; peroxidases and dehydrogenases; cytochrome oxidases; dynamics of intracellular enzyme systems; and oxidative phosphorylation.

Dr. Chance is a member of the National Academy, the American Society of Biological Chemists, the American Chemical Society, the Harvey Society, the American Philosophical Society, and the American Physiological Society, and a fellow of the Physical Society and the American Academy of Arts and Sciences.

❖

Murray Eden was born in Brooklyn, N. Y., in 1920. He received the B.S. from The College of the City of New York in 1939 and the Ph.D. in physical chemistry from the University of Maryland in 1951.

From 1943 to 1949, he was associated with the National Bureau of Standards, Washington, D. C., working in the physical chemistry and the mechanical instruments sections. Then, in 1949, he joined the biophysics section of the National Cancer Institute, Bethesda, Md. He was a special fellow of the U. S. Public Health Service in the mathematics department of Princeton University, Princeton, N. J., from 1953 to 1955, and in 1955 joined the laboratory of technical development of the National Heart Institute, Bethesda, Md., where his principal interests are instrumentation and the application of mathematical models to biology. He is presently a research associate in the center for communication sciences at the Massachusetts Institute of Technology, Cambridge.

M. Eden

Dr. Eden is a member of the American Physical Society and Sigma Xi.

❖

Richard M. Ellis was born in Mexico City, Mexico, on July 24, 1922. He obtained his first electronic experience as a Navy radio technician during World War II. In 1949 he received the B.S. degree in electrical engineering from the University of Washington, Seattle.

R. M. Ellis

After graduation, he continued at the University Medical School, working on the design and development of instrumentation for cardiac research. In 1954, he joined the engineering staff of the Seattle Development Laboratory of Minneapolis-Honeywell Regulator Company, where he is presently employed, working primarily on the design of military electronic systems and underwater acoustics.

❖

Leslie E. Flory (A'35–SM'46–F'57) was born in Sawyer, Kans. on March 17, 1907. He received the B.S. degree in electrical engineering from the University of Kansas, Lawrence, in 1930.

L. E. Flory

He joined the RCA Manufacturing Company, Camden, N. J., in 1930 and worked largely on television pickup tubes, particularly on the development of the iconoscope, and on earliest types of photomultipliers. Later, he was concerned with the development of basic circuits for electric computers, some of which form the basis of present systems. During World War II, he was engaged on a project on infrared image tubes and equipment. In 1942, he was transferred to the RCA Laboratories at the David Sarnoff Research Center, Princeton, N. J., where, as a member of the technical staff, he continued work on infrared until the end of the war. This work included development of the "snooperscope." In 1945, he took over a project on sensory aids for the blind, developing reading devices under a Veterans Administration contract. From 1946 to 1954, he was in charge of work on storage tubes and, after 1950, in charge of work on industrial television. Since then, he has pioneered in the development of industrial television systems. Currently, he supervises the General Research Laboratory, which incorporates work on industrial television as well as medical electronics. More recently, he also has been engaged in a project on electronic control of traffic vehicles.

Mr. Flory is the recipient of three RCA Incentive Awards and is a member of Sigma Xi, Kappa Eta Kappa, and Sigma Tau.

❖

Lawrence J. Fogel (S'47–A'49–M'53–SM'56) was born in Brooklyn, N. Y., on March 2, 1928. He received the B.E.E. degree from New York University, N. Y., in 1948 and the M.S. degree from Rutgers University, New Brunswick, N. J., in 1952. He is currently completing work toward his doctorate degree at the University of California, Los Angeles.


L. J. Fogel

From 1948–1953 he was employed in various engineering activities at Fort Monmouth, N. J., where he was primarily concerned with aspects of communication, including antenna design and evaluation. He was subsequently project engineer on several human engineering research studies for Stavid Engineering, Inc., Plainfield, N. J. There a mathematical model for the human operator was devised which included representation for anticipation, quantization, and sampling operation. This was further pursued at Convair, San Diego, Calif., when he was employed there in 1956 for the investigation of cockpit display techniques. He is presently head of the Reliability Group at Convair, San Diego Division.

Mr. Fogel is a member of the New York Academy of Science, the Acoustical Society of America, the American Rocket Society, and the Human Factors Society.

❖

Dean L. Franklin was born on April 18, 1929, in Cheney, Wash. He received the B.S. degree in physics in 1953 from the University of Washington, Seattle; while there he was employed as chief electronic technician for the Department of Physiology and Biophysics. He began his medical education in 1958 at McGill University, Faculty of Medicine, Montreal, Can. He is currently a member of the cardiovascular research techniques training program at the University of Washington School of Medicine.

His initial training in electronics was obtained while he was in the Army at the Signal School, Fort Monmouth, N. J., where he served as senior instructor of the advanced radar electronics course. Following his discharge he was employed by Boeing Aircraft Company as experimental electronic technician on the Bomarc project.


D. L. Franklin

❖

Walter H. Freygang, Jr., was born in Jersey City, N. J., on December 27, 1924. He received the degrees of M.E. from Stevens Institute of Technology in 1945 and M.D. from the University of Pennsylvania in 1949.


W. H. Freygang

He was assistant resident in medicine at Bellevue Hospital in New York from 1949 to 1951 and then assistant resident in neurology at at New York Neurological Institute until 1952. Since then he has held the position of neurophysiologist with the Laboratory of Neurophysiology at the National Institutes of Health, Bethesda, Md. His fields of research have ranged from blood circulation studies in the brain through nerve junctions to propagation along single nerve fibers.

Dr. Freygang is a member of the Association for Research in Nervous and Mental Disease, American Physiological Society, Biophysical Society, New York Academy of Sciences, AAAS, and the Marine Biological Association.

❖

Robert C. Gesteland was born in Madison, Wis., on July 1, 1930. He received the B.S. degree from the University of Wisconsin in 1953 and the M.S. degree from Massachusetts Institute of Technology in 1957, both in electrical engineering. He is presently a Ph.D. candidate in biology at Massachusetts Institute of Technology, and is a predoctoral fellow of the Institute of Neurological Diseases and Blindness of the National Institutes of Health. During 1953 and 1954 he was an engineer with General Radio Company and was a research assistant at the Research Laboratory of Electronics, M.I.T., from 1955 to 1957.


R. C. Gesteland

Mr. Gesteland is a member of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu.

Joseph E. Hayes, Jr., was born in El Dorado, Ark., in 1927. In 1948, he received the A.B. degree in chemistry and in 1953 the Ph.D. degree in biological chemistry from Washington University, St. Louis. Mo. After serving with the U. S. Army Medical Service Corps from 1953 to 1957, he joined the Laboratory of Technical Development of the National Heart Institute, Bethesda, Md.


J. E. Hayes, Jr.

Dr. Hayes is a member of the Biophysical Society.

❖

Julia F. Herrick (M'46) was born in North Saint Paul, Minn., on September 14, 1893. She received the B.A. degree in 1915 from the University of Minnesota. After three years as a high-school teacher, she spent one year as a teaching fellow in physics at the Graduate School of the University of Minnesota, where she received the M.A. degree in 1919.


J. F. Herrick

She returned to high-school teaching for one year, and then taught mathematics and physics at the Northrup Collegiate School, Minneapolis, Minn., from 1920 to 1922. From 1922 to 1927 she was head of the Department of Physics at Rockford College, Rockford, Ill.

The Mayo Foundation, Rochester, Minn., granted her a fellowship in biophysics in 1927, and for more than five years she pursued extensive studies in biophysical research, physiology, and biophysics, and also special courses at the University of Minnesota Medical School. She was awarded the Ph.D. degree in biophysics from the University of Minnesota in 1931.

Dr. Herrick was appointed a member of the staff of the Mayo Clinic and Mayo Foundation as a consultant in biophysics at the Institute of Experimental Medicine in 1936. She became an assistant professor of physiology in the Mayo Foundation in 1938, an associate professor in 1945, and professor in 1958. Her designation was later changed to professor of biophysics.

Dr. Herrick's studies of blood flow in various mammalian blood vessels have resulted in some modifications in the Rein thermostromuhr which extended the use of this instrument to the measurement of the flow of blood under more normal conditions.

In 1942, at the request of the War Department, Dr. Herrick joined the Signal Corps Engineering Laboratories, Fort Monmouth, N. J., where her work was concerned largely with radio direction-finding. Returning to the Mayo Foundation in 1946, her research activities were devoted to the biologic effects of microwaves and ultrasound, physiologic thermometry, and the circulation of the blood. She retired from the Mayo Clinic in 1958.

During the last several years Dr. Herrick has worked extensively on the development of an ultrasonic flowmeter and, with the assistance of electronic engineers and physicists, a first model of the flowmeter has been constructed. The instrument is presently being used for measuring the velocity of blood outside the natural circulation in the body; further development is necessary for measurements in the living body. She also has periodically given a course of lectures in electricity and magnetism, including electronics, to the general staff in the Section of Physical Medicine and Rehabilitation.

Dr. Herrick is a member of the American Physiological Society, the American Physical Society, the American Association for the Advancement of Science, the Alumni Association of the Mayo Foundation, Sigma Xi, Phi Beta Kappa, Alpha Gamma Delta, and Mortar Board. She has been chairman of the IRE Professional Groups on Medical Electronics and Ultrasonics Engineering, and is former editor of IRE TRANSACTIONS ON MEDICAL ELECTRONICS.

❖

John P. Hervey (A'27–VA'39–M'55) was born in Montclair, N. J., on January 1, 1905. He received the Master's degree in engineering from Harvard University, Cambridge, Mass., in 1929. He has been engaged in the field of electronic instrumentation for the biomedical sciences since 1935, when he joined the Johnson Foundation at the University of Pennsylvania, Philadelphia. Except for a year at Cornell Medical College as assistant professor of biophysics, he was with the Johnson Foundation until 1949, when he became associate professor of biophysics at The Johns Hopkins University. Since 1954 he has been senior electronic engineer at the Rockefeller Institute for Medical Research, Woods Hole, Mass.

J. P. HERVEY

❖

Bradford Howland was born in Lafayette, Ind., on October 8, 1927. He received the B.S. degree in electrical engineering from Purdue University in 1945 and the M.S. degree in applied science from Harvard University in 1948.

B. HOWLAND

In 1946 and 1947 Mr. Howland was with the rocket sonde section of the Naval Research Laboratory, Washington, D. C., where he worked on electronics design for cosmic ray experiments. He was at the Research Laboratory of Electronics, Massachusetts Institute of Technology, from 1950 to 1956, working on speech recognition and on instrumentation for neurophysiology. He was with Hycon Eastern Company, Cambridge, Mass., in 1956 and since 1957 he has been working with the pattern recognition group at Lincoln Laboratory, M.I.T.

Mr. Howland is a member of Sigma Xi.

❖

Kenneth E. Jochim was born in St. Louis, Mo., on July 30, 1911. He received the A.B. degree in 1939 and the Ph.D. degree in physiology in 1941 from the University of Chicago, Ill.

K. E. JOCHIM

He was appointed research assistant and research associate in the Cardiovascular Department of Michael Reese Hospital, Chicago, from 1931 to 1942. In 1942, he joined the Physiology Department of St. Louis University School of Medicine as instructor and was promoted to assistant professor in 1945. In 1946, he went to the University of Kansas, Lawrence, as professor of physiology and chairman of the department. In 1956–1957, he spent a year at the Physiological Institute of the University of Munich, Germany, while on a Fulbright Research Scholarship. His major research interest is in the field of cardiovascular physiology particularly in the area of hemodynamics and cardiovascular regulatory mechanisms.

Dr. Jochim is a member of the American Physiological Society, the Biophysical Society, and Sigma Xi.

❖

Paul E. Klopsteg was born in Henderderson, Minn., on May 30, 1889. He attended the University of Minnesota, and received the Ph.D. degree in physics in 1916. He also holds honorary Sc.D. degrees from Northwestern and Wesleyan Universities.

P. E. KLOPSTEG

During World War I he was a development engineer in the U. S. Army Ordnance Department.

His career as a research administrator has included service during World War II with the Office of Scientific Research and Development and the Office of Field Service. He was awarded the Medal of Merit with Presidential Citation for his wartime work.

He is a former president of the Central Scientific Company. He has also been a member and chairman of the board of governors of the Argonne National Laboratory. Since 1953 he has served on the Personnel Security Review Board of the Atomic Energy Commission.

Dr. Klopsteg participated in the establishment of the Committee on Artificial Limbs in the National Research Council and was chairman of the committee for eleven years. He is also a member of the Committee on Meteorology, National Research Council, and alternate member of the Defense Science Board.

He is president of the American Association for the Advancement of Science, former president of the American Association of Physics Teachers, chairman of the governing board of the American Institute of Physics, a fellow of the American Physical Society, and a member of the Optical Society of America.

❖

Robert S. Ledley (M'58) was born on June 28, 1926, in New York, N. Y. He received the M.A. degree in mathematical physics from Columbia University, New York, N. Y., in 1949, and the D.D.S. degree from New York University in 1948. He joined the National Bureau of Standards, Washington, D. C. in 1951, working in biophysics and the logical design of digital computers. In 1954 he became a staff member of the Operations Research Office, where he worked on computer simulations and the application of logic to operations research problems.

R. S. LEDLEY

Since 1956 he has been an assistant professor of electrical engineering at The George Washington University, Washington, D. C. He is also a staff member of the National Academy of Sciences-National Research Council project on the use of computers in biomedical sciences. He is consultant mathematician at the Data Processing Systems Division of the National Bureau of Standards, and consultant to the National Library of Medicine. Currently, he is principal investigator of several NIH and NSF grants on information retrieval systems, computational methods in symbolic logic, and computers in biomedical research.

Dr. Ledley is a member of the American Mathematical Society, American Physical Society, Operations Research Society of America, and the American Museum of Natural History.

❖

Jerome Y. Lettvin was born in Chicago, Ill., on February 23, 1920. He received the B.S. and M.D. degrees in 1942 and 1943, respectively, from the University of Illinois, Chicago, and interned at Boston City Hospital, Harvard Nerve Service, Boston, Mass., from 1943 to 1944.

J. Y. LETTVIN

He was with the U. S. Army, from 1944 to 1946, and was head of the Neuropsychiatry 237th General Hospital. From 1948 to 1951, he was senior psychiatrist, Manteno State Hospital, Manteno, Ill.

Since 1951, he has been a member of the staff of the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge. Since 1957, he has also been part-time senior psychiatrist at the Boston Psychopathic Hospital, and has been appointed Visiting Associate Professor of Physiology at Massachusetts Institute of Technology for 1959–1960.

Dr. Lettvin is a member of the American Board of Psychiatry and the American Physiological Society.

❖

Martin Lubin was born in New York, N. Y., on March 30, 1923. He received the B.A. degree from Harvard University, Cambridge, Mass., in 1942, the M.D. degree from Harvard Medical School, Boston, Mass., in 1945, and the Ph.D. degree from the Massachusetts Institute of Technology, Cambridge, in 1954.

M. Lubin

Since 1954, he has been a member of the faculty at Harvard Medical School, and is now an assistant professor of pharmacology. In his research work, he has been concerned with various aspects of muscle physiology, transport of electrolytes, and microbial genetics.

❖

Lee B. Lusted (F'59) was born in Mason City, Iowa on May 22, 1922. He received the B.A. degree from Cornell College, Mount Vernon, Iowa in 1943, and the M.D. degree from Harvard Medical School, Cambridge, Mass., in 1950.

L. B. Lusted

From 1943 to 1946 he was a special research associate at the Radio Research Laboratory, Harvard University, working on radar counter measures and during this time he spent thirteen months in the European Theatre of Operations as a member of the American British Laboratory. After graduation from medical school, he took specialty training in radiology at the University of California Hospitals, San Francisco and subsequently was assistant professor of radiology. From 1956 to 1958, he was assistant radiologist at The National Institutes of Health, Bethesda, Md., and he is now associate professor of radiology at the University of Rochester School of Medicine and Dentistry, Rochester, N. Y.

Dr. Lusted was national chairman of the Professional Group on Medical Electronics in 1957 and at present is editor of the IRE Transactions on Medical Electronics. He is a diplomate of the American Board of Radiology, and a member of the American College of Radiology, the Association of University Radiologists, and Sigma Xi.

Edward F. MacNichol, Jr. (A'46–M'55) was born in Toledo, Ohio, on October 24, 1918. He received the B.A. degree in physics from Princeton University, Princeton, N. J., in 1941. His graduate study at the Johnson Foundation of the University of Pennsylvania, Philadelphia in 1947 and 1948 was supported by a predoctoral fellowship of the American Cancer Society. He received the Ph.D. degree in biophysics from The Johns Hopkins University, Baltimore, Md., in 1952.

E. F. MacNichol, Jr.

He was a staff member of the Radiation Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., from 1941 to 1946, and worked on precision radar ranging equipment, especially automatic range tracking, radar navigation and bombing devices, relay radar, and an early missile-guiding system. He was also a contributor and volume editor of volumes 19 and 20 of the Radiation Laboratory Series. After receiving his doctorate, he remained at The Johns Hopkins University, and is now an associate professor of biophysics. He is currently doing research on the physiology of the retina of the eye and is active in the design of instruments for biological research. He is also teaching physiology of vision and elementary electronics for biologists.

Dr. MacNichol is a member of the American Physical Society and the Biophysical Society.

❖

Humberto R. Maturana was born in Santiago, Chile, on September 14, 1929. He attended the Medical School, University of Chile, Santiago, from 1950 to 1954 and studied anatomy at University College, London University, England, from 1954 to 1956.

H. R. Maturana

He received the Ph.D. degree in biology from Harvard University, Cambridge, Mass., in 1959 and is a member of the staff of the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, on leave from the University of Chile.

❖

Warren S. McCulloch (SM'54) was born on November 16, 1898, in Orange, N. J. He received the B.A. degree in 1921 from Yale University, New Haven, Conn., and the M.A. and M.D. degrees in 1923 and 1927, respectively, from Columbia University, New York, N. Y. He became a diplomate of of the National Board of Medicine and served his internship from 1927 to 1928 at the Neurophysical Service, Bellevue Hospital, New York, N. Y., and his residency from 1933 to 1934 at Rockland State Hospital, N. Y. He was an honorary research fellow in the Laboratory of Neurophysiology at Yale from 1934 to 1935, a Sterling Fellow at Yale from 1935 to 1936, an instructor at Yale from 1936 to 1940, an assistant professor at Yale from 1940 to 1941, an associate professor of psychiatry at the University of Illinois, Urbana, from 1941 to 1945, and a professor of psychiatry and clinical professor of physiology at the University of Illinois from 1945 to 1952. Since 1952, he has continued research as a staff member of the Research Laboratory of Electronics of the Massachusetts Institute of Technology, Cambridge. He is the author of numerous articles on functional organization of the brain.

W. S. McCulloch

Dr. McCulloch is a Fellow of the American Association for the Advancement of Science and a member of the American Academy of Arts and Sciences, the American Academy of Neurology, the American Anatomical Society, the American Mathematical Society, the American Neurological Association, the American Physiological Society, Alpha Omega Alpha, the Association for Research in Nervous and Mental Diseases, the EEG Society (of England), the New York Academy of Science, Sigma Xi, the Society for Biological Psychiatry, and the Society for Experimental Biology and Medicine.

❖

L. H. Montgomery (A'45–SM'50) was born on January 18, 1907 in Nashville, Tenn. He attended public schools and Wallace University School in Nashville.

L. H. Montgomery

He has been connected with WSM and WSM-TV in various technical capacities since 1925 and also worked as an instructor at Vanderbilt University, Nashville, teaching radio engineering from 1941 to 1943. He then moved to the Medical School to help with an OSRD research project and there became interested in the application of electronic engineering techniques to the problems of the medical profession.

He is now medical electronics consultant and research associate in the Department of Anatomy.

Mr. Montgomery is a member of the American Institute of Electrical Engineers, Southern Electroencephalographic Society, and Sigma Xi.

For the past eight years he has been a member of the Committee on Electrical Techniques in Medicine and Biology of the AIEE. He is also a member of the IRE Standards Committee and chairman of the IRE Technical Committee on Medical Electronics.

John W. Moore (SM'56) was born in Winston-Salem, N. C., on November 1, 1920. He received the Bachelor's degree in physics from Davidson College in 1941 and the Ph.D. degree in physics from the University of Virginia in 1945.

J. W. Moore

After a year at the RCA Laboratories in Princeton, N. J., he took a teaching and research position at the Medical College of Virginia in Richmond, where he made electrical impedance measurements on the human body. In 1950 he joined the biophysics division of the Naval Medical Research Institute, Bethesda, Md., and studied impedance and potentials in muscle and nerve. Since 1954, he has been a member of the Biophysics Laboratory, National Institute of Neurological Diseases and Blindness, National Institutes of Health, Bethesda, Md., where he has been concerned largely with electronic control of nerve membranes and the resulting measurements.

Dr. Moore is a member of Sigma Xi, the American Physical Society, the Biophysical Society, New York Academy of Science, and AAAS.

❖

Frank W. Noble (SM'57) was born in Ithaca, N. Y., in 1920. He received the B.E.E. degree in 1942, and the M.E.E. degree in 1948 from Cornell University, Ithaca, N. Y.

F. W. Noble

He served with the U. S. Army Signal Corps and Air Corps from 1942–1946 and worked at Cornell as a research associate from 1948–1950. Then he joined the Laboratory of Technical Development of the National Heart Institute, Bethesda, Md., and has engaged in the design and development of electronic instrumentation for use in heart research.

Mr. Noble is a member of the Biophysical Society.

❖

Walter H. Pitts was born in Detroit, Mich., on April 23, 1923. He studied at the University of Chicago, Ill., from 1940 to 1943 and at M.I.T. from 1943 to 1951, holding a Guggenheim scholarship in 1946 and 1947.

W. H. Pitts

He was associated with the Manhattan Project in 1944 and 1945. Since 1951, he has been at Research Laboratory of Electronics, M.I.T., as a mathematician and physiologist.

John M. Reid (S'48–A'51–M'56) was born in Minneapolis, Minn., on June 8, 1926. After serving two years as an electronic technician in the U. S. Navy, he attended the University of Minnesota and received the B.E.E. degree in 1950 and the M.S. degree in 1957. From 1950 to 1953 he was a research fellow in the Department of Surgery and then in the Department of Electrical Engineering, where he also did part-time graduate work. His research project on the development of equipment for detection and diagnosis of cancer using pulsed ultrasound was continued at St. Barnabas Hospital, Minneapolis, until 1957, when he became an associate in the Electromedical Division of the Moore School of Electrical Engineering, University of Pennsylvania.

J. M. Reid

Mr. Reid is a member of Eta Kappa Nu, and is a registered professional engineer in Minnesota.

❖

Robert F. Rushmer was born in Ogden, Utah, on November 30, 1914. He received his medical training at the University of Chicago, Ill., and Rush Medical College, and his postdoctoral training in pediatrics at the Mayo Clinic, Rochester, Minn.

R. F. Rushmer

During World War II he conducted research in the effects of radial acceleration and crash injury at the School of Aviation Medicine, Randolph Field, Tex. Since 1947, he has worked in cardiovascular physiology at the University of Washington, Seattle, where he is presently professor of physiology in the Department of Physiology and Biophysics in the School of Medicine. His principal interest is the study of the function and control of the heart in intact experimental animals.

❖

Herbert S. Sawyer (S'50–A'51–M'57) was born in Fort Fairfield, Maine, on August 12, 1927. In 1951 he received the B.A. degree from Bowdoin College, Brunswick, Md., and the B.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, Mass.

H. S. Sawyer

From 1951 to 1954 he was employed at Wheeler Laboratories, Inc., Great Neck, N. Y., where he was engaged in radar system development. Since 1954 he has been employed at Airborne Instruments Laboratory, Mineola, N. Y., where he has been engaged in the development of electronic equipment for medical research.

Mr. Sawyer is a member of Phi Beta Kappa and Sigma Xi.

❖

Herman P. Schwan (M'53–SM'55–F'59) was born on August 7, 1915, in Aachen, Germany. He studied physics, electrical engineering, and biophysics in Goettingen and Frankfurt, and spent two years in industry as an electrical engineer with Siemens Telefunken. He received the Ph.D. degree in physics and biophysics in 1940 and 1946, respectively, from the University of Frankfurt, and was engaged in biophysical research and ultrahigh frequency development work from 1938 to 1947 at the Kaiser Wilhelm Institute at Frankfurt. From 1946 to 1947, he held positions as assistant director and assistant professor at the same institute.

H. P. Schwan

He came to this country in 1947 and worked for the United States Navy's Aero-Medical Equipment Laboratory as a research specialist. Since 1950, he has been with the University of Pennsylvania, Philadelphia, and holds appointments as professor of physical medicine and physics in medicine in the Graduate School and School of Medicine, and as professor of electrical engineering in the Moore School of Electrical Engineering. He heads the electromedical research team which has been organized at the University of Pennsylvania by the Electrical Engineering and Medical Schools, and is conducting research in the fields of biophysics and medical electronics. During the summer of 1956 he held a visiting professorship at the University of California, Berkeley.

Dr. Schwan is a Fellow of the American Association for the Advancement of Science, and a member of the Physical Society, the New York Academy of Science, the AIEE, the Biophysical Society, and Sigma Xi.

❖

Richard B. Shackelford (A'58) was born in Orange, N. J., on January 28, 1928. While attending the University of Colorado, Boulder, he was transmitter engineer for station KBOL from 1948 to 1950, and then continued his studies at the University of Kansas, Lawrence, until 1952.

R. B. Shackelford

In 1952, he joined the General Communications Company of Salina, Kansas. In 1954, he spent a year as computer maintenance and operation engineer at the Air Force Guided Missile Test Center, Cocoa, Fla. In 1955, he joined the Physiology Department of the University of Kansas in his present position as research assistant. His major interests have been in the field of instrumentation.

❖

Hampton W. Shirer (A'51) was born in Newton, Mass., on August 8, 1924. He graduated from Washburn University, St. Louis, Mo., with the B.S. degree in 1945 and received the M.D. degree from the University of Kansas, Lawrence, in 1948.

H. W. Shirer

After a year of internship, he spent two years as a U. S. Public Health Service Postdoctoral Research Fellow at the University of Kansas Medical Center.

From 1951 to 1953 he was appointed to the Biophysics Division of the Naval Medical Research Institute, Bethesda, Md., as a lieutenant of the Medical Corps. Following release to inactive duty he returned to the University of Kansas and, after a year as an instructor in the Department of Surgery, he was appointed as a research associate in the Department of Physiology. In 1956 he received a three-year Lederle Medical Faculty Award and was promoted to his present position as assistant professor in the Physiology Department. He has been working chiefly in the field of cardiovascular dynamics and in the development of biophysical instruments.

Dr. Shirer is a member of Sigma Xi and the Biophysical Society.

❖

Lawrence Stark was born in New York, N. Y., on February 21, 1926. He received the A.B. degree from Columbia College in 1945, and the M.D. degree from Albany Medical College in 1948. In 1957 he became a Board Certified Neurologist, following training at the Neurological Institute, Columbia University, the U. S. Naval Medical Corps, and Yale University.

L. Stark

His research interests began with biochemistry at Oxford in 1949 and neuromuscular physiology at University College, London, in 1950–1951, and continued during his term as Assistant Professor of Physiology at New York Medical College in 1952. Following Naval service during the Korean War he became interested in the application of concepts from servoanalysis and information theory to the study of the normal functioning of the central nervous system and abnormalities caused by disease processes. He is presently Assistant Professor of Neurology at Yale University School of Medicine, New Haven, Conn.

Dr. Stark is a member of the Biophysics Society of the American Academy of Neurology, the American Physiological Society, the American Electroencephalographic Society, and Sigma Xi.

❖

Walter E. Tolles (M'46–SM'56) was born in Moline, Ill., on February 1, 1916. He received the B.S. degree in physics from Antioch College, Yellow Springs, Ohio, in 1939, and the M.S. degree in physics from the University of Minnesota, Minneapolis, Minn., in 1942.

W. E. Tolles

From 1938 to 1942 he was a teaching fellow in physics, first at Antioch College, and later at the University of Minnesota. In 1942 he joined the Airborne Instruments Laboratory, Division of War Research, Columbia University, New York, N. Y., where he worked on the development of magnetic airborne detectors and electronic counter measures systems. From 1945 to 1955 he continued this work as supervisor of the transmitter section of Airborne Instruments Laboratory, Inc., Mineola, N. Y. From 1955 to the present he has been the Head of the Department of Medical and Biological Physics at AIL, where he has directed research on high-speed scanning systems and associated data reduction equipment, physiological monitoring systems, clinical instrumentation, and computer methods in medical diagnosis. He has served as expert consultant for the War Department and the National Institutes of Health.

Mr. Tolles is a fellow of the American Association for the Advancement of Science and the New York Academy of Sciences. He is a member of the American Physical Society, Inter-Society Cytology Council, the Biophysical Society, and Sigma Pi Sigma.

❖

Homer R. Warner (A'58) was born in Salt Lake City, Utah, on April 18, 1922. He received the B.A. degree in zoology in 1946 from the University of Utah, Salt Lake City. During this time he served for two years as a pilot in the Navy, and was discharged with the rank of Lieutenant (j.g.). He graduated from the University of Utah College of Medicine in 1949 with the M.D. degree, and interned at South Western University in Dallas, Tex., in 1949 and 1950. From 1950 to 1953 he served as resident in internal medicine and Fellow in physiology at the University of Minnesota Hospital, Minneapolis, and the Mayo Clinic, Rochester, Minn. In 1953 he received the Ph.D. degree in physiology from the University of Minnesota.

H. R. Warner

Following this, he returned to Salt Lake City, where he is now director of the Cardiovascular Laboratory of the Latterday Saints Hospital.

Dr. Warner is also an Established Investigator of the American Heart Association and an assistant research professor of physiology at the University of Utah College of Medicine.

❖

Vladimir K. Zworykin (M'30–F'38) was born in Mourom, Russia, in 1889. He received the E. E. degree from the Petrograd Institute of Technology in 1912, the Ph.D. from the University of Pittsburgh in 1926, and the D.Sc. from Polytechnic Institute of Brooklyn in 1938.

V. K. Zworykin

He came to the United States in 1920 and joined the research staff of Westinghouse Manufacturing Company. In 1930 he became associated with the Radio Corporation of America as director of the electronic research laboratory, and in 1947 was elected vice-president and technical consultant of the RCA Laboratories Division.

In 1954 he was named Honorary Vice-President. Since then he has continued as technical consultant to RCA Laboratories and, in addition, has directed the Medical Electronics Center at the Rockefeller Institute.

In World War II he served on the Scientific Advisory Board to the Commanding General of the United States Army Air Forces, the Ordnance Advisory Committee on Guided Missiles, and three subcommittees of the National Defense Research Committee, also directing important research work.

Dr. Zworykin received the IRE Morris Liebmann Memorial Prize in 1934 and the IRE Medal of Honor in 1951, as well as numerous other awards for basic contributions to television, electron microscopy, and other phases of radio and electronics. He is a Fellow of the American Institute of Electrical Engineers, the American Physical Society, and the American Association for the Advancement of Science, and a member of the National Academy of Science, American Academy of Science, Society of Motion Picture and Television Engineers, the Electron Microscope Society of America, and Sigma Xi.

Scanning the Transactions

Will pay checks become obsolete? The day may come when every payday a payroll machine in your office will "phone" the data processing center of the local bank to transfer funds from the company account into each employee's account, all automatically. Certainly automatic data processing, in one form or another, will play a key role in the banking system of the future, assisting with and reducing the tremendous volume of paper and accounting work that is now carried on. It is estimated that from 12 to 15 billion checks flow through the nation's banking system each year, and that each check is handled on an average of 5 to 7 times during its journey. This journey is complicated by the fact that a check drawn on one bank may be deposited in any one of the other 14,000 banks in the system and must be able to flow back to the originator of the check. Added to this are the differences in the forms and sizes of checks which create problems, not only under the manual system of today, but in applying automatic electronic processing techniques in the future. Nevertheless, there are two primary areas where the assistance of the electronics industry is sorely needed by the banking industry: 1) machines for processing the tremendous volume of paper automatically and with much greater speed and capacity, and 2) techniques for eliminating as much paper as is feasible, especially checks. We are beginning to see some progress on the first item, but all too little has been accomplished with respect to the second. (F. B. Miller, "Automation trends in the banking industry," IRE Trans. on Industrial Electronics, July, 1959.)

How deep the moon? Current efforts to place a satellite in orbit around the moon will be rewarded, it is hoped, by providing us with much valuable information about our lunar neighbor. We may even hope to learn more about the origin and age, not only of the moon, but of the earth and other parts of the solar system. Considerable thought is, therefore, being given to the most profitable types of experiments that might be performed with a moon-circling vehicle. One interesting suggestion that has been made is that if we equipped the satellite with a radar we could calculate the depth of the surface layer of the moon by measuring the reflected power at different wavelengths. The method is based on the fact that the power reflected at different wavelengths by a layered structure is a function of the number and depth of any layer present, and of the electromagnetic constants associated with this layer. These quantities could be calculated from power reflection measurements telemetered to earth from the satellite. (W. E. Fensler, T. B. A. Senior, and K. M. Siegel, "Exploring the depth of the surface layer of the moon from a space radar observatory," Abstract of an URSI paper, IRE Trans. on Antennas and Propagation, July, 1959.)

Transatlantic dialing is being studied by AT&T and in all probability will be added to their overseas public telephone service in the not-too-distant future. Experience in the United States, where long distance dialing is now coming into wide use, has shown that a dialing system enables calls to be handled more speedily and accurately. The resulting improvement in efficiency is of substantial interest to the overseas telephone system because the volume of overseas messages, which nearly doubled just in the last five years, is increasing so rapidly. In adapting a dial signaling system to overseas use some new problems must be faced. Although a substantial volume of transatlantic traffic is handled by submarine cable, high-frequency radio continues to be the major method of transoceanic communication. Multipath fading encountered on radio circuits imposes more stringent requirements on the signalling system. However, a system meeting these requirements has been developed, and is now being used successfully on radiotelephone circuits to Hawaii. There remain a number of problems in introducing dial signalling to foreign countries, but when these are resolved it is expected that overseas dial signalling will be adopted. (D. D. Donald and T. A. Chandler, "Global public telephone service—1958," IRE Trans. on Communications Systems, June, 1959.)

A micro-micro-microampere may seem like a current too minute to bother measuring, much less to be able to measure. Nevertheless, a vacuum tube electrometer has been developed for measuring currents as low as 6×10^{-18} amperes, or about 37 ions per second. This current is so small that it takes from 15 to 25 minutes to measure it. The device is intended for measuring very low ion densities in gases. The current arises both from the diffusion of ions (including electrons) to a probe and from any relative motion existing between the probe and the gas. The device is decidedly more than a laboratory curiosity; it should prove most useful in connection with upper atmosphere and space exploration programs that are currently the center of so much attention. (R. L. Ramey and R. L. Overstreet, "Electrometer measurements of very low ion densities in gases," IRE Trans. on Instrumentation, September, 1959.)

Optical telemetry systems are being seriously studied by communications experts in their quest for a telemetry system for hypersonic vehicles that won't black out during re-entry into the earth's atmosphere. Thermally ionized gas in the shock wave surrounding the re-entry vehicle absorbs signals over the major portion of the radio spectrum, causing a temporary communication blackout during a part of the re-entry flight. The absorption is most pronounced in the VHF frequency range allocated to standard telemetry systems. As a result, efforts are being made to develop systems in the HF and millimeter ranges where conditions are somewhat more favorable. However, the infrared and visible light regions of the spectrum appear to offer more promising advantages because optical radiation is entirely free of plasma absorption. There still remain problems of background noise, in the form of diffused sky light and optical radiation emitted by the shock wave, and adequate detector sensitivity. However, the rapid strides being made in photoconductor detectors and the availability of electronic filtering techniques for reducing noise in the detector output make optical communication a distinct possibility in the future. (E. Langberg, "Optical communication during hypersonic re-entry," IRE Trans. on Communications Systems, June, 1959.)

The latest in high-speed printers for data processing and communications equipment employs an electrostatic technique to produce reading material at a rate of 10,000 words per minute or more. Printing is accomplished by rows of closely spaced pins positioned near the surface of the recording paper. When a voltage is applied to appropriate pins, an electrostatic image of letters, in the form of closely spaced dots, is formed on the paper, to which ink particles will adhere when the paper is fed through an ink bath. The speed of print-out equipment has long lagged far behind the speed with which electronic data processing equipment has been able to furnish information. It appears that electronics has now had to take over the printing operation, too. (K. M. Kiel, "High-speed terminal printers using the Burroughs electrostatic technique," IRE Trans. on Communications Systems, June, 1959.)

Noise measurement in amplifiers using negative-resistance devices poses a problem because the usual methods of measurement have failed when used with such devices. The growing use of amplifiers using semiconductor diodes in a negative-resistance mode has stimulated the development of new measurement techniques at the Solid State Electronics Branch of the U. S. Naval Research Laboratory. By using the idea of exchange power as proposed by Haus and Adler, a measurement system has been successfully used wherein a passive linear 4-pole having only thermal noise is combined with the negative-resistance amplifier to provide a composite amplifier having positive resistance characteristics. By making measurements on this amplifier and using certain calculation procedures, values of noise figure and gain for the negative-resistance amplifier can be obtained. Thus our methods of measurement have been extended into a new region by the careful application of some mathematical and physical ideas. (A. Brodzinsky and A. C. Macpherson, "Noise measurement of negative resistance amplifiers," IRE TRANS. ON INSTRUMENTATION, September, 1959.)

Books

Les Principes de la Théorie Electromagnetique et de la Relativité, by Marie-Antoinette Tonnelat

Published (1959) by Masson et Cie, Éditeurs, 120 Boulevard Saint-Germain, Paris, VI^e, France. 386 pages+2 bibliography pages. Illus. 6¼×9¾. 5.000 fr. (In French.)

The author's goal is to describe the principles upon which electromagnetic and special and general relativity theories have been established. It gives a very clear account of the transition from the classical Maxwell field equations to the latest developments in field theory, including the general relativity problem and recent attempts at unifying electromagnetic field theory and gravitation into one basic set of equations as well as resolving the duality between the particle and the field concepts.

This book is destined for post-graduate students. After a brief derivation of the basic Maxwell equations, the author shows how these equations lead to the Einstein special relativity concept and how fruitful this has been for the explanation of many experimental facts such as Doppler effects, Ives and Stilwell experiments, lifetime of mesons, equivalence of mass and energy and many others.

The third part of the book is devoted to general relativity. The reviewer has found that the introduction to this complex subject is very enlightening. He feels that the book can be highly recommended to all those who seek a guide for their study of the modern approach to relativistic theories and their consequences, such as the red shift of spectral lines and the deviation of light in gravitation fields.

The mathematics used include tensors and four dimensional and noneuclidean geometries. A discussion of these mathematical tools is given in a separate part of the work so that the reader is not interrupted in his study of the main subject if he is already familiar with this kind of symbolism.

The reviewer thinks that the value of the book is not only in a clear and progressive introduction to theoretical subjects of fundamental interest, but also in the fact that it contains thought provoking discussions of the nature of the problems involved which are couched in simple language and which are conducive to re-examination of principles and historical trends in scientific theories. In this respect the book is more than purely didactic or tutorial. It will be found very helpful as an introduction to the latest attempts toward a unitized theory of the electromagnetic and gravitational fields.

Each chapter is followed by a series of problems with clues leading to their solutions. The reviewer has found them very well chosen. The bibliographical references, however, have been kept to a minimum and it might be useful to have them extended in a new edition of the book.

A. G. CLAVIER
ITT Labs.
Nutley, N. J.

Control Engineering, by Gordon J. Murphy

Published (1959) by D. Van Nostrand Co., Inc., Princeton, N. J. 372 pages+5 index pages+8 appendix pages+xii pages. Illus. 6×9¼. $7.50.

Dr. Murphy has written this book to meet the need for "a text in which both the elementary and the advanced topics are treated carefully and in some detail at a mathematical level which will challenge the good student without unnecessarily interfering with his efforts to learn." In this effort the author has been successful. He has selected topics of basic value in the theory of servomechanisms and process control, and has treated each concisely and logically.

The basic theory treated in virtually every modern elementary book in the field is disposed of in the first five chapters. In these chapters are introduced the root locus, time domain, and frequency response approaches. Then, one chapter is devoted to each of such advanced topics as linear systems with dead time, modulated carrier (a-c) control, sampled-data systems, and noise. Finally, nonlinear control systems are analyzed by describing functions and, in the final chapter, by the phase plane. In this book the treatment of design is largely mathematical in nature, and is limited to s-plane and frequency-response design of linear continuous systems and modulated carrier systems. However, there is a chapter giving a brief treatment of typical control system components.

Given the sensible limitation of 372 pages, it is difficult to suggest any more logical and connected sequence of material that could better fulfill the aim of a textbook of this kind, although obviously much important material could not be included. The author states that the book is for use at junior or senior level. While doubtless a junior could be led through this book, it is doubtful whether he possesses the engineering background to gain much in the way of "Control Engineering" from such early study of it. However, a senior or young graduate working in allied fields, having a good mathematics and electric circuit theory background, and knowing something of the practical aspects of electric machines, electronic circuits and feedback, could rapidly gain from this book the material needed for an up-to-date acquaintance with the control engineering field. There are good problems given at the end of each chapter.

Considering the brevity of coverage, one could wish the text were better annotated as to references. Some are given in footnotes, and the original papers and standard textbooks pertinent to the chapter are listed at the end of each. But references to more advanced and off-shooting topics are for the most part omitted. On the other hand, in the economy of treatment in this book, Dr. Murphy has successfully avoided verboseness, and has in great part made explicit the too-general mathematical formulas which weighed down the student of his earlier book, "Basic Automatic Control Theory," by the same publisher.

FREDERICK A. RUSSELL
Newark College of Engrg.
Newark, N. J.

An Approach to Electrical Science, by Henry G. Booker

Published (1959) by McGraw-Hill Book Co., 330 W. 42 St., N. Y. 36, N. Y. 727 pages+12 index pages +84 problems pages+xviii pages. Illus. 6×9. $9.50.

This book, according to a note on its wrapper, "transcends conventional lines of demarcation between electrical engineering, physics, and mathematics." Its objective is to present an integrated approach to electrical science as represented by a thorough and careful treatment of the topics usually covered in the elementary physics course in electricity and magnetism, plus those in the elementary electrical engineering courses in circuits and fields. Thus, it is intended to provide in one year the basis necessary for 1) the study of electric networks, including any combination of inductors, capacitors, resistors, and vacuum tubes, or 2) the study of transmission lines, or 3) the study of electromagnetic theory.

In presenting each topic the author has used what he regards as the "best" approach, whether that be the approach of the physicist, the electrical engineer, or the mathematician; or a composite, unconventional approach. More often than not, the point of view turns out to be that of the physicist.

The book is divided into four parts as follows:

Part 1—Electric Fields and Circuits in Electrostatics.
Part 2—Steady Electric Currents and Vacuum Tubes.
Part 3—Magnetic Fields and Circuits.
Part 4—Relations between Electric and Magnetic Fields.

Part 1 begins with Coulomb's Law, runs through the main ideas of static electric fields, develops a circuit theory applicable only to capacitance networks and concludes with electrostatics of transmission lines. Along the way, Kirschhoff's laws, equivalent T and π networks, Thevenin's and Norton's theorems, as well as other familiar topics, are modified for special application to networks of capacitors alone.

A similar organization is used for the other parts. A conspicuous feature is that an elementary circuit theory is developed separately for capacitors, for resistors, and for inductors. This presentation possesses the virtue of a certain logic, but it suppresses the elegant unity of electromagnetic theory and results in many wasted words about circuits. Such topics as equivalent T and π networks and Thevenin's and Norton's theorems are discussed in three different contexts, each without reference to the others.

A studied effort is made to provide an intuitive approach to each new topic. This attempt is apparently what is responsible for repetitive explanations that are much more detailed than is really necessary, and for circuitous reasoning that leaves the reader floundering.

The book makes its main contribution as an introductory course in electrostatics. This subject is indeed given a more scientific status than the conventional elementary physics course, and it is carried to the point where a reasonably sophisticated treatment of electromagnetic theory can begin. As an introduction to circuit theory it leaves much to be desired.

B. J. Dasher
School of Elec. Engrg.
Georgia Inst. Tech.
Atlanta, Ga.

Production and Field Reliability, edited by the Technical Publications Committee

Published (1959) by the American Society for Quality Control, Electronics Division. 294 pages +11 appendix pages+1 bibliography page. Illus. 6×9. $3.50. Paperbound.

This handbook presents a survey of the problems and methods of achieving product reliability through a judicious selection of fourteen papers from the myriad that have been published in diverse periodicals, proceedings, and documents. It is intended for the engineer who has little, or no, background in formal reliability study, and takes him through the basic principles and theory, measurement and improvement, of reliability in production, then through the means of utilizing field failure and maintenance information in the reliability program.

The selection of papers is good. None are too theoretical to be understood by a neophyte to the subject, and, at the same time, each author presents his topic from a broad enough viewpoint to give the reader a real appreciation of the scope and importance of the problem.

The format of the book being a collection of papers, it is obvious that it is neither an ideal textbook nor reference book. However, it contains enough material for either, and may be used to fulfill such a need during the current period when no suitable text or reference books have yet reached the market.

In the opinion of the reviewer, this book is an excellent introduction to a topic which has achieved major importance in today's defense picture, and which is becoming increasingly important in the commercial field. It should certainly be read by those directly concerned with, or about to be concerned with, any reliability program. It is probably the easiest and fastest means available today by which a beginner can familiarize himself with the concepts, methods, and terminology encountered in the reliability literature.

Leon Bass
Manager JED Quality Control
General Electric Co.
Cincinnati 15, Ohio

Solid State Magnetic and Dielectric Devices, edited by Harold W. Katz

Published (1959) by John Wiley and Sons, Inc., 440 Fourth Ave., N. Y. 16, N. Y. 500 pages+6 index pages+xxi pages+36 appendix pages. References each chapter. Illus. 6×9. $13.50.

This book, which includes contributions by a number of authors specialized in their individual field, serves a need in the field of technical literature on solid state magnetic and dielectric devices. It provides a thorough introductory treatment of the theory and application of some new solid state materials, such as ferrites and titanates, used for device applications. It is a practical and authoritative book which should be a valuable aid both to the electronic engineer and those in other fields who are interested in the design and operation of solid state devices.

The book begins with a review of electrostatic and magnetostatic field theory. This is followed by a classical treatment of magnetization and polarization and the application of these principles to electromechanical phenomena. The last half of the book starts off with a basic introduction to the magnetic properties exhibited by ferrites at microwave frequencies, including a derivation of the basic tensor permeability and application of the Faraday effect for obtaining isolators, switches, circulators, etc. This is followed by separate chapters describing magnetic and dielectric amplifiers, digital techniques using square loop materials and magnetic recording. The last chapter describes present measurement techniques used to obtain the dielectric constant and tensor permeability of ferrites at microwave frequencies, the measurement of electrostrictive and magnetostrictive constants of ceramic materials, and the design of a hysteresigraph for the measurement of square loop materials.

An extensive bibliography is given at the end of each chapter. Also included are several appendixes giving brief treatments on the tensor dielectric constant of a plasma, magnetoresistance, and parametric devices.

In conclusion, this worthy contribution to the engineering literature could profitably be used with supplementary problems as a textbook in a one-semester course for senior students majoring in electrical engineering and interested in the application of electronic principles to special dielectric and magnetic devices.

F. Reggia
Diamond Ordnance Fuze Labs.
Washington, D. C.

Feedback Control Systems, by J. C. Gille, M. J. Pelegrin, and P. Decaulne

Published (1959) by the McGraw-Hill Book Co., Inc., 330 W. 42 St., N. Y. 36, N. Y. 718 pages+11 index pages+7 bibliography pages+18 glossary pages +xx pages+38 problems pages. Illus. 6×9. $16.50.

This is an unusually complete and well written book on feedback control systems. Although the original text, "Theorie et Technique des Asservissements," was written by three French authors in their native language, it has been adapted into a very clear English translation, and it has a rather cosmopolitan flair because the authors have drawn freely from significant American, English, French, Russian, and German papers. This is a particularly significant contribution because the features of the papers are not only well described, but they are often compared with one another to stress the advantages and disadvantages of each method.

Another fine feature of the book is its organization. Each chapter begins with a very brief summary which simply itemizes the main points to be discussed in the chapter. An introduction follows, then the subject is developed rather quickly, but clearly and surprisingly thoroughly for the large amount of material presented. The clarity of the book and its significance is enhanced by the philosophical discussions in each chapter which attempt to evaluate the material and put it in its proper perspective with respect to its practical and theoretical usefulness. Because this is done frequently, the reader is constantly reminded of his progress and the direction in which the text is advancing. Most of the chapters end with an appropriate concluding paragraph or section. Details and proofs of the important points in the development of a concept are given in half-size type. This has the advantages of emphasizing the important material and of providing extra space for additional subject matter.

The book is divided into five basic parts: I, Dynamics of Linear Systems (190 pages); II, Linear Servo Systems (159 pages); III, Nonlinear Servo Systems (129 pages); IV, Components of Servo Systems (195 pages); and V, Basic Design of a Servo System (21 pages). There is also a 20-page introduction, 38 pages of problems, 18 pages of an English, Spanish, French, German, and Russian Glossary of Automatic Control terminology (transliteration is used to write Russian words in Latin letters), and an 8-page selected bibliography.

The first part is concerned with well known and useful tools and terminology which are used in the field of the automatic control. Differential equations, analogs and duals are reviewed, transients are described, and the Laplace Transform is discussed. Then the significance of first and second order system characteristics is considered. This leads to the concept of transfer functions, the harmonic approach, Nyquist and Nichols charts, the pole-zero approach, stability, transfer materices, impedance matching, the linearity domain, and statistical considerations, including the fundamentals of information theory and its future in the control field.

The second part of the book is naturally derived from the material in the first part, which is combined with the concept of the feedback loop and its transfer functions. Several examples are given to illustrate the use of the Nyquist diagram, the Hall chart, and the Nichols' chart. Included are the all-important resonant frequency terms which are often neglected in textbook presentations. The root locus and its construction is also developed in detail, and it is compared with the frequency response approach. Various steady-state errors are considered, methods of determining stability mathematically and graphically are developed, and performance criteria are discussed. This leads to the subjects of compensation and synthesis, including the methods Guillemine-Truxal and the optimization methods of Wiener. The second section ends with a chapter on sampled-data systems and multiple feedback control systems and remarks on the general theories involved.

Part three begins with general remarks, definitions, and classifications of nonlinear elements and systems. It follows with a discussion of nonlinear transients. The describing function concept is then introduced with references to the works of Kochenburger (U.S.A.), Dutilh (France), and Goldfarb (Russia). Several examples are given and the validity and usefulness of the method is reviewed. The Poincaré approach is presented in a separate chapter of part three, and this is extended in the next chapter to the determination of oscillations in on-off control systems using the loci of Hamel and Tsypkin. Additional methods of analyzing nonlinear systems are given in another chapter including Poincaré's theorems, a discussion of limit cycles, singular points, and one of the most significant Russian contributions: the second method of Liapunov. Part three concludes with examples of linearization problems and final remarks.

Part four, essentially a 200-page textbook in itself, discusses servo system components: error sensing devices, the determination and design of servomotors both theoretically and practically (this is one of the most complete single discussions available in the literature), electric servomotors, hydraulic servomotors, and amplifiers and preamplifiers including a section on the design of transistor amplifiers. Practical considerations are always given, particularly with respect to microscopic and macroscopic noise and saturation. In general, the material is presented in more detail than it is in most of the American textbooks. This section ends with a table comparing 20 common characteristics of nine types of power amplifiers: relays, vibrating relays, vacuum tubes, transistors, thyratrons, magnetic amplifiers, Ward-Leonard systems, amplidynes, and hydraulic motors or actuators.

Part five includes two practical examples of basic system design: a control to automatically determine the bearing of an airplane with respect to a runway, and an autopilot control for a guided missile.

It should be noted that the book has a pocket on the inside back cover with several enlarged charts consisting of iso-power hyperboles, log amplitude response of first and second order systems, log phase response of first and second order systems, and a Nichols' chart.

The authors' note in the preface that the "... book is intended (a) as a servomechanism textbook for seniors and first-year graduate students, and (b) as a reference book for the engineer working in the automatic-control field. The aim of the authors is, by progressing from the elements of the subject, to provide in a single book a treatment both of the over-all theory of feedback control systems and of their components." Certainly, the authors have more than met their goal, and in doing this, they have produced a textbook which all serious students and engineers in the control field should obtain for the authors' European viewpoint, theoretical but tempered with practical considerations, and their unifying contribution, the practical philosophies, often neglected in textbooks on automatic control.

George S. Axelby
Westinghouse Air Arm
Baltimore, Md.

The Properties, Physics and Design of Semiconductor Devices, by John N. Shive

Published (1959) by D. Van Nostrand Company, Inc., 120 Alexander St., Princeton, N. J., 471 pages+13 index pages+1 appendix page+xxi pages. Illus. 6¼×9¼. $9.75.

This book is highly recommended for the author's avowed purposes, which are as follows:

> "In its general level of presentation, this text is directed toward advanced undergraduate and graduate university students. Particular attention has been paid to the development of descriptive pictures of the various electronic processes occurring in semiconductors, so that persons seeking to enter the semiconductor field as specialists may find this text a valuable introduction to more advanced physical and analytical study. In level, depth, and topical content this book attempts to bridge the gap between elementary expositions of the semiconductor story and the highly theoretical treatments found, for example, in Seitz's *Modern Theory of Solids* or Shockley's *Electrons and Holes in Semiconductors.*"

The author's writing is lucid, often entertaining and has a personal flavor that the publisher has (thankfully) not edited out. In developing the subject matter, physical reasoning is given in substantial detail, with much healthful exercising of the reader's intuitive grasp of the situation. Mathematical derivations are eschewed in the first half of the book, which presents "a discussion of semiconductor devices from an empirical point of view.... The several devices, their fabrication, properties, characteristics, and typical applications [are treated] with just enough background material on fundamentals to make the operation of the devices descriptively understandable." It is a pleasure to read such material when presented by a competent writer who evidently knows the physical theory well, and is simply selecting the choice morsels from the carcass for you, the reader, to chew on.

As might be expected from John Shive, the discussion is not at all restricted to transistors. Thermistors, varistors (the term includes diode rectifiers as well as symmetrical elements such as thyrite and silicon carbide), semiconductor photoelectric cells, and odd-ball devices, as well as transistors, get sympathetic treatment.

The solid-state electronics of semiconductors is discussed in Part II of the book, with careful attention to correctness in the theory and a lean (but reasonably muscular) use of mathematics. Mathematical results seem to be considered more worthy of physical justification than of detailed derivation from some other obscure equation. This point of view will be applauded by many readers, I am sure. The discussion progresses from the band theory of solids through conduction mechanisms, hole and electron statistics, and the continuity equation ("whatever goes up comes down") to *p-n* junction theory, metal-semiconductor contact theory, and the junction triode. Then there are lucid chapters on the theory of optical properties, thermoelectric effects, and the Hall effect, and finally a discussion of topics in device fabrication technology.

The penalty paid for the fluent, clear style of the book is that the amount and depth of material covered in 471 text pages is less than in a more conventionally dense textbook. A worthwhile penalty, in this case.

H. E. Tompkins
The Moore School of Electrical Engineering
Philadelphia 4, Pa.

Audio Measurements, by Norman Cowhurst

Published (1958) by Gernsback Library, 154 W. 14 St., N. Y. 11., N. Y. 212 pages+5 index pages+vi pages. 1029 Figs. 5½×8½. $2.90 softcover, $5.00 hardcover.

This book contains a very complete review of the instruments and practical techniques used in making measurements at audio frequencies. It also points up some of the fallacies of system specification based upon inadequate or misleading measurements. Although it was written primarily for the hobbyist and technician, most engineers will find the practical hints very helpful. A lack of detail and theory, necessitated by reasonable size and the intended audience, has been largely compensated for by the inclusion of a bibliography at the end of each chapter.

Chapters include: Measurement Techniques, Test Equipment, Basic Measurements, Basic Amplifiers, Output Transformers, Preamplifiers, Pickups and Arms, Turntables and Changer, Tape Recorders, and Microphones. The index is complete and well cross referenced.

Roy A. Long
Stanford Research Institute
Menlo Park, Calif.

Programming Business Computers, by D. D. McCracken, H. Weiss, and Tsai-Hwa Lee

Published (1959) by John Wiley and Sons, Inc., 440 4th Ave., N. Y. 16, N. Y. 428 pages+8 index pages+14 bibliography pages+39 appendix pages +19 glossary pages+xvii pages. Illus. 6×9¼. $10.25.

This book presents a thorough coverage of the principles and techniques of programming for data processing applications. The discussion of computers and their characteristics is well developed and affords a good primer for persons acquainted with no or only one machine. The development of the programming art through the remainder of the book is done with ease and continuity.

Especially useful to the new programmer and provocative to the experienced are the chapters on loops, subroutines and program verification. The last of these chapters and those on audit trail and systems design should be of special interest to executives and managers.

The book presents one of the best coverages for both programmers and associated fields that is available in one volume form. It reads easily and can be used as a text for basic orientation classes.

V. E. Henriques
H. J. Bracken, Jr.
National Bureau of Standards
Washington, D. C.

Technology of Printed Circuits, by P. Eisler

Published (1959) by Academic Press, Inc., 111 5th Ave., N. Y. 3, N. Y. 320 pages+6 index pages +50 bibliography pages+29 appendix pages+x pages. 151 figures. 5¾×8¾. $12.00.

This volume by P. Eisler is an authoritative and comprehensive treatise on the subject of foil type "printed circuits" written for a broad audience—the engineer, technician and designer. It contains discussions which should satisfy both the expert in the field and the layman who desires information on a relatively new technology. The wealth of illustrations, tables, graphs, and detailed procedures aid understanding the subject, and the extensive bibliography and appendixes should be most helpful for those who wish additional detail. Even though there are some instances of undue repetition, some subjects too finely subdivided and some chapter arrangements disorganized, the volume is a good reference due to its comprehensiveness.

The following listing of the chapter headings gives an idea of the comprehensive nature of the volume:

1. The Basic Story of Printed Circuits
2. The Confused Start of a New Technology
3. The Principles of Method Selection and the Foil Technique
4. Watching Printed Circuit Production
5. From the Drawing to the Photo-Stencil
6. Litho Printing
7. Photomechanical Printing Processes
8. Plating and the Foil Technique
9. Ancillary Operations
10. Laboratory Routine
11. Design
12. Conventional Components for Assembly With Printed Circuits
13. Printed Circuit Trouble Shooting
14. Automatic Assembly in the Electronics Industry
15. Regeneration of Ferric Chloride
16. Weight Saving and Heat Dissipation
17. Miniaturization
18. Printed Components
19. Microwave Printed Circuits: A Short Survey
20. Printed Circuits and the Electronics Industry

Appendix 1: Examples of Commercially Available Adhesives for Foil Bonding
Appendix 2: The Control of Strength of Etching Solutions
Appendix 3: Double Layer Process
Appendix 4: Plating Solutions and Some Principles of the Automatic Plating Plant
Appendix 5: Description of RCA Programmed Punching Machine
Appendix 6: Approximate Analysis of the Characteristic Impedance of Flat-Strip Transmission Lines
Appendix 7: Characteristics of Flat-Strip Transmission Lines of Importance in Applications

In general, the first two chapters cover the historical background of printed circuits. The third chapter covers general descriptions of well-known printed circuit fabrication methods. The fourth chapter contains a description of printed circuit production methods. Chapters 5 through 9 cover in detail the many techniques used for the fabrication of printed circuits and some general assembly operations for the fabrication of printed circuit assemblies. Chapter 10 appears to be a bit out of place and describes various laboratory techniques and equipment necessary for the fabrication of printed circuits. Chapter 11 discusses considerations for the design and layout of electronic circuits for use of printed circuitry techniques. Chapters 12 and 14 cover various methods for the assembly of electronic parts to the printed circuit boards. Chapter 13 covers trouble shooting of printed circuits both on the production line and in the field. Chapter 15 covers details of regenerating the ferric chloride etchant for the etched circuit process. Chapters 16 and 17 cover advantages to be gained by printed circuit techniques such as weight saving and miniaturization. Chapter 18 discusses printed electronic circuit elements such as resistors. Chapter 19 covers the fabrication of microwave circuits by the foil technique. Chapter 20 deals with developmental possibilities presented by the new technology of printed circuitry.

L. K. Lee
Space Technology Labs.
Los Angeles, Calif.

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*
Antennas and Propagation	AP-7, No. 3	$1.75	$2.60	$5.25
Communications Systems	CS-7, No. 2	1.85	2.75	5.55
Industrial Electronics	PGIE-10	1.20	1.80	3.60
Instrumentation	I-8, No. 2	.65	.95	1.95
Production Techniques	PGPT-6	0.80	1.20	2.40

*** Libraries and colleges may purchase copies at IRE Member rates.**

Antennas and Propagation

Vol. AP-7, No. 3, July, 1959

A Variational Expression for the Terminal Admittance of a Semi-Infinite Dielectric Rod— C. M. Angulo and W. S. C. Chang (p. 207)

The reflection that the abrupt termination of a semi-infinite circular dielectric rod produces on the $TM_{0,1}$ surface wave is determined here by calculating the terminal admittance of the surface wave at the end of the rod. The semi-infinite dielectric rod is regarded as an open waveguide partially filled with dielectric. The half space at the end of the rod is analyzed as a homogeneous open waveguide. The complete sets of proper and improper eigenfunctions are found for both waveguides and used to represent the transversal fields. Finally, a variational expression is set up for the admittance and an an approximate value is obtained using as trial field the pure $TM_{0,1}$ surface wave.

Radiation from Slot Arrays on Cones—R. F. Goodrich, *et al.* (p. 213)

A method is obtained for determining far field patterns, sidelobes as well as the main beam, for an array of slots on the surface of a cone. It is found that accurate results can be obtained for a single slot by using geometric optics for the main beam and an extension of Fock theory for fields in the shadow region. The tip contribution is computed by physical optics and, for reasonably thin cones, is found to be negligible. The array pattern is obtained by appropriately summing the single slot fields. To test the validity of the method and to test the ease with which computations could be performed, a radiation pattern from a linear array of 65 slots on the surface of a cone was computed and compared with experiment. The agreement is excellent.

The major theoretical part of this paper is the generalization and simplification of Fock theory as applied to the surface of a cone.

A Study of Spherical Reflectors as Wide-Angle Scanning Antennas—T. Li (p. 223)

A study is made of spherical reflectors for use as wide-angle scanning antennas. In order to keep the effects of spherical aberration within tolerable limits, the approach of using a restricted aperture is adopted. This approach is suitable for applications requiring very wide angles of scan.

Experimental results show that the phase error over the illuminated aperture of a spherical reflector should not exceed one-sixteenth of a wavelength. This requirement determines the beamwidth of the primary source. A square-aperture horn with diagonal polarization is found to satisfy the requirements of a suitable feed for the reflector. Secondary patterns of a 10-foot-diameter hemispherical reflector illuminated by this horn at 11.2 kmc have a 3-db beamwidth of 1.76° and a relative sidelobe level of about −20 db throughout a total useful angle of scan of 140°. The measured gain is 39.4 db, which is equivalent to the gain of a uniformly illuminated circular aperture of 31-inch diameter.

Analysis and Reduction of Scattering from the Feed of a Cheese Antenna—W. A. Cumming, C. P. Wang, and S. C. Loh (p. 226)

The far field of a cheese antenna can be described in terms of three components: 1) one due to the unobstructed aperture field, 2) one due to that portion of the primary feed energy not intercepted by the reflector, and 3) a scattered component due to the fact that the feed acts as an obstacle in the path of energy emerging from the reflector. This scattered component is usually calculated by considering a perturbation in the aperture field, in the form of an out-of-phase component sufficient to produce zero field in the geometrical shadow region behind the feed. A more exact analysis of a reflector excited by a longitudinally-slotted circular cylinder shows this engineering approach to be satisfactory, provided the out-of-phase component is assumed to exist over an area of width 1.5 times the projected width of the feed. An empirical investigation shows a similar result for a slotted rectangular waveguide feed and for a horn feed. An antenna is described in which the scattered field and the back-lobe of the primary feed are made to partially cancel. An additional control over these field components is provided by a series of vanes or waveguides located either side of the feed.

On the Phase Velocity of Wave Propagation Along an Infinite Yagi Structure—D. L. Sengupta (p. 234)

An approximate expression is derived for the phase velocity of wave propagation along an infinite Yagi structure and its dependence on the various parameters of the structure is discussed in detail. The structure is at first studied qualitatively by applying the transmission line analogy. The problem is next treated from the viewpoints of linear antenna and field theories. It is assumed that a traveling wave is propagating along the axial direction and it induces an axially symmetric and sinusoidal current distribution in each element. The electric field at any point due to this current distribution is calculated by the Hertz vector method. After applying boundary condition to the electric field, an expression for the propagation constant is derived. The results are compared with existing experimental values. The agreement between theory and experiment is found to be within 5 per cent. The accuracy of the expression given is sufficient for practical purposes.

Effect of Relatively Strong Fields on the Propagation of EM Waves Through a Hypersonically Produced Plasma—W. B. Sisco and J. M. Fiskin (p. 240)

The simple classical theory employed in the analysis of electromagnetic waves propagating through an ionized gas is not sufficiently general to take into account the variation of the complex conductivity of the plasma with the magnitude of the impressed EM field. Problems of this nature arise when it becomes necessary to transmit radar signals of high energy density through the ionized shock wave produced by a hypersonic vehicle.

The exact theory of conductivity developed by Margenau is for impressed, relatively high field strengths too difficult to handle from an analysis and computational standpoint. By making two simplifying assumptions in the general velocity distribution function and graphically interpolating between them, two relations, one nearly exact and one employing the simple theory, are obtained for the conductivity. The accuracy of these relations is then examined analytically for a typical case, and graphical comparisons between the methods are made.

Effects of neglecting Coulomb interactions and higher order components in the velocity distribution function are considered briefly. For an example, the complex conductivity of a typical ionized shock wave as a function of field strength and frequency is calculated and plotted.

Microwave Scattering by Turbulent Air—C. E. Phillips (p. 245)

A system for measuring the power scattered by a region of thermally turbulent air is described, and its performance is analyzed. Also described is the equipment used to generate the turbulent region and the temperature fluctuations therein.

The experimental procedure can be summarized by stating that it consisted in measuring the fluctuations of the null in the interference pattern of an antenna array composed of two antennas excited out of phase when thermally turbulent air, of known properties, was interposed between the receiving and transmitting systems.

The measured and calculated receiver response due to scattering by the turbulent air is compared, and it is noted that they agree at least within an order of magnitude.

The experiment described shows that a measurement of the angular dependence of the signal scattered by turbulent air is within the possibilities of available techniques, although a more elaborate measuring system and higher microwave power levels would be required.

Correction to "Radiation from Ring Sources in the Presence of a Semi-Infinite Cone"—L. B. Felsen (p. 251)

Influence of an Atmospheric Duct on Microwave Fading—F. Ikegami (p. 252)

The results of continuous observation of a duct carried out utilizing a tower 312 meters high are presented together with those of measurements of microwave fading conducted simultaneously.

The variation of duct height with time, as well as the influence of the duct on fading for a horizontal and an oblique propagation path, are investigated in detail.

A ray-theoretical analysis is given, indicating that fading may be attributed to the divergence or the convergence of radio waves and to the interference of two or more rays, caused by existence of a duct, or, more generally, of nonlinear M-profile.

A comparison of calculation with experimental results shows that many of the characteristics of microwave fading are well explained by means of this interpretation.

Comparison of Computed with Observed Atmospheric Refraction—W. L. Anderson, N. J. Beyers, and B. M. Fannin (p. 258)

Ray tracing methods have been applied in the computation of atmospheric refraction for a path at White Sands Missile Range, N. M., with a range of about 48 miles and an elevation angle of 14.5 milliradians. The atmosphere was assumed to be horizontally stratified. Refractive index profiles were derived from meteorological data obtained from surface observations, wiresondes, radiosondes, and airborne refractometer soundings. The profiles were classified "A," "B," or "C," in descending order of reliability, prior to radar refraction computations The classification system considered the variety of data available, the time lag between radar and weather observations, and the proximity of the sounding to the propagation path. A good correlation between observed and computed angles resulted and the correlation was directly related to the classification.

Radar observations were made in the X band and instrumental precision maintained to within 0.25 milliradian. Total bending ranged between 0.56 and 2.23 milliradians, with standard deviation 0.38 milliradian.

The rms deviation of computed from observed angles ranges from 0.19 to 0.41 milliradian for Class A and Class C data, respectively. The correlation coefficient ranges from 0.81 to 0.13. It is concluded that within the limits of this experiment: 1) ray tracing methods are justified, 2) horizontal stratification may be assumed and 3) the accuracy of bending predictions is increased by improving the meteorological data.

Diffraction Theory of Tropospheric Propagation Near and Beyond the Radio Horizon. Part I—Theory; Part II—Comparison with Experiments—O. Tukizi (p. 261)

By means of the earth-flattening approximation, the problem of radio diffraction by the earth is treated as that of refraction in an atmosphere with a linear profile of modified index. Use is made of the saddle point method, whose special case, known as the method of stationary phase, proves useful for analysis of the normal propagation within the horizon.

It is found that the classical diffraction theory is as valid at ranges beyond the horizon as in its neighborhood, if account is taken of the contribution of other terms than the first of the residue series, and that exclusive use of the saddle point method makes it possible to deal systematically with the tropospheric propagation for all regions within, near, and beyond the horizon.

Theoretical derivations are presented in Part I. Part II gives a comparison of the theory with experiments, showing a fairly good agreement between them.

Communications

A Note on Surface Waves Along Corrugated Structures—L. O. Goldstone and A. A. Oliner (p. 274)

Comments on "Scanning Surface Wave Antennas—Oblique Surface Waves Over a Corrugated Conductor"—R. E. Collin (p. 276)

The Filling in of an Antenna Hull by Off-Path Scattering on a Tropospheric Scatter Circuit—H. Staras (p. 277)

The concept and characteristics of Antenna-to-Medium Coupling loss have already been discussed in the literature in some detail. These earlier analyses show that the coupling loss is not significant unless the beamwidth of the antenna becomes equal to or less than the angular width of the scatter volume as seen from the antenna. An inverse phenomenon occurs if one tries to design an antenna null to protect a given location beyond the horizon; namely, if the width of the null is narrower than the angular width of the scatter volume, the null can be filled in by off-path scattering and the desired protection might not be achieved. This paper outlines the analysis used in evaluating the "filling in" of the antenna null by off-path scattering and presents curves which should prove useful in obtaining quantitative estimates.

An examination of the interrelation between the scattering process and the free-space antenna pattern reveals that the scattering process averages the free-space antenna pattern, the averaging interval being of the order of several degrees. Thus a qualitative conclusion can be drawn that if it is desired to protect an installation with a 40 db null, this null must be maintained over at least several degrees in order to be effective. A cosine squared pattern is not satisfactory.

An Investigation of the Complex Mutual Impedance Between Short Helical Array Elements—A. R. Stratoti and E. J. Wilkinson (p. 279)

Gains of Finite-Size Corner-Reflector Antennas—E. F. Harris (p. 281)

A Method to Achieve a Collimated Circularly Polarized Beam—C. L. Gray and J. C. Huber, Jr. (p. 281)

Abstracts of Papers from the IRE-URSI Symposium Held May 4–7, 1959—Washington, D. C. (p. 283)

Contributors (p. 304)

Communications Systems

Vol. CS-7, No. 2, June, 1959

1958 IRE-PGCS Awards (p. 65)

Frontispiece and Guest Editorial—H. P. Westman (p. 66)

Optical Communication During Hypersonic Re-Entry—E. Langberg (p. 68)

Communication during the re-entry of a space vehicle is a formidable problem because thermally-ionized gas in a shock wave surrounding such a vehicle absorbs the signal over the major portion of the radio spectrum. This paper describes some general features of re-entry communications. The entire electromagnetic spectrum is examined for selection of a suitable signal frequency.

An Experimental Automatic Communication System for Air Traffic Control—W. R. Deal (p. 71)

The "Experimental Automatic Ground/Air /Ground Communication System (AGACS)" under development for the Bureau of Research and Development, Federal Aviation Agency, will provide a two-way traffic control data link permitting automatic mechanized communication for essential routine flight information between the air and ground environments. The system selected is operable in both VHF and UHF aircraft communication bands and permits choice of FSK-carrier or FSK-AM modulation. Up to 500 aircraft may be interrogated and reply on a single channel in a time period of two minutes or less. Information is binary coded and transmitted by time division multiplex techniques at a data rate of 750 bits per second.

Factors Affecting Modulation Techniques for VHF Scatter Systems—J. W. Koch (p. 77)

An experimental program at approximately 50 mc has been carried out over a 1295 km east-west ionospheric-scatter path to determine the communication capacity of the propagation medium using ordinary modulation techniques. Binary error rate studies using dual-narrow-band frequency-shift-keyed terminal equipment were made to observe the dependence of error rate in scatter propagation on signal-to-noise ratio, multipath factors, and Doppler shifts. With respect to multipath limitation, using antennas of 6° beamwidth, error rate is found to be independent of transmission speed to 500 binary digits per second. Tests were made using dual-diversity narrow-beam rhombic antennas, and dual- and quadruple-diversity broad-beam Yagi antennas. The use of four Yagi antennas gave very promising results as compared to use of two narrow-beam antennas. Frequency shifts of 2, 4, and 6 kc were used to determine the effect of meteor Doppler components in the received signal, leading to recommendations for minimum frequency shift. Though actual modulation tests have so far been made at a frequency of 49.6 mc only, systematic observations of long-delay signals of the order of 10 to 80 msec were made at 30 and 40 mc, using pulse techniques. These delayed signals arrived at the receiver via the F_2 backscatter mode, and were occasionally of sufficient strength to cause binary errors for circuits operating in these frequency ranges. Studies of intelligibility for voice transmission were made using narrow-band frequency modulation and single-sideband modulation. Using 20 kw power, the frequency modulation tests with dual-diversity receivers and narrow-beam rhombic antennas gave good results at levels of signal-to-noise ratio exceeded for 90 per cent of the year at 50 mc.

Double-Sideband Suppressed-Carrier Multiplex Equipment for Cable and Microwave Applications—W. S. Chaskin and G. L. Curtis (p. 92)

Single-sideband suppressed carrier modulation has been the backbone system for telephone multichannel carrier equipment for many years. However, particularly in recent years, a number of other modulation systems have been applied in lesser or greater degrees with the emergence of new problems in telephone transmission. These problems are related to economics, bulk of equipment, power consumption, phase distortion as it is related to data, etc.

Integrated Data Systems—R. Filipowsky (p. 95)

Two theoretical data systems (Shannon's ideal channel with average power limitation and the ideal symmetrical binary channel) are compared with the recent results of HF-data transmission tests over a 5000-mile link. The 30-to-40-db increase in signal power required in the practical case is analyzed. Non-Gaussian noise, multipath disturbances (fading), and lack of coding are reasons usually mentioned to explain this poor efficiency of presently available data systems. The present paper points towards one additional reason: Low-order signal alphabets. Binary or quaternary sets of transmission signals can not achieve the ideal channel capacity as long as decision systems have to operate at 8-to-10-db "energy contrast" (E/N_0) for a sufficiently low error rate.

To more nearly approach the ideal capacity, it appears necessary to follow most of ten important principles when designing an "integrated" data system. Operating with higher-order signaling systems and applying error correcting codes are the two principles which promise the largest increase in transmission efficiency. All the rest of these ten "Cardinal Principles" are discussed briefly.

Pulse Phase-Change Signaling in the Presence of Ionospheric Multipath Distortion—S. G. Lutz, F. A. Losee, and A. W. Ladd (p. 102)

Ionospherically propagated high-frequency signals generally arrive by multiple paths having different and gradually changing time delays which cause multipath distortion. Interference fading and delay-difference smearing are the principle manifestations of this distortion.

The technique of alternately transmitting long (1 ms) and short (20 μs) pulses was developed to study multipath smearing and to determine whether the rates of change of phase would permit the use of phase-change signaling systems. Tests were conducted which confirmed the frequent occurrence of abrupt phase changes during reception of any one pulse but established that the rate of phase-change between corresponding portions of successive pulses always was gradual.

The phase-change signaling system that was developed as a result of this study is described. Data transmission (with milli-second pulses) over distances greater than 3000 km with negligible errors for prolonged periods was observed.

Synchronization of Single-Sideband Carrier Systems for High-Speed Data Transmission—T. Gombellick (p. 110)

This paper sets forth the requirements for synchronization of single-sideband carrier systems and the methods of accomplishing absolute synchronization. The problems of application to existing systems and alternative solutions are discussed.

Global Public Telephone Service—1958—D. D. Donald and T. A. Chandler (p. 115)

At the Symposium of Global Communications in Washington, D. C. in June, 1954, the author presented a paper discussing the global public telephone service as provided by the American Telephone and Telegraph Company. It is proposed here to review some aspects of public telephone service in the light of the important developments that have taken place since then.

Propagation of HF and VHF in the Arctic Region—R. Penndorf and S. C. Coroniti (p. 121)

All the available ionospheric records for stations located north of 60° geographic latitude and for the period 1954 to 1957 were critically analyzed. The F2 region shows two distinct types: one, north of 75° with very little or no diurnal and seasonal variation in the critical frequency; the second, south of 70°N, exhibits diurnal and seasonal variations as well as those correlated with sunspot activity. Communication by means of abnormal ionization, such as Es, is known. We found three types of sporadic E abnormalities; namely, the Thule type, the Auroral Belt type, and the Mixed type. These types show a distinct geographic distribution which is of utmost importance for planning communication links within the Arctic as well as between continental U. S. and the Arctic. Auroras can be used as reflectors or scatterers.

High-Speed Terminal Printers Using the Burroughs Electrostatic Technique—K. M. Kiel (p. 125)

The need for high-speed, reliable, low-cost terminal printers in communications and data processing has been forcefully brought home with the advent of electronic computers, telemetry and electronic data processing systems. Our ability to print-out has not kept pace with our ability to furnish information at high speed.

Electromechanical devices are being pushed beyond their practical ultimate limit, resulting in unreliable, high cost operation. The solution lies in ELECTRONICS—nonimpact printing and recording that takes these terrific speeds in stride!

This article deals with Burroughs new electrostatic printing technique and the high-speed, low-cost equipments utilizing this technique, such as communications teleprinters, printed tape recorders, and oscillographs.

A Synchronous Communications Receiver for the Military UHF Band—R. H. Wood and W. P. Whyland (p. 129)

A synchronous receiver for double-sideband suppressed carrier signals in the 225-mc to 400-mc band is described. The receiver is suitable for voice or data link operation and provides automatic correction of Doppler shifts in the received signal. The "local carrier" oscillator is crystal stabilized at the IF frequency, and electronically tuned by a closed servo loop in which the locally generated carrier is compared with the received double-sideband signal. The parameters of the AFC loop have been designed to provide correction from plus or minus five kilocycles. The development of this receiver represents a significant advance in the state of the communications art since it extends the advantages of suppressed carrier communications to high-performance aircraft in the UHF or higher frequency bands.

The receiver development was sponsored by an Air Force (RADC) contract for development of a multiple channel UHF transceiver for synchronous communication and is designated the AN/GRC-88 (XW-1).

A New Simplified Aircraft Data Link—M. Cooper (p. 133)

The problem of traffic control in this country's airways has been discussed at length in recent years. It is generally agreed that an automatic data link type of system will be needed to prevent saturation of radio channels and of the airways themselves as the number of aircraft in high density areas increases.

Microwave and Scatter Communications System for the Eglin Gulf Testing Range—T. J. Heckelman (p. 136)

A seven-hop 295-mile combined microwave and tropospheric forward-scatter communications system is being installed for carrying voice, timing, telemetering, radar data, and controls for the Eglin Gulf Testing Range. System design of the frequency diversity 7000 mc microwave system is reviewed, as well as the design of the four fold diversity 2000 mc scatter system.

Problems of special interest on transmission of timing signals, telemetry, and radar digital data are discussed and solutions indicated. Considerations for reliability, flexibility and ease of maintenance are indicated. Integration of the range operations by use of the communications system is described.

Contributors (p. 142)

Industrial Electronics

PGIE-10, July, 1959

Message from the Publications Chairman (p. 1)

Introduction to Session on Frontiers in Industrial Electronics—C. G. Suits (p. 2)

Automation Trends in the Banking Industry—F. B. Miller (p. 3)

Data Gathering in Continuous Chemical Processing—P. D. Schnelle (p. 8)

Application of Data-Logging and Programming Techniques to Steel Mill Processes—R. W. Barnitz and G. E. Terwilliger (p. 24)

Application of Numerical Control for Automatic Manufacturing in General Industry—W. E. Brainard (p. 33)

The First Yugoslav Machine Tool with Digital Control—G. L. Lukatela (p. 40)

Design Considerations for an Automatic Rearview Mirror—G. Platzer and L. Gau (p. 43)

Roster of PGIE Members Listed by IRE Regions and Sections, As of June 19, 1959 (p. 48)

Instrumentation

Vol. I-8, No. 2, September, 1959

Abstracts (p. 38)

Broad-Band Radio Frequency Interferometer—J. W. Carr (p. 39)

In a steady state monochromatic interferometer the interference pattern is scanned by the detector and the absolute wavelength thus determined. In the system described here the interference pattern sweeps past stationary detectors in a manner which is related to the change in frequency. Thus when the total net change in the phase of the interference pattern is measured with respect to a given absolute reference, an absolute measure of frequency can be obtained. The band being scanned is divided into an arbitrarily large number of unit cells determined by the required resolution. Pulse forming networks and logic circuits deliver the information in a manner readily adaptable to a binary counting system and substantially independent of the rate of change of RF frequency and signal amplitudes.

Noise Measurements of Negative Resistance Amplifiers—A. Brodzinsky and A. C. Macpherson (p. 44)

The usual methods for measuring the gain and noise figure of amplifiers using standard noise sources are not suitable when these devices have output impedances with a negative real part or when they are driven by negative resistance sources. Using the extended definition of noise figure of Haus and Adler, a modified procedure is described which employs an auxiliary passive dissipative network following the negative resistance device to transform the real part of the output impedance to a positive value. The necessary measurement and computational steps are outlined and the calculations are carried out for three specific networks covering the low frequency and microwave regions. It is also shown that, under certain conditions, the use of a passive dissipative network at very low temperature results in an over-all noise figure (for a cascade containing negative resistance devices) which approaches that of the first stage.

Electrometer Measurements of Very Low Densities in Gases—R. L. Ramey and R. L. Overstreet (p. 46)

Use of a vacuum tube electrometer to measure electric current as low as 6×10^{-18} amperes arising from the relative motion existing between a probe and an ion bearing gas is described. The measurement techniques are applicable to most gases at ambient pressures ranging from many atmospheres down below 10^{-7} millimeters Hg, provided that high vacuum practices are observed. Ion densities as low as 10 ions per cubic millimeter at standard conditions and flow rates of 5 to 10 cubic millimeters per second at standard conditions are used. The problems of balancing the electrometer to compensate for background, available measurement time, and probe selection are considered.

Two Economical Circuits for High-Speed Checking of Contact Closures—K. Enslein (p. 51)

This paper describes electronic circuits to be used in the checking of relay arrays to ascertain whether or not the contacts have operated properly. The circuits described are faster than relay checking systems, yet they are not as fast as the microsecond units used in computers. The check is made by examining the contacts one at a time, and counting the contact closures.

Optimization of the Aperture Corrective Systems—J. Otterman (p. 55)

The corrective networks that come to offset the aperture distortion have been described in a previous paper. In this paper the systematic distortions in an aperture system without a corrective network, and with a corrective network, are compared. Optimization of the aperture size is discussed under different conditions of noise and signal and under different optimization criteria, *i.e.*, criteria for minimization of both the random and the systematic errors. It is shown that the optimization generally will result in aperture length ranging from 0.75 π/ω_s, to 1.35 π/ω_s, where ω_s is the limiting frequency of the signal.

Production Techniques

PGPT-6, September, 1959

Message from the Editor (p. i)

Scanning the Issue (p. ii)

Technique for Automatic Testing Electronic Components—V. W. Walter and S. W. Nelson (p. 1)

The technique for automatic testing of electronic components must be adapted to the requirements for each individual test program. This paper describes the methods employed in instrumenting two test programs. The first involves approximately 75,000 components of four types—resistors, diodes, transistors and capacitors—all being tested for an extended period under environmental and electrical load conditions. The second program, in contrast, describes the parameter measurements which have to be made on small quantities of semiconductor devices under nuclear radiation conditions.

The Semiautomatic Circuit Component Tester—F. C. Brammer (p. 12)

Minimizing Production Costs Through Modular Automatic Test Equipment—J. Tampico and H. B. Rose (p. 19)

The dominant consideration influencing the trend toward automatic test and checkout techniques for production-line test stations is cost reduction. Testing costs include 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 sub-assemblies, 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 tests 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.

Some Aspects of the Thermal Design of Electronic Equipment Operating at 300–500°C Environmental Temperature—J. P. Welsh (p. 25)

Aspects of electronic part-cooling techniques applicable to 300–500°C equipments are presented. Contrary to some beliefs, the minimization of thermal resistance at these temperatures is as important as at lower temperatures. The significant shifts in the natural modes of heat transfer which occur with high-temperature electronic parts are outlined, together with some recommended methods of cooling high-temperature parts and assemblies.

Reliable Design and Development Techniques—J. E. McGregor (p. 30)

Mechanical concepts must be applied to modify electronic circuit design in order that large scale computer equipment can attain practical levels of reliability. This paper cites some specific mechanical design problems with regard to the reliability of a particular computer system, the AN/FSQ-7.

Management Views Automation—I. Travis (p. 35)

Abstracts and References

Compiled by the Radio Research Organization of the Department of Scientific and Industrial Research, London, England, and Published by Arrangement with that Department and the *Electronic and Radio Engineer*, incorporating *Wireless Engineer*, London, England

NOTE: The Institute of Radio Engineers does not have available copies of the publications mentioned in these pages, nor does it have reprints of the articles abstracted. Correspondence regarding these articles and requests for their procurement should be addressed to the individual publications, not to the IRE.

The number in heavy type at the upper left of each Abstract is its Universal Decimal Classification number. 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.

The Index to the Abstracts and References published in the PROC. IRE from February, 1957 through January, 1958 is published by the PROC. IRE, May, 1958, 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.

ACOUSTICS AND AUDIO FREQUENCIES

534-8 **3177**
Recent Researches in the Division of Acoustics, National Physical Laboratory of India—S. Parthasarathy. (*J. Sci. Industr. Res.*, vol. 17A, pp. 405-411; October, 1958.) A summary of recent work on problems at ultrasonic frequencies.

534.15 **3178**
Reciprocity Method for Vibration Measurements—H. G. Diestel and W. Mühe. (*Arch. tech. Messen.*, no. 272, pp. 183-186; September, 1958.)

534.32:621.396.625.3 **3179**
The Measurement of Pitch Fluctuations—E. Belger. (*Rundfunktech. Mitt.*, vol. 2, pp. 168-169; August, 1958.) A compromise proposal is made for the curve giving the factor of subjective assessment of pitch fluctuations as a function of modulating frequency.

534.64 **3180**
Simple Method of Measuring Acoustic Impedances of Closed Air Spaces—L. Keibs and W. Tismer. (*Tech. Mitt. BRF, Berlin*, vol. 2, pp. 31-36; May, 1958.) The ratio of sound pressure inside the enclosure to the sound pressure in front of the absorbing wall of the enclosure is determined; the method is quicker than an interferometric one.

534.76:621.396.5 **3181**
A Compatible Stereophonic Sound System—Becker. (See 3474.)

534.84:621.396.712.3 **3182**
The Architectural Design of the New Südwestfunk Studio Building at Kaiserslautern—D. Eber and A. Straub. (*Rundfunktech. Mitt.*, vol. 2, pp. 149-153; August, 1958.) Special attention has been given to the sound insulation of the building, particularly against noise from aircraft.

534.84:621.396.712.3 **3183**
The Acoustic Design of the New Südwestfunk Studio Building at Kaiserslautern—R. Thiele. (*Rundfunktech. Mitt.*, vol. 2, pp. 154-157; August, 1958.)

534.843 **3184**
The Build-Up of Sound Oscillations in Enclosed Spaces—D. Barbaro. (*Alta Frequenza*, vol. 27, pp. 472-485; October, 1958.) The theory of the Larsen effect is developed, including also the stepwise building-up of oscillations. Suggestions are made for avoiding the oscillations without a reduction in sound level.

534.86:621.3.018.78 **3185**
Investigations of the Audibility of Nonlinear Distortion—R. Reichl. (*Tech. Mitt. BRF, Berlin*, vol. 2, pp. 60-63; July, 1958.) Tape recordings with various degrees of distortion were used in a series of subjective tests.

534.86.089.6(083.74) **3186**
Researches on the Errors of Electroacoustic Standards—T. Hayasaka and M. Suzuki. (*Rep. elect. Commun. Lab., Japan*, vol. 6, pp. 309-332; September, 1958.) Equipment for calibration by the reciprocity method is described and system errors are evaluated.

534.861 **3187**
The Measurement of Program Level of High-Quality Sound Transmissions—E. Belger. (*Rundfunktech. Mitt.*, vol. 2, pp. 170-172; August, 1958.) A comparison of VU-meter and peak program meter and a proposal for the standardization of both types so that measurements made by the two methods may be compared.

534.861:621.396.712.3 **3188**
The Audio-Frequency Equipment in the New Südwestfunk Studio Building at Kaiserslautern—A. Weingärtner. (*Rundfunktech. Mitt.*, vol. 2, pp. 160-167; August, 1958.) The design of a control desk and ancillary equipment is described.

621.395.613.32 **3189**
On the Deterioration Phenomena of Microphone Carbon Powder—H. Hirabayashi. (*Rep. elect. Commun. Lab., Japan*, vol. 6, pp. 447-463; December, 1958.) A study of the physical and chemical properties of carbon powder by the measurement of contact resistance under various conditions. See also 1758 of June (Hirabayashi *et al*).

621.395.625.3:534.86 **3190**
Modern Production Methods for Multiple-Track Magnetic Sound Recording—L. Heck. (*Rundfunktech. Mitt.*, vol. 2, pp. 145-148; August, 1958.) A comparison of tape-recording techniques for achieving various sound effects in the recording of music.

ANTENNAS AND TRANSMISSION LINES

621.315.212.011.3 **3191**
Inductance of an Eccentric Tubular Conducting System—E. E. Jones. (*Brit. J. Appl. Phys.*, vol. 10, pp. 230-232; May, 1959.) A study of the effect on inductance of noncollinearity of the axes of two magnetic cylindrical tubes.

621.372.22 **3192**
Properties of Lossy Inhomogeneous Lines under Matched Conditions—H. Meinke. (*Nachrichtentech. Z.*, vol. 11, pp. 333-339; July, 1958.) Practical applications of inhomogeneous lines and methods of producing losses in them are summarized. A simple solution of the line equations is derived.

621.372.221 **3193**
The Significance of Phase and Group Delay—F. Kirschstein and H. Krieger. (*Nachrichtentech. Z.*, vol. 11, pp. 370-371; July, 1958.) Two comments on 698 of March and authors' replies.

621.372.823 **3194**
Attenuation in Circular Waveguides with Absorbing Walls—L. Caprioli. (*Alta Frequenza*, vol. 27, pp. 510-527; October, 1958.) A general expression giving attenuation also for frequencies close or equal to the critical frequencies is discussed in comparison with other approximate expressions. The attenuation characteristic is investigated for frequencies below the critical one.

621.396.67 **3195**
Pattern of an Antenna on a Curved Lossy Surface—J. R. Wait and A. M. Conda. (IRE TRANS. ON ANTENNAS AND PROPAGATION,, vol. AP-6, pp. 348-359; October, 1958. Abstract, PROC. IRE, vol. 46, p. 1976; December, 1958.)

621.396.67:621.372.837.3 **3196**
Resonant-Ring Diplexing in Forward-Scatter Systems—(*Electronics*, vol. 32, pp. 54-56; July 3, 1959.) For simultaneous transmission and reception on a single antenna, a diplexer constructed of symmetric directional couplers and a resonant ring provides isolation of about 60 db between transmitter and receiver.

621.396.67:621.396.61 **3197**
Methods of Automatic Tuning of Transmitter Aerials—O. Metze. (*Nachr. Tech.*, vol. 8, pp. 324-331, 361-364; July and August, 1958.) Equipment is described in which servo-

mechanisms are used for antenna tuning and subsequent impedance matching for an 800-watt, medium-wave ship's transmitter. The modification of a magnetic variometer for power tuning applications is discussed, particularly with regard to the choice of a suitable core material. A 100-watt automatic antenna tuning circuit using a magnetic variometer is described.

621.396.677.31:621.396.11.029.64 3198

A Rapid Beam-Swinging Experiment in Transhorizon Propagation—A. T. Waterman, Jr. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 338–340; October, 1958. Abstract, PROC. IRE, vol. 46, pp. 1975–1976; December, 1958.)

621.396.677.44 3199

The Radiation Field of the Delta Aerial, a Long-Wire Aerial with Vertical Main Beam—W. Knappe. (*Frequenz*, vol. 12, pp. 261–267; August, 1958.) The field distribution is calculated for a receiving antenna using a number of simplifying assumptions, such as a plane and perfectly conducting earth. This type of antenna has characteristics similar to those of a vertical rhombic antenna but erection costs are lower.

621.396.677.5:621.397.62 3200

Frame Aerials for Distant Television Reception—S. Sotnikov. (*Radio. Mosk.*, no. 4, pp. 31–32; April, 1959.) A brief description of an antenna using two loops disposed one above the other. This type of antenna ensures a clearer image and more stable reception over a range of 150–200 km than that obtained with conventional types. Antenna dimensions for different channels are tabulated and the general arrangement is illustrated.

621.396.677.7:621.372.826 3201

Scanning Surface-Wave Antennas: Oblique Surface Waves over a Corrugated Conductor—R. W. Hougardy and R. C. Hansen. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 370–376; October, 1958. Abstract, PROC. IRE, vol. 46, p. 1976; December, 1958).

621.396.677.71 3202

Nonresonant Slotted Arrays—A. Dion. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 360–365; October, 1958. Abstract, PROC. IRE, vol. 46, p. 1976; December, 1958.)

621.396.677.71 3203

Radiation and Reception Properties of a Wide Slot in a Parallel-Plate Transmission Line: Parts 1 and 2—R. F. Millar. (*Can. J. Phys.*, vol. 37, pp. 144–169; February, 1959.) Solutions are obtained for radiation from a slot in a parallel-plate transmission line for both E-polarized and TEM modes, and expressions are derived for the reflection and transmission coefficients and the polar diagram of the radiated field. The amplitude and phase of the propagated waves are determined from reciprocity arguments, and the dependence of the field on the slot width and distance between the plates is examined.

621.396.677.75.001.57 3204

Wax Models Speed Antenna Design—W. E. Junker. (*Electronics*, vol. 32, pp. 58, 60; June 26, 1959.) The practical details of producing a model of a wide-band dielectric antenna in paraffin wax, which has a dielectric constant similar to teflon, are described.

621.396.677.832 3205

Gains of Finite-Size Corner-Reflector Antennas—H. V. Cottony and A. C. Wilson. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 366–369; October, 1958. Abstract, PROC. IRE, vol. 46, p. 1976; December, 1958.)

AUTOMATIC COMPUTERS

621.142 3206

Study of a Digital-Analogue Converter—J. C. Bauwens. (*Rev. HF, Brussels*, vol. 4, no. 2, pp. 39–42; 1958.) A technique is described, using bistable multivibrator circuits, for generating a current proportional to the numbers on a binary scale.

681.142 3207

An Analogue Computer for the Prediction of Acoustic Propagation in the Atmosphere—F. A. Key and W. G. P. Lamb. (*Electronic Engrg.*, vol. 31, pp. 398–402; July, 1959.) Ray patterns are presented on a cathode-ray tube when air-temperature and wind-velocity data are fed into the computer.

681.142:061.3 3208

New Horizons in Computing—(*Wireless World*, vol. 65, pp. 311–314; July/August, 1959.) A report of some of the computer engineering techniques which were described in papers read at the International Conference on Information Processing held in Paris, June 15–20, 1959.

681.142:537.312.62 3209

Superconducting Computer Elements—E. H. Rhoderick. (*Brit. J. Appl. Phys.*, vol. 10, pp. 193–198; May, 1959.) Switching elements working at 30 mμs and storage elements of the persistent-current type with a transition time of 15 mμs have been developed under laboratory conditions. Their successful commercial possibilities depend mainly on the development of thin metallic films of consistent quality used in the construction of these elements.

681.142:621.318.4 3210

Flux-Controlling-Type Adder—N. Kuroyanagi. (*Rep. elect. Commun. Lab., Japan*, vol. 6, pp. 468–475; December, 1958.) Experimental circuits using magnetic cores to control switching are described.

681.142:621.318.57:537.311.33 3211

The Cryosar—a New Low-Temperature Computer Component—A. L. McWhorter and R. H. Rediker. (PROC. IRE, vol. 47, pp. 1207–1213; July, 1959.) The cryosar is a high-speed component whose operation, at about 4°K, is based on impact ionization of impurities in Ge. Two types, using compensated and uncompensated Ge, can be used as bistable and normal diodes respectively. Turn-on time is about 10^{-8} seconds and they are very small; present results suggest excellent reliability.

681.142:621.318.57:538.221 3212

Thin Magnetic Films for Digital Computer Memories—D. O. Smith. (*Electronics*, vol. 32, pp. 44–45; June 26, 1959.) Characteristics of magnetic films and ferrite cores are tabulated.

CIRCUITS AND CIRCUIT ELEMENTS

621.3.049.7:537.311.33 3213

Semiconductor Devices for Microminiaturization—J. T. Wallmark and S. M. Marcus. (*Electronics*, vol. 32, pp. 35–37; June 26, 1959.) Integrated semiconductor devices are illustrated, in which active and passive components are incorporated in a single piece of semiconductor material with no interconnecting metallic leads; their application to logic circuits is explained.

621.314.22:621.396.822 3214

Design of Random-Noise Transformers—M. P. Vore. (*Commun. and Electronics*, no. 41, pp. 59–63; March, 1959.) The design is developed from basic principles and verified by experiment.

621.316.5 3215

Representation of Switching Circuits by Binary-Decision Programs—C. Y. Lee. (*Bell Sys. Tech. J.*, vol. 38, pp. 985–999; July, 1959.)

621.316.8 3216

The Total-Excursion Resistor—R. H. W. Burkett. (*Electronic Engrg.*, vol. 31, pp. 393–397; July, 1959.) The design of resistors that will remain within a guaranteed tolerance range throughout their life is discussed.

621.318.4.042:538.221 3217

Behaviour of Square-Loop Magnetic Cores in Circuits—D. A. H. Brown. (*Electronic Engrg.*, vol 31, pp. 408–411; July, 1959.) The switching time with a loaded secondary winding is derived. The behaviour of ferrite cores under various conditions of loading and drive is shown to agree with theory.

621.318.57 3218

High-Speed Multiplexing with Closed-Ring Counters—K. L. Berns and B. E. Bishop. (*Electronics*, vol. 32, pp. 48–50; June 26, 1959.) The action of a 20-channel electronic switch, which consists of twenty semiconductor diode gates each controlled by a flip-flop circuit, is explained. The switch can sample from 2 to 20 voltage sources at rates as high as 50 kc.

621. 319.42 3219

Probable Life of Paper Capacitors—J. P. Mayeur. (*Câbles & Trans.*, vol. 12, pp. 136–143; April, 1958.) A study based on statistical laws of probability. Experimental results are in good agreement with theory.

621.372:517.512.2 3220

Study of the Transfer Properties of a Linear System based on the Concept of Rate of Growth of Spectral Energy—C. Lafleur. (*Rev. HF, Brussels*, vol. 4, no. 3, pp. 59–66; 1958.)

621.372.01 3221

Elements of Electronic Circuits: Part 4—Use of Short-Time-Constant Circuits with Diodes and Triodes—J. M. Peters. (*Wireless World*, vol. 65, pp. 346–347; July/August, 1959.) Part 3: 2494 of August.

621.372.54 3222

Network Transformations for Wave Filter Design—A. Zverev and H. Blinchikoff. (*Electronics*, vol. 32, pp. 52–54; June 26, 1959.) Calculations are simplified by charts which enable the configuration of two-terminal networks to be altered while retaining their impedance characteristic, and also to alter the element values without changing the configuration or impedance characteristics.

621.372.543.2 3223

Band-Pass Ladder-Filter Half-Sections—J. Bimont. (*Câbles & Trans.*, vol. 12, pp. 96–120; April, 1958.) Considering purely reactive half-sections, general formulas for branch and image impedances and reflectionless interconnections are examined.

621.372.543.2 3224

Generalization of Zobel-Type Ladder Filters—J. E. Colin. (*Câbles & Trans.*, vol. 12, pp. 185–205; July, 1958.) A new and simplified method of design in which the filter is considered as an indivisible unit rather than a combination of elementary structures as in the Zobel method.

621.372.543.2:[538.652+537.228.1 3225

Physical Theory and Technique of Electromechanical Wave Filters—W. Poschenrieder. (*Frequenz*, vol. 12, pp. 246–255; August, 1958.) Equivalent circuits are derived and the design of some practical filters is described with details of their performance.

621.372.552 3226

Constant-S Equalizers—B. D. Solomon and C. S. Broneer. (*J. Audio Engrg. Soc.*, vol. 6, pp.

210–215; October, 1958.) Equations are derived for high-, medium-, and low-frequency equalizers based on a bridged-T configuration. As equalization is varied from step to step, the point of maximum equalization changes and the insertion loss S of unequalized frequencies remains constant.

621.373.51:621.314.63 3227
Tunnel Diodes as High-Frequency Devices—Sommers. (See 3505.)

621.374:621.314.7 3228
Transistor Applications in Pulse Techniques—E. Baldinger. (*Bull. Schweiz. elektrotech. Ver.*, vol. 50, pp. 2–9; January 3, 1959.) Details are included of high-speed divider circuits and a stabilized dc power unit.

621.374.3:621.314.7:681.142 3229
Transistor Circuits for a 1 Mc/s Digital Computer—I. Krajewski. (*Electronic Engrg.*, vol. 31, pp. 403–407; July, 1959.) "The requirements of amplifiers used for pulse regeneration are considered and blocking oscillators and regenerative amplifiers employing transistors suitable for reshaping and retiming pulses in a digital computer with a 1 Mc/s digit rate are described. Some details of their practical application are given."

621.374.32:621.314.7 3230
Simplified Coincidence Circuits using Transistors and Diodes—R. H. Miller. (*Rev. Sci. Instr.*, vol. 30, pp. 395–398; June, 1959.) A simple coincidence circuit is described which is as fast as a standard valve circuit. The measured threshold and resolving-time characteristics are given.

621.374.32:621.314.7 3231
Transistor Circuitry for Radiation Counting—J.R. Gilland. (*Rev. Sci. Instr.*, vol. 30, pp. 479–484; June, 1959.) Four instruments are described: a pulse amplifier, a pulse-height analyzer, a scaler, and a linear count-rate meter.

621.374.44:621.3.018.41(083.74) 3232
Using Divider Vernier to Synchronize Pulses—E. F. Wilson. (*Electronics*, vol. 32, pp. 44–45; July 3, 1959.) The vernier adds or deletes pulses at the input of frequency dividers, thus enabling the output to be precisely synchronized, *e.g.*, to the WWV 1-second pulses.

621.375.024:537.311.33 3233
Present Applications of Semiconductors—J. Florine. (*Rev. HF, Brussels*, vol. 4, no. 4, pp. 89–102; 1958.) Description of semiconductor applications with particular reference to dc amplification for analog computers.

621.375.121:621.396.61 3234
Wideband Amplification at High Frequencies—V. O. Stokes and B. M. Sosin. (*Point to Point Telecommun.*, vol. 3, pp. 11–23; June, 1959.) A 1-kilowatt transmitter incorporating a wideband amplifier covering 2–24 mc is described. Distributed amplification is used with artificial transmission lines in which the shunt capacitances are provided by the interelectrode capacitances of the tubes.

621.375.132.3 3235
The Design of Cathode-Follower Circuits—R. Gärtner. (*Nachr. Tech.*, vol. 8, pp. 313–323; July, 1958.) A detailed summary of design principles and operating characteristics with a description of several circuit applications.

621.375.2.029.63 3236
The Fundamentals of a Low-Noise H.F. Preamplifier for 70 to 40 cmλ—H. Rieck. (*Nachr. Tech.*, vol. 8, pp. 306–313; July, 1958.) The circuit described uses a disk-seal triode in grounded-grid connection and is intended for radio-astronomy applications. At a wavelength of 54 cm and with IF bandwidth 4 mc the input noise figure is 4.2 kT_0.

621.375.3 3237
Contribution to the Study of the Operation of Magnetic Amplifiers—E. Bernard and E. Pio. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 394–397; January 19, 1959.) The operation of a basic magnetic amplifier circuit is analyzed assuming a square-wave input and a rectangular hysteresis loop, the area of which is increased due to eddy currents. An equivalent circuit is derived that is applicable for the case of a periodic or continuous control signal.

621.375.3 3238
Volt-Second Transfer Efficiency in Fast-Response Magnetic Amplifiers: Part 2—N^2/R as a Design Parameter—J. T. Pula, G. E. Lynn and J. F. Ringelman. (*Commun. and Electronics*, no. 41, pp. 8–11; March, 1959.) The relation between volt-seconds and ampere-turns is confirmed by measurements on a specially-designed magnetic amplifier. Part 1: 2171 of July (Pula).

621.375.43 3239
Some Aspects of Feedback in Transistor Amplifiers—R. Dessoulavy. (*Bull. schweiz. elektrotech. Ver.*, vol. 50, pp. 233–239; March 14, 1959. In French.) The design of an amplifier with high input impedance is described and problems peculiar to the use of feedback in transistor circuits are discussed.

621.375.9 3240
Parametric Energy Conversion in Distributed Systems—G. M. Roe and M. R. Boyd. (PROC. IRE, vol. 47, pp. 1213–1218; July, 1959.) Analysis of most types of the travelling-wave parametric amplifier with little or no dispersion shows that no exponential amplification at the fundamental frequency of an applied signal is possible. Although the total energy increases with distance it is spread among a large number of frequency cross-products. For the "beam" type, system amplification of the fundamental component may become possible as dispersion increases, but a more detailed analysis is required.

621.375.9:537.311.33 3241
The Grain-Boundary Amplifier—O. A. Weinrich, H. Mataré and B. Reed. (*Proc. Phys. Soc., London*, vol. 73, pp. 969–972; June 1, 1959.) The construction and characteristics are described of an amplifier which is based on modulation of the sheet conductance of a grain boundary in *n*-type Ge. At 78°K, a power amplification of 21 db was obtained at a frequency of 1 kc, which approached zero at 50 kc.

621.375.9:538.569.4.029.64 3242
A New Class of Materials for Bloembergen-Type Masers—B. Bleaney. (*Proc. Phys. Soc., London*, vol. 73, pp. 937–939; June 1, 1959). A class of substances which have four levels degenerate in zero field exists. These levels show a linear Zeeman effect, but are not necessarily equally spaced when a magnetic field is applied, and there are allowed transitions which are not limited to those between adjacent energy levels.

621.375.9:621.3.011.23 3243
Superregenerative Reactance Amplifier—B. B. Bossard. (PROC. IRE, vol. 47, pp. 1269–1271; July, 1959.) The method of operation of an L-band variable-reactance amplifier as a superregenerative amplifier is clearly described. The results indicate the feasibility of a single-stage low-noise microwave receiver.

621.375.9:621.3.011.23:538.221 3244
Parametric Amplifiers and Cancellation of Ferromagnetic Hysteresis by Orthogonal Polarization—Y. Angel and G. A. Boutry. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 384–386; January 19, 1959.) Measurements have been made on tubular samples of Ni-Zr ferrites showing the effect of a uniform magnetic field H_t parallel to the axis on the magnetization due to alternating current in a toroidal winding. For a sufficiently high field intensity H_t, hysteresis becomes negligible and the magnetization curve tends to become linear. This principle has been applied in the design of a variable-inductance parametric amplifier operating at 2.2 mc with a gain of 20 db.

621.375.9:621.3.011.23:621.376.23 3245
Parametric Amplifiers as Superregenerative Detectors—J. J. Younger, A. G. Little, H. Heffner and G. Wade. (PROC. IRE, vol. 47, pp. 1271–1272; July, 1959.) The details of superregenerative operation are described and effects permitting self-quenching are explained. Two cavity-type parametric amplifiers using semiconductor diodes are investigated, one in the L-band near 780 mc and the other in the S-band near 3 kmc.

621.375.9:621.314.63 3246
Low-Noise Tunnel-Diode Amplifier—K. K. N. Chang. (PROC. IRE, vol. 47, pp. 1268–1269; July, 1959.) An amplifier consisting of a tuned circuit shunted by a negative-conductance "tunnel" diode is described, and experimental gains (20–40 db), bandwidths (0.2–0.8 mc) and noise figures (4–8 db) obtained at 30–100 mc are quoted. Greatly improved noise figures and microwave operation should be possible ultimately. See 3505 below.

621.376.32:621.319.43:537.227 3247
Ferroelectric Capacitors—T. W. Butler, Jr, and G. A. Roberts. (*QST*, vol. 43, pp. 32–36, 144; July, 1959.) A VFO is described which incorporates a ferroelectric capacitor as a frequency-modulating element.

GENERAL PHYSICS

530.112:530.12 3248
The Deflection of Waves by Movement of the Media of Propagation—A. Metz. (*Compt. Rend. Sci., Paris*, vol. 248, pp. 1615–1617; March 16, 1959.) Optical observations of aberrations of fixed stars are held to be inconsistent with the hypothesis that light waves are affected by movements of the ether. This disproves Datzeff's theory (1381 of 1958) which has already been criticized (see 3753 of 1958).

534.2+538.566 3249
Related Experiments with Sound Waves and Electromagnetic Waves—W. E. Kock. (PROC. IRE, vol. 47, pp. 1192–1201; July, 1959.) Analogous situations in the propagation of sound and electromagnetic waves are described and demonstrated experimentally; these include cutoff effects, polarization rotation, and external guiding. Superdirective arrays are also examined.

537.291 3250
Stable Orbits of Charged Particles in an Oscillating Electromagnetic Field—E. S. Weibel. (*Phys. Rev.*, vol. 114, pp. 18–21; April 1, 1959.) In the field of a cylindrical waveguide driven in the TE_{01} mode, the particle moves around the axis of the guide. Sufficient conditions for stable orbits are derived.

537.311.31 3251
Charge Penetration into a Conductor in Equilibrium—R. Fürth and E. Morris. (*Proc. Phys. Soc., London*, vol. 73, pp. 869–872; June 1, 1959.) "Statistical mechanics is used to show that the electric charge on a conductor in equilibrium is not restricted to its surface but penetrates into the interior in the form of a space charge density, the effective penetration depth

being of the order of a few atomic distances. Similarly, the electric potential in a charged conductor is not strictly constant but varies in space which results in a small but finite potential difference between the interior and the surface layers."

537.311.32 3252
The Determination of a Criterion for the Surface Conduction of Insulating Solids—R. Lacoste. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 655–657; February 2, 1959.) An electrode system is described for determining a curve of the variation of specific resistance of a specimen as a function of distance between electrodes. A criterion of surface conduction may be deduced by extrapolation.

537.523:621.396.822 3253
Electron Temperature and Electron Noise in High-Frequency Torch Discharges—L. Mollwo. (*Ann. Phys.* (*Lpz.*), vol. 2, pp. 97–129; October 12, 1958.) The noise temperature of an electron torch [see, *e.g.*, 2785 of 1951 (Cobine and Wilbur)] in air and nitrogen at atmospheric pressure is measured at 20 cmλ and found to be 14,000°K, of which about 13,500°K is due to electron temperature and 380°K to shot noise. These results and the measurements of other plasma characteristics are interpreted with reference to the findings of other authors. 85 references.

537.525 3254
Nature and Role of Ionizing Potential Space Waves in Glow-to-Arc Transitions—R. G. Westberg. (*Phys. Rev.*, vol. 114, pp. 1–17; April 1, 1959.)

537.533.7:539.23 3255
Influence of the Diffusion Angle on the Energy Spectrum of an Electron Beam after passing through a Thin Film—C. Fert and F. Pradal. (*Compt. Rend. Acad. Sci. Paris*, vol. 248, pp. 666–669; February 2, 1959.) Measurements have been made with a magnetic spectrograph on films of Al, Ge and Mn.

537.533.73 3256
The Three-Dimensional Intensity Distribution near the Focus of Waves Diffracted by Slit and Rectangular Apertures—B. J. Thompson. (*Proc. Phys. Soc., London*, vol. 73, pp. 905–911; June 1, 1959.)

537.533.8 3257
Theory of Secondary Electron Emission of Metals—H. Stolz and H. W. Streitwolf. (*Z. Naturforsch*, vol. 13a, pp. 1100–1101; December, 1958.) Preliminary note.

537.533.8 3258
Electron Emission from the Surface of Pure Molybdenum after Bombardment by Electrons—M. S. Sinel'nikov. (*Dokl. Ak. Nauk, SSSR*, vol. 126, pp. 554–556; May 21, 1959.) After low-energy electron bombardment for a period of 6 seconds, a pure Mo surface at room temperature in the dark was found to emit electrons at energies up to 7 keV. The electron current decreased from 10^{-11} amps to 10^{-14} amps after 10 hours.

537.533.9 3259
The Excitation of Molecular Vibration and Rotation by Impact of Slow Electrons—H. S. W. Massey. (*Phil. Mag.*, vol. 4, pp. 336–340; March, 1959.) "A new approximate formula is derived for the cross-sections for excitation of molecular vibration and rotation by impact of slow electrons. Account is taken of the fact that in such impacts the incident electron is moving much faster than the nuclei in the molecule, while its wave function suffers considerable distortion by the molecular field."

537.56 3260
Electron Energy Distributions in Plasmas: Part 1—R. L. F. Boyd and N. D. Twiddy. (*Proc. roy. Soc. London A*, vol. 250, pp. 53–69; February 24, 1959.) Report of a method of carrying out a Druyvesteyn analysis electronically. The results show very wide deviations from the Maxwellian distribution in striated discharges and suggest that a high anomalous rate of energy exchange between the electrons of a low-pressure discharge is a much rarer phenomenon than might be supposed.

537.56 3261
Langevin Equation and the A.C. Conductivity of Non-Maxwellian Plasmas—P. Molmud. (*Phys. Rev.*, vol. 114, pp. 29–32; April 1, 1959.)

537.56:538.56 3262
Microwave Investigations of Plasma—M. D. Raĭzer and I. S. Shpigel'. (*Uspekhi fiz. Nauk*, vol. 64, pp. 641–667; April, 1958.) The basic theory and the experimental details of three methods are described: resonator method, waveguide method and cross-modulation method. The ambipolar diffusion and recombination coefficients and the electron collision frequency are determined and experimental results are reported. 55 references.

538.221 3263
Excluded-Volume Problem and the Ising Model of Ferromagnetism—M. E. Fisher and M. F. Sykes. (*Phys. Rev.*, vol. 114, pp. 45–58; April 1, 1959.)

538.561:537.533:538.6 3264
Stimulated Emission of Radiation by Relativistic Electrons in a Magnetic Field—J. Schneider. (*Phys. Rev. Lett.*, vol. 2, pp. 504–505; June 15, 1959.) An expression is derived for the transfer of energy between an alternating electric field and electrons moving at right angles to a magnetic field. It is shown that for relativistic electrons a net stimulated emission is possible.

538.566:535.42]+534.26 3265
Diffraction by a Smooth Object—B. R. Levy and J. B. Keller. (*Commun. Pure Appl. Math.*, vol. 12, pp. 159–209; February, 1959.) A geometrical theory of diffraction, a theory of wave propagation based on an extension of geometrical optics, is explained and applied to diffraction problems associated with smooth convex opaque objects.

538.566:535.42 3266
Diffraction of a Skew Plane Electromagnetic Wave by an Absorbing Right-Angled Wedge—F. C. Karal, Jr., and S. N. Karp. (*Commun. Pure Appl. Math.*, vol. 11, pp. 495–533; November, 1958.)

538.566:537.56 3267
Growth of Electric Space-Charge and Radio Waves in Moving Ion Streams—J. H. Piddington. (*Phil. Mag.*, vol. 3, pp. 1241–1255; November, 1958.) In a medium comprising two identical interpenetrating ion streams, four different space-charge phenomena may occur. Two of these are discussed: 1) an instability, described as a "growing ion cloud"; this is thought to be the basic mechanism of a two-stream amplifier and to be important in solar rf emission; and 2) an evanescent wave, sometimes wrongly interpreted as a spatially growing wave. See 2702 of 1956 and back references.

538.566.2:539.32 3268
Thin Absorbent Films: the Problem of Optimum Thickness in connection with a Classification of these Films—M. Gourceaux. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 392–394; January 19, 1959.) Theoretical note on determining the film thickness for which the absorption of an incident plane electromagnetic wave is a maximum.

538.569.4:538.222:535.34 3269
An Electronic Paramagnetic-Resonance Spectrometer: Influence of the Amplitude of Modulation of the Magnetic Field on the Line Shape obtained by Synchronous Detection—J. Roch. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 663–666; February 2, 1959.)

538.569.4:539.2 3270
Cross-Relaxation in Spin Systems—N. Bloembergen, S. Shapiro, P. S. Pershan and J. O. Artman. (*Phys. Rev.*, vol. 114, pp. 445–459; April 15, 1959.) Analysis of the transition region of nearly equally spaced levels.

538.63 3271
Galvanomagnetic Properties of Cylindrical Fermi Surfaces—J. M. Ziman. (*Phil. Mag.*, vol. 3, pp. 1117–1127; October, 1958.)

GEOPHYSICAL AND EXTRATERRESTRIAL PHENOMENA

523.152.2 3272
On Highly Ionized Regions of Interstellar Matter—A. M. Rozis-Saulgeot. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 650–653; February 2, 1959.) A parallel is established between the properties of heavily-ionized regions and the superconducting properties of metals at very low temperatures. This analogy leads to a new description of certain regions of the interstellar medium.

523.164 3273
Principles and Present Results of Radio Astronomy—R. Coutrez. (*Rev. HF, Brussels*, vol., 4, pp. 71–84; 1958.)

523.164 3274
The Nature of the Cosmic Radio Sources—M. Ryle. (*Proc. roy. Soc. London A*, vol. 248, pp. 289–308; November 25, 1958.) A detailed review of observational data and discussion of theory. 50 references.

523.164.3 3275
Radio Emission of the Sun and Planets—V. V. Zheleznyakov. (*Uspekhi Fiz. Nauk.*, vol. 64, pp. 113–154; January, 1958.) A description of the sun's radio emissions during periods of low and high activity on 21 and 60 cmλ. Radio bursts are considered and the thermal and sporadic theories are discussed. Radio emissions from Jupiter, Venus and Mars are also examined. 158 references.

523.164.32:523.75 3276
Radio Emission following the Flare of August 22, 1958—A. Bioschot and J. W. Warwick. (*J. Geophys. Res.*, vol. 64, pp. 683–684; June, 1959.) Two types of emission appear to have occurred: 1) a burst at 2800 mc and 470 mc, decreasing in intensity with decreasing frequency, and 2) a long-lasting continuum, on frequencies below 200 mc, delayed with respect to 1).

523.165 3277
Measurement of the Cosmic-Ray Intensity in the Stratosphere at Various Heights and Latitudes—A. N. Charakhch'yan and T. N. Charakhch'yan. (*Zh. eksp. teor. Fiz.*, vol. 35, pp. 1088–1102; November, 1958.) Investigation of the altitude dependence of particles of soft the cosmic-ray soft component in the latitude range 51° to 31°. Results indicate that at a latitude of 31° the majority of particles of soft components are electrons generated by Π-mesons. At latitude 51° an excess of electrons is found possessing energies below 2–3 g/cm². Expressions are derived for the primary particle energy spectrum. Near the equator (latitude 2°) the primary cosmic-ray particle flux has been found equal to 0.48 ±0.004 particles/min. cm². sterad.

523.165:523.746.5 3278
On Recurrent Variations in the Intensity

of Primary Cosmic Radiation in a Period of Maximum Solar Activity—A. Fréon, J. Berry and J. P. Coste. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 674–677; February 2, 1959.) Recordings of variation in the intensity of cosmic radiations, made during the period October, 1956 to July, 1957, show a cyclic variation with a stable period of 27.35 ±0.1 mean solar days, the maximum amplitude being reached in October, 1957.

523.165:523.746.5 3279
Flux and Energy Spectrum of Cosmic-Ray α Particles during Solar Maximum—P. S. Freier, E. P. Ney and C. J. Waddington. (*Phys. Rev.*, vol. 114, pp. 365–373; April 1, 1959.) Flux measured over Minnesota and Texas are significantly lower than those observed at solar minimum. A mechanism is proposed to explain the changes.

523.165:550.38 3280
Cosmic Ray Cut-off Rigidities and the Earth's Magnetic Field—J. J. Quenby and W. R. Webber. (*Phil. Mag.*, vol. 4, pp. 90–113; January, 1959.)

523.745:(551.510.535+523.164.32 3281
Some Indices of Solar Activity based on Ionospheric and Radio Noise Measurements—C. M. Minnis and G. H. Bazzard. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 213–228; June, 1959.) Methods are described for producing indices of solar activity based on the ionization of the E and F_2 layers and monthly mean figures thus derived are tabulated. The two indices are compared with each other and with solar noise flux at 10.7 cmλ and the correlation is high. Practical applications of the indices are discussed briefly.

523.746.5 3282
Prediction of Sunspot Numbers until the End of the Present Cycle—P. Herrinck. (*Nature, London*, vol. 184, pp. 51–52; July 4, 1959.) A modified equation derived from recent data has been used for the determination of sunspot numbers for the period 1959–1968.

523.75:523.164.32 3283
Relation between Type-IV Emissions and other Forms of Solar Activity—Y. Avignon and M. Pick. (*Comp. Rend. Acad. Sci., Paris*, vol. 248, pp. 368–371; January 19, 1959.) Observations indicate a correlation between optical centers associated with type-IV radio emissions [see, *e.g.*, 3801 of 1958 (Boischot and Denisse)] and radio storm centers.

550.385:523.75 3284
Geomagnetic Activity following Large Solar Flares—C. S. Warwick and R. T. Hansen. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 287–295; June, 1959.) There is a marked tendency for magnetic disturbance to follow the largest solar flares. No association was shown in the period 1950–1954 near sunspot minimum; this confirms the negative result found by Watson (3853 of 1957) for roughly the same period.

.385.43 3285
The Propagation of World-Wide Sudden Commencements of Magnetic Storms—V. B. Gerard. (*J. Geophys. Res.*, vol. 64, pp. 593–596; June, 1959.) Analysis of differences in the recorded time of the first impulse at ten widely separated observatories suggests that the position of the sun controls the hemisphere in which the sudden commencement first occurs.

550.389.2:629.19 3286
Circuits for Space Probes—R. R. Bennett, G. J. Gleghorn, L. A. Hoffman, M. G. McLeod and Y. Shibuya. (*Electronics*, vol. 32, pp. 55–57; June 19, 1959.) Brief description of instrumentation in the payload stage of the Pioneer vehicle and at tracking stations.

550.389.2:629.19 3287
Theory of the Effect of Drag on the Orbital Inclination of an Earth Satellite—J. P. Vinti. (*J. Res. Nat. Bur. Stand.*, vol. 62, pp. 79–88; February, 1959.) It is assumed that the drag is in the direction of the air velocity relative to the satellite, and that its magnitude diminishes so rapidly with altitude that it is appreciable only near perigee in an elliptic orbit. Results are deduced for the secular changes in inclination up to, but not including the final ballistic stage.

550.389.2:629.19 3288
Optical Observations of Artificial Earth Satellites—I. S. Shklovskiĭ and P. V. Shcheglov. (*Uspekhi. Fiz. Nauk*, vol. 64, pp. 417–424; March, 1958.) Observations are made visually using telescopes and are also recorded photographically. An oscillographic method is described for recording the exact instant of the opening and closing of the camera shutter. These observations are valuable in the determination of the exact shape of the earth.

550.389.2:629.19 3289
Radio Observations of the First Two Artificial Earth Satellites—H. Fleischer. (*Nachrichtentech. Z.*, vol. 11, pp. 340–347; July, 1958.) Review of some of the early measurements made by stations in West Germany.

550.389.2:629.19 3290
Radar Observations of the Second Russian Earth Satellite (Sputnik II: 1957β)—J. G. Davies, J. V. Evans, J. S. Greenhow, J. E. Hall, E. L. Neufield and J. H. Thomson. (*Proc. Roy. Soc. London A*, vol. 250, pp. 367–376; March 24, 1959.) An analysis of observations made with the 80-meter steerable radio telescope at Jodrell Bank at frequencies of 36 and 100 mc. Results suggest that the observed fading characteristics were due mainly to rotation of the satellite; from a study of the scattering area it is concluded that the satellite was a long object having a reradiation polar diagram with one major and many minor lobes. There was no evidence that the satellite produced any ionization.

550.389.2:629.19 3291
Recording of Sputnik II on 40 002 kc/s—H. A. Hess. (*Nachrichtentech. Z.*, vol. 11, pp. 347–348; July, 1958.) Field-strength recordings taken at Breisach, West Germany, of day-time transits during the period November 4–8, 1957, are discussed.

550.389.2:629.19:523.164 3292
Radio-Astronomical Investigations by Means of Artificial Earth Satellites—G. G. Getmantsev, V.L. Ginsburg and I.S. Shklovskiĭ. (*Uspekhi. Fiz. Nauk*, vol. 66, pp. 157–161; October, 1958.) The ionosphere cuts off extraterrestrial radio emissions over a wide range of frequencies. Thus observations which would not be possible on the ground could be made from satellites.

550.389.2:629.19:523.72 3293
Investigation of Solar Corpuscular Radiation using Artificial Earth Satellites—V. I. Krasovskiĭ. Yu. M. Kushnir and G. A. Bordovakii. (*Uspekhi. Fiz. Nauk*, vol. 64, pp. 425–434; March, 1958.)

550.389.2:629.19:551.510.535 3294
Discovery of 10-keV Electrons in the Upper Atmosphere by means of the Third Soviet Earth Satellite—V. I. Krasovskiĭ, I. S. Schklovskiĭ, Yu. I. Gal'perin and E. M. Svetlitskii. (*Dokl. Ak. Nauk SSSR*, vol. 127, pp. 78–81; July 1, 1959.) Note of electron energies recorded between 470 and 1880 km above sea level. The maximum electron energy registered was 40 keV.

550.389.2:629.19:551.510.535 3295
Investigations of the Ionosphere by means of an Artificial Earth Satellite—Ya. L. Al'pert. (*Uspekhi. Fiz. Nauk.* vol. 64, pp. 3–14; January, 1958.) A description of methods based on the Doppler effect in which correct results depend on the appropriate choice of the two frequencies used.

550.389.2:629.19:551.510.535 3296
Some Deductions of Ionospheric Information from the Observations of Emissions from Satellite 1957 α2: Part 1—The Theory of the Analysis—G. J. Aitchison and K. Weekes. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 236–243; June, 1959.) The effect of the ionosphere on the Doppler shift frequency and on the frequency of Faraday fading is expressed theoretically in terms of the difference in phase path of the waves for two positions of the satellite.

550.389.2:629.19:551.510.535 3297
Some Deductions of Ionospheric Information from the Observations of Emissions from Satellite 1957 α2: Part 2—Experimental Procedure and Results—G. J. Aitchison, J. H. Thomson and K. Weekes. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 244–248; June, 1959.) The application of the approximate analysis to the small number of suitable observations suggests that the ionization density decreases slowly with height above the maximum of F_2 at night. Part 1: 3296 above.

550.389.2:629.19:551.510.535 3298
Some Results of the Determination of the Electron Concentration of the Outer Regions of the Ionosphere by Observations of the Radio Signals from the First Earth Satellite—Ya. L. Al'pert, F. F. Dobryakova, R. F. Chudesenko, and B. S. Shapiro. (*Uspekhi. Fiz. Nauk*, vol. 65, pp. 161–174; June, 1958.) The method of calculating the path of the radio waves in the outer ionosphere is described. The electron concentrations determined for heights between 200 and 3120 km, which vary between $10^5/cm^3$ and $10^2/cm^3$, are tabulated.

550.389.2(54):551.510.535 3299
The Indian Program for the International Geophysical Year: Part 1—Some Results of the Program on Ionosphere—A. P. Mitra. (*J. Sci. Industr. Res.*, vol. 17A, pp. 395–401; October, 1958.) A report of some ionospheric effects, particularly those observed in equatorial regions.

551.510.53 3300
A Preliminary Model Atmosphere based on Rocket and Satellite Data—H. K. Kallmann. (*J. Geophys. Res.*, vol. 64, pp. 615–623; June, 1959.) A model atmosphere is described for the height range 100–800 km and values of density, pressure and scale height are tabulated as functions of altitude. Discrepancies between data from various sources are examined and it is estimated that they are within the limits to be expected from latitude variations and the uncertainties of experimental observations.

551.510.535 3301
"Chapman Behaviour" in the Lower Ionosphere—C. H. Cummack. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 229–235; June, 1959.) "It is shown that Chapman's relations for the variation of critical frequency and height of a layer with solar zenith angle can be applied to any point on that layer and that it is not necessary to assume monochromatic radiation. These results are then applied to electron density *vs* height profiles and it is shown that 'Chapman behaviour' can be detected to a height of 160 km being consistent with a scale height of 10 or 15 km."

551.510.535 3302
Some Wind Determinations in the Upper Atmosphere using Artificially Generated

Sodium Clouds—E. Manring, J. F. Bedinger, H. B. Pettit and C. B. Moore. (*J. Geophys. Res.*, vol. 64, pp. 587–591; June, 1959.) Results for the height range 77–200 km are reported.

551.510.535 3303

Coefficients for the Rapid Reduction of h'-f Records to N-h Profiles without Computing Aids—E. R. Schmerling and C. A. Ventrice. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 249–261; June, 1959.) The tables of coefficients can be used for the ordinary ray at any station having a magnetic dip less than 80°. The sensitivity of the coefficients to dip angle and gyrofrequency is discussed and comparisons are made with the profiles produced using the Budden matrix method.

551.510.535 3304

Spiral Occurrence of Sporadic-E—E. L. Hagg, D. Muldrew and E. Warren. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 345–347; June, 1959.) Intense sporadic-E ionization at high latitudes in the northern hemisphere is distributed along a curve similar to a Störmer precipitation spiral for negative particles.

551.510.535 3305

An Additional Lunar Influence on Equatorial E_s at Huancayo—R. W. Knecht. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 348–349; June, 1959.) The time of first appearance of equatorial E_s during the period May, 1957—April, 1958 is shown to vary with lunar age. This supports Matsushita's hypothesis on the relation between S_q currents and equatorial E_s (see 2124 of 1957).

551.510.535:523.78 3306

Some Further Analyses of E-Layer Measurements in South Africa during the Solar Eclipse of 25 December 1954—M. W. McElhinny. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 273–286; June, 1959.) A re-examination of data [see 2061 of 1956 (Szendri and McElhinny) for preliminary results] suggests that the measurements could be accounted for by assuming that 15 per cent of the solar radiation comes from outside the visible disk. At the same time it is necessary to postulate an effective recombination coefficient of $4\times10^{8-}$ cm^3 sec^{-1}.

551.510.535:551.594.6 3307

The Electron Density Distribution in the Outer Ionosphere derived from Whistler Data G. McK. Allcock. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 185–199; June, 1959.) A method of successive approximations is developed to determine the electron density distribution with height in the outer ionosphere when the variation of whistler dispersion with geomagnetic latitude is known. The method is applied to specific whistler data and the results infer that the electron density distribution decreases exponentially with height in the range 1000–13,000 km. An equation is derived for the distribution in this height range during January–May, 1957.

551.510.535:621.396.11 3308

Some Possible Effects caused by Strong Gyro-Waves in the Ionosphere: Part 1—V. A. Bailey. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 299–324; June, 1959.) A theory is developed for the effect of an extremely powerful pulse of extra-ordinary gyro-waves on the nocturnal E-region near 92 km and the day-time D-region near 75 km. The changes in collisional frequency and electron density at these levels are derived with their consequent effects on the propagation characteristics of a transverse, ordinary wave at 590 kc. It is shown that, with radio pulses of sufficient power, observable changes may be produced in air glow, meteor trails, fast particle tracks and local magnetic elements. It would also be possible to use such a pulse to determine the coefficient of attachment at the 92-km and 75-km levels.

551.594.5 3309

Auroral X-Rays, Cosmic Rays, and Related Phenomena during the Storm of February 10–11, 1958—J. R. Winckler, L. Peterson, R. Hoffman and R. Arnoldy. (*J. Geophys. Res.*, vol. 64, pp. 597–610; June, 1959.) Balloon observations of X-ray bursts coincided with magnetic bays, radio noise absorption and auroral luminosity.

551.594.5 3310

Photoelectric Measurements of Hydrogen Emissions in Aurorae and Airglow—R. Montalbetti. (*J. Atmos. Terr. Phys.*, vol. 14, pp. 200–212; June, 1959.)

551.594.6 3311

Theory of the Origin of the Very-Low-Frequency Radio Emissions from the Earth's Exosphere—J. W. MacArthur. (*Phys. Rev. Lett.*, vol. 2, pp. 491–492; June 15, 1959.) The gyro-rotation of solar protons about the earth's magnetic field, corrected for Doppler shift, is used to derive a general equation of the theory [see 1575 of May (Gallet)].

551.594.6 3312

Interpretation of Smooth-Type Atmospheric Waveforms—F. Hepburn. (*J. Atmos. Terr. Phys.*, Vol. 14, pp. 262–272; June, 1959.) Accurate waveguide or pulse reflection analysis of the regular type of smooth waveform reveals a subdivision termed "pseudoregular." This type can be satisfactorily explained using waveguide or pulse reflection theory of propagation when complicated source conditions are considered and the discharge in the lightning channel is not assumed to be effectively instantaneous.

551.594.6:621.396.11 3313

Very-Low-Frequency Spectra of Atmospherics Propagated through the Ionosphere—T. Obayashi. (*Nature, London*, vol. 184, pp. 34–36; July 4, 1959.) A report of experimental results which support the mode theory suggested by Budden (1652 of 1951).

LOCATION AND AIDS TO NAVIGATION

621.396.933.2 3314

Radio Direction Finding in Three Dimensions—F. Adcock. (*Proc. IRE Aust.*, vol. 20, pp. 7–11; January, 1959.) "Methods of determining the direction of propagation of radio signals, in three dimensions, using spaced aerial systems are discussed and a direct reading system is proposed. The errors caused by the proximity of the earth are indicated."

621.396.962.2.029.65 3315

Phantom Radar Targets at Millimetre Radio Wavelengths—C. W. Tolbert, A. W. Straiton and C. O. Britt. (IRE Trans. on Antennas and Propagation, vol. AP-6, pp. 380–384; October, 1958. Abstract, Proc. IRE, vol. 46, p. 1976; December, 1958.)

621.396.962.3:621.374.3 3316

Using Magnetic Circuits to Pulse Radar Sets—A. Krinitz. (*Electronics*, vol. 32, pp. 42–43; July 3, 1959.) The size and weight of a pulse generator are reduced by the use of transistors and magnetic circuits with new core materials. An experimental generator described weighs 10 pounds, and gives a peak output power of 24 kw in 1 μsec pulses at a repetition frequency of 1700 cps.

621.396.963.3:621.314.7 3317

Transistorized Radar Sweep Circuits using Low Power—C. E. Veazie. (*Electronics*, vol. 32, pp. 46–47; June 26, 1959.) The deflection provided is sufficient for a 50° radar cathode-ray tube. Recovery time of transistors is less than 100 μsec. The sweep is initiated by a trigger pulse and the cathode-ray tube is blanked during flyback time.

621.396.968 3318

Reducing Mutual Radar Interference—K. H. Chase and J. L. Pierzga. (*Electronics*, vol. 32, pp. 39–43; July 10, 1959.) Considers gating, synchronization and blanking, pulse-width discrimination and pulse-shaping. Recent techniques use FM pulse and frequency diversity.

621.396.968.82 3319

The Elimination of Background Noise and Interference in Radar—J. Troude. (*Rev. HF, Brussels*, vol. 4, pp. 25–31; 1958.) The principle of operation and basic circuits of equipment for the elimination of random interference are described. A barrier-grid storage-tube circuit differentiates between wanted and unwanted components of the signal.

621.396.969.36 3320

The Nose-On Radar Cross-Sections of Conducting Right Circular Cones—J. E. Keys and R. I. Primich. (*Can. J. Phys.*, vol. 37, pp. 521–522; April, 1959.) A brief report of measurements made at wavelengths of 0.8565 cm and 3.426 cm on cones with diameters between 0.1λ and 3λ and nose angles between 8° and the limiting case of a thin disk.

MATERIALS AND SUBSIDIARY TECHNIQUES

535.215 3321

The Evaluation of Conductivity Glow Curves—K. W. Böer, S. Oberländer and J. Voigt. (*Ann. Phys.*, (*Lpz.*), vol. 2, pp. 130–145; October 12, 1958.) Glow curves were obtained for a number of activated single-crystal CdS specimens. A system of approximation equations is developed and applied successfully to the evaluation of photoconductivity parameters.

535.215:546.482.21 3322

Lifetime Measurements on Photoelectrons in CdS Crystals using a Kerr-Cell Apparatus—R. Caspary. (*Ann. Phys.* (*Lpz.*), vol. 2, pp. 182–190; October 12, 1958.) Short-duration phenomena occurring within 10^{-4} seconds after the start or the end of illumination are investigated with the apparatus described.

535.215:546.817.221:539.23 3323

Proportionality between Photoconductance and Time Constant in Lead Sulphide Films—H. E. Spencer. (*J. Appl. Phys.*, vol. 30, pp. 927–929; June, 1959.) The result that the relative change in photoconductance is proportional to the relative change in time constant is taken as evidence that this time constant is identical to the lifetime of the holes.

535.37:546.482.21 3324

Multiplet Structure of Excitons in CdS—R. G. Wheeler. (*Phys. Rev. Lett.*, vol. 2, pp. 463–465; June 1, 1959.)

535.376 3325

Excitation of Luminescent Materials by Ionizing Radiation—H. Kallmann and J. Dresner. (*Phys. Rev.*, vol. 114, pp. 71–79; April 1, 1959.) The effective volume of phosphor excited by a single ionizing event is calculated from experimental data.

537.226/.227 3326

Dielectric Properties of some Metaniobate and Metatantalate Ceramics—R. V. Coates and H. F. Kay. (*Phil. Mag.*, vol. 3, pp. 1449–1459; December, 1958.) Fifteen new metaniobates and metatantalates have been prepared. $MnTa_2O_6$, $BaNb_2O_6$ and $AlNb_2O_9$ have been found to be ferroelectric in addition to $PbNb_2O_6$, discovered by Goodman (see 1448 of 1954). $SrNb_2O_6$ and $BaTa_2O_6$ may have transition points below −183°C.

537.226/.227.2:546.431.824–31 3327
On the Anomaly in Residual Polarization of $BaTiO_3$ Ceramics—Y. Saito and S. Yamanaka. (*Commun. and Electronics*, no. 41, pp. 70–76; March, 1959.) The causes of anomalous polarization and methods of reducing it are investigated. See also 3400 of 1956 (Blood, *et al*).

537.226.33 3328
Contribution to the Phenomenological Theory of the Dielectric After-Effect in Ceramic Dielectrics—A. Hersping and K. Blank. (*Z. angew. Phys.*, vol. 10, pp. 371–376; August, 1958.) Results of measurements on four specimens are anzlyzed to obtain locus curves of complex polarizability which conform to the requirements of phenomenological theory.

537.227 3329
Growth of Ferroelectric Hysteresis Loops—J. C. Burfoot and R. V. Peacock. (*Proc. Phys. Soc., London*, vol. 73, pp. 973–975; June 1, 1959.) Discussion of the mechanisms effective in high-speed switching, and of the phenomenon of an increase in the amount of polarization reversal, observed with sinusoidal voltages applied to *c*-domain $BaTiO_3$ crystals.

537.227 3330
Study of Ferroelectric Transitions of Solid-Solution Single Crystals of $KNbO_3$-$KTaO_3$—S. Triebwasser. (*Phys. Rev.*, vol. 114, pp. 63–70; April 1, 1959.)

537.227 3331
On the Ferroelectric Behaviour of Potassium Dihydrogen Phosphate—J. Grindlay and D. ter Haar. (*Proc. Roy. Soc. London A*, vol. 250, pp. 266–285; March 10, 1959.)

537.227:546.431.824–31 3332
Asymmetric Hysteresis Loops and the Pyroelectric Effect in Barium Titanate—R. C. Miller and A. Savage. (*J. Appl. Phys.*, vol. 30, pp. 808–811; June, 1959.) A study of the asymmetry in the polarization direction of hysteresis loops as measured by the dynamic pyroelectric method.

537.228.1/.2:546.431.824–31 3333
The Electromechanical Properties of Barium Titanate—H. L. Allsopp and D. F. Gibbs. (*Phil. Mag.*, vol. 4, pp. 359–370; March, 1959. Plates.) Measurements in the form of oscillograms have been made of the electric field, the electric displacement and the strain on both single-crystal and ceramic $BaTiO_3$ over a wide temperature range.

537.228.1:546.431.824–31 3334
Electromechanical Behaviour of Single Crystals of Barium Titanate from 25 to 160°C—E. J. Huibregtse, W. H. Bessey and M. E. Drougard. (*J. Appl. Phys.*, vol. 30, pp. 899–905; June, 1959.) The strains induced by the polarization in a stress-free crystal are shown to be proportional to the square of the polarization, for values of polarization up to one-third of the spontaneous polarization at room temperature.

537.311.3:546.841.4–31:541.135.4 3335
Polarization Studies with Thorium Oxide—W. E. Danforth. (*J. Franklin Inst.*, vol. 266, pp. 483–491; December, 1958.) Studies of electrical polarization in the temperature range 600°–1300°C [2419 of 1956 (Danforth and Bodine)] have been extended to detailed measurement of the onset of polarization in an attempt to determine carrier density and mobility. Reasonable values of carrier density are obtained but mobility values appear to be too low and to vary with temperature in the wrong sense.

537.311.31:534.23–8 3336
Ultrasonic Attenuation by Electrons in Metals—E. I. Blount. (*Phys. Rev.*, vol. 114, pp. 418–436; April 15, 1959.) Calculation for a general band structure.

537.311.31:538.63 3337
Longitudinal Magnetoresistance of Metals in High Fields—B. Lüthi. (*Phys. Rev. Lett.*, vol. 2, pp. 503–504; June 15, 1959.) Several polycrystalline metals show a saturation effect in high fields.

537.311.33 3338
Dislocation Planes in Semiconductors—H. F. Mataré. (*J. Appl. Phys.*, vol. 30, pp. 581–589; April, 1959.) Dislocations and mainly dislocation planes, as the most important nonchemical imperfections in semiconductor crystals, are discussed from the point of view of their influence on carrier transport. The general properties of grain boundary planes of medium angle of misfit are reviewed and electrical effects treated. Device structures based on these properties are described.

537.311.33 3339
Capacitance and Barrier Height in Grain Boundaries—R. K. Mueller. (*J. Appl. Phys.*, vol. 30, pp. 546–550; April, 1959.) The capacitance of a grain-boundary is shown to depend on the height of the potential barrier ϕ for $\phi 2kT \log_e (N_d n_i)$. ($N_d$ is the number of uncompensated donors per cm^3, n_i the intrinsic carrier density.) For higher values of ϕ, the capacitance to a good approximation depends only on the bulk properties. Experimental data give a lower l mit for ϕ for specific samples.

537.311.33 3340
Drift and Hall Mobilities of Electrons in Nondegenerate Impure *n*-Semiconductors—M. S. Sodha and P. C. Eastman. (*Progr. theor. Phys.* (*Kyoto*), vol. 21, pp. 214–216; January, 1959.) A study of the variation with temperature of drift and Hall mobility, taking into account electron-electron scattering, and making use of the equations derived by Brooks (*Phys. Rev.*, vol. 83, p. 879; August 15, 1951.) and Herring (2642 of 1955).

537.311.33 3341
Auger Effect in Semiconductors—A. R. Beattie and P. T. Landsberg. (*Proc. Roy. Soc. London A*, vol. 249, pp. 16–29; January 1, 1959.) Lifetimes of excess electrons and holes are calculated, assuming the Auger effect between bands to be the only recombination mechanism. These are compared with experimental lifetimes in InSb. The mechanism envisaged may dominate radiative recombination above 240°K and accounts for the order of magnitude of the observed lifetimes ($\sim 10^{-8}$ sec) in the neighborhood of 330°K.

537.311.33 3342
Measurements of Photoconductivity and Lifetime of Charge Carriers in Semiconductors—G. Bedendo and D. Sette. (*Alta Frequenza*, vol. 27, pp. 437–471; October, 1958.) Detailed description of experimental apparatus and of measurements of minority-carrier lifetime and mobility made with it.

537.311.33 3343
Interdiffusion in Binary Ionic Semiconductors—R. F. Brebrick. (*J. Appl. Phys.*, vol. 30, pp. 811–815; June, 1959.) An analysis applicable to binary ionic conductors which are nondegenerate and in the exhaustion range.

537.311.33 3344
Lattice Screening in Polar Semiconductors—S. Doniach. (*Proc. Phys. Soc. London*, vol. 73, pp. 849–855; June 1, 1959.) The influence of space-charge screening on electron-phonon scattering is discussed, having regard to the high optical-mode vibration frequency.

537.311.33:537.228.1 3345
Interelectron Collisions and the "Temperature" of Hot Electrons—S. H. Koenig. (*Proc. Phys. Soc. London*, vol. 73, pp. 959–962; June 1, 1959.) The assumption made by Paige (518 of February) in interpreting piezoresistance measurements on *n*-type Ge, that the electron concentration is sufficiently high for an electron temperature to be defined for all carriers, is shown to be invalid. A large part of the observed electric-field dependence of the piezoresistance is caused by differences in electron temperature in different valleys.

537.311.33:537.228.5 3346
Stark Effect for Cyclotron Resonance in Degenerate Bands—J. C. Hensel and M. Peter. (*Phys. Rev.*, vol. 114, pp. 411–417; April 15, 1959.) Calculations indicate that line shifts may be sufficient to be useful in the study of the valence bands of Ge and Si.

537.311.33:537.32 3347
Thermoelectric Properties of Semiconductors—D. A. Wright. (*Electronics*, vol. 32, pp. 70–71; June 19, 1959.) Tabulated characteristics of ten common semiconductor materials.

537.311.33:537.322 3348
Electrical Determination of the Thermal Parameters of Semiconducting Thermoelements—C. Hérinckx and A. Monfils. (*Brit. J. Appl. Phys.*, vol. 10, pp. 235–236; May, 1959.) A procedure, for which only the mean temperature of the semiconductor is required, is described for determining the heat conductivity using the Peltier effect.

537.311.33:538.22 3349
Vapour-Deposited Films of Semiconducting III–V Compounds—K. G. Günther. (*Z. Naturforsch.*, vol. 13a, pp. 1081–1089; December, 1958. Plate.) Details are given of a "three-temperature method" of obtaining InAs and InSb films suitable for use as Hall generators. The properties of these films are investigated and compared with those of the bulk material.

537.311.33:539.23 3350
Comparative Investigation of the Variations of Resistance and Noise in Films of Gold, Silver and Copper—C. Uny. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 1655–1658; March 16, 1959.) Report of measurements of deviations from Ohm's law and scintillation noise in Au and Cu films. See 3495 of 1958 (Uny and Nifontoff).

537.311.33:[546.28+546.289 3351
Valence Semiconductors: Germanium and Silicon—G. I. Fen. (*Uspekhi. Fiz. Nauk.*, vol. 64, pp. 733–779; and vol. 65, pp. 111–132; April/May, 1959.) A general survey of the theory of semiconductors in which the crystal properties of Ge and Si, their conductivity in strong fields and at low temperature, the recombination and capture of carriers, and the drift of excess carriers, are investigated. 205 References.

537.311.33:[546.28+546.289 3352
Slow Capture of Holes and Electrons by Surface States on Germanium and Silicon at Low Temperatures—S. R. Morrison. (*Phys. Rev.*, vol. 114, pp. 437–444; April 15, 1959.) Field-effect decay curves are given for *n* and *p*-type Ge and for *p*-type Si. Decay depends on several parameters. It is suggested that the trapping centers involved are the fast states or recombination centers at the surface. Observed results are compared with a model.

537.311.33:[546.28+546.289 3353
The Heat Capacity of Pure Silicon and Germanium and Properties of their Vibrational Frequency Spectra—P. Flubacher, A. J. Lead-

better and J. A. Morrison. (*Phil. Mag.*, vol. 4, pp. 273–294; March, 1959.)

537.311.33:546.28 3354
Growth of Silicon Crystals Free from Dislocations—W. C. Dash. (*J. Appl. Phys.*, vol. 30, pp. 459–474; April, 1959.) Sources of dislocations which are important at particular stages of growth are isolated, and means found to minimize or eliminate them. Residual dislocations can be removed by a process of climb, caused by vacancy supersaturation, during growth along certain crystallographic axes.

537.311.33:546.28 3355
Temperature Dependence of Fractional Velocity Changes in a Silicon Single Crystal—F. Stein, N. G. Einspruch and R. Truell. (*J. Appl. Phys.*, vol. 30, pp. 820–825; June, 1959.) Results are given for compressional wave propagation along the 100 axis of Si, from 10 mc to 170 mc over the temperature range −60°C to +15°C.

537.311.33:546.28 3356
Ohmic Aluminium-*n*-Type Silicon Contact—S. L. Matlow and E. L. Ralph. (*J. Appl. Phys.*, vol. 30, pp. 541–543; April, 1959.) By careful heat control, an ohmic alloyed Al contact can be obtained. The theory is discussed.

537.311.33:546.28:535.215:538.65 3357
Photomagnetomechanical Forces on *p-n* Junctions Subjected to Nonuniform Illumination—P. Gosar. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 658–660; February 2, 1959.) Results are given of experiments with two differently constructed Si junction rings suspended in a magnetic field. Factors influencing sense and speed of rotation are discussed.

537.311.33:546.289 3358
Growth Phenomena in Ge Crystals—G. Zielasek. (*Z. Naturforsch* vol. 13a, pp. 1097–1098; December, 1958. Plate.) Various effects observed in crystals grown by the Czochralski method are illustrated and discussed.

537.311.33:546.289 3359
Measurements of the Dember Potential in Bulk Germanium—L. Stubbe and B. Gossick. (*J. Appl. Phys.*, vol. 30, pp. 507–508; April, 1959.) Measurements of the Dember voltage and photoconductivity enable the ratio of mobilities for electrons and holes to be determined. A value of 2:1 was obtained.

537.311.33:546.289 3360
Contact Potential Measurements on Cleaned Germanium Surfaces—A. B. Fowler. (*J. Appl. Phys.*, vol. 30, pp. 556–558; April, 1959.) The contact potential of the (110) face of the *n* and *p* regions of a *n-p* Ge crystal was measured in a vacuum of $<10^{-9}$ mm Hg using a Kelvin bridge technique. Results are discussed in terms of several surface state distributions.

537.311.33:546.289 3361
Drift Velocity Saturation in *p*-Type Germanium—R. D. Larrabee. (*J. Appl. Phys.*, vol. 30, pp. 857–859; June, 1959.) A method is described for making an ohmic contact to a semiconductor that does not inject minority carriers at high current densities.

537.311.33:546.289 3362
Anisotropic Mobilities in Plastically Deformed Germanium—R. A. Logan, G. L. Pearson and D. A. Kleinman. (*J. Appl. Phys.*, vol. 30, pp. 885–895; June, 1959.) An investigation of the effects of edge dislocations on mobility in *n*-type Ge.

537.311.33:546.289 3363
Observation of Thermally Induced Glide in Germanium—R. L. Cummerow. (*J. Appl. Phys.*, vol. 30, p. 946; June, 1959.)

537.311.33:546.289 3364
Exciton and Magneto-absorption of the Direct and Indirect Transitions in Germanium—S. Zwerdling. B. Lax, L. M. Roth and K. J. Button. (*Phys. Rev.*, vol. 114, pp. 80–89; April 1, 1959.) Low-temperature, high-resolution experiments on magneto-absorption effects in Ge have resolved fine structure in the direct interband transition and revealed structure in the indirect transition. Experimental data agree with theoretical predictions.

537.311.33:546.289 3365
Theory of Optical Magneto-absorption Effects in Semiconductors—L. M. Roth, B. Lax and S. Zwerdling. (*Phys. Rev.*, vol. 114, pp. 90–104; April 1, 1959.) Theory is developed on the basis of the effective-mass approximation. A detailed treatment of the direct transition in Ge is given. Results agree with experiment.

537.311.33:546.289 3366
Optical and Magneto-optical Absorption Effects of Group III Impurities in Germanium—P. Fisher and H. Y. Fan. (*Phys. Rev. Lett.*, vol. 2, pp. 456–458; June 1, 1959.) Measurements on single-crystal Ge at 5°K, and comparison with theory.

537.311.33:546.289 3367
Theory of Lattice Vibrations of Germanium—W. Cochran. (*Phys. Rev. Lett.*, vol. 2, pp. 495–497; June 15, 1959.) An extension of the Born-von Kármán theory to apply to a particular model of the Ge crystal.

537.311.33:546.289 3368
Impurity Conduction in Indium-Doped Germanium—J. S. Blakemore. (*Phil. Mag.*, vol. 4, pp. 560–576; May, 1959.) Studies were made of electrical conductivity at low temperatures in Ge crystals grown with In and Sb as majority and minority impurities respectively. The In densities were kept small to ensure that impurity conduction was controlled by the degree of compensation. Results illustrate trapping of electrons on preferred acceptor sites at the lowest temperature in accordance with the model proposed by Mott (see 2088 of 1957) and elaborated by Price (3117 of 1958). The trapping energy is $\sim 10^{-3}$ ev. The mobility of electrons appears to be independent of temperature which is at variance with Carwell's results (155 of 1957), although the increase of mobility with impurity density does follow the Carwell model.

537.311.33:546.289 3369
Light due to Recombination by means of Impurities in Germanium—A. Gosnet, O. Parodi and C. Benoit a la Guillaume. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 1628–1631; March 16, 1959.) An electron bound to an impurity atom can be recombined with a free hole. Radiation observed at energy levels of 0.707 and 0.734 ev in Sb-doped Ge is attributed to such recombinations.

537.311.33:546.289 3370
A Contribution to the Investigation of Chemical Action and Oxidation on Oriented Surfaces of Germanium Single Crystals—L. Gouskov and N. Nifontoff. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 1499–1502; March 9, 1959.)

537.311.33:546.289 3371
Microwave Field Dependence of Drift Mobility in Germanium—K. Seeger. (*Phys. Rev.*, vol. 114, pp. 476–481; April 15, 1959.) Mobility charges due to the application of both strong and weak fields depend on the carrier density, suggesting that carrier-carrier interaction plays an important part in the carrier-phonon scattering mechanism. Data agree well with those obtained with dc fields.

537.311.33:546.289 3372
Change of Surface Recombination Velocity of Germanium by γ-ray Irradiation, using *P-N-P* Alloy Transistor—K. Komatsubara. (*J. Phys. Soc. Japan*, vol. 14, pp. 383–384; March, 1959.)

537.311.33:546.289:534.23–8 3373
Acoustoelectric Effect in *n*-Type Germanium—G. Weinreich, T. M. Sanders, Jr., and H. G. White. (*Phys. Rev.*, vol. 114, pp. 33–44; April 1, 1959.) Theory and experiment are presented for the drag exerted on electrons in a solid by a travelling ultrasonic wave. The effect is generally very small but under certain conditions in *n*-type Ge may be appreciable.

537.311.33:546.289:537.32 3374
Strain-Induced Changes in the Seebeck Coefficient of *n*-Type Germanium—J. R. Drabble and R. D. Groves. (*Phys. Rev. Lett.*, vol. 2, pp. 451–452; June 1, 1959.) The results are used to deduce data on the anistoropy of the Seebeck tensor.

537.311.33:546.289:538.614 3375
Free-Carrier Faraday Effect in *n*-Type Germanium—A. K. Walton and T. S. Moss. (*J. Appl. Phys.*, vol. 30, pp. 951–952; June, 1959.) The experiments confirm the validity of a theoretical expression and show that effective masses can be determined accurately by Faraday-effect measurements.

537.311.33:546.289:548.73 3376
X-Ray Integrated Intensity of Germanium: Effect of Dislocations and Chemical Impurities—B. W. Batterman. (*J. Appl. Phys.*, vol. 30, pp. 508–513; April, 1959.)

537.311.33:546.289:548.73 3377
X-Ray Measurement of Microstrains in Germanium Single Crystals—L. P. Hunter (*J. Appl. Phys.*, vol. 30, pp. 874–884; June, 1959.)

537.311.33:546.289:621.314.7 3378
Avalanche Breakdown in *n-p* Germanium Diffused Junctions—R. Yee, J. Murphy. A. D. Kurtz and H. Bernstein. (*J. Appl. Phys.*, vol. 30, pp. 596–597; April, 1959.) Comparison of experimental data with theory.

537.311.33:546.3–1' 289'28 3379
Germanium-Silicon Alloy Junctions—Y. Matukura. (*J. Phys. Soc. Japan*, vol. 14, p. 374; March, 1959.) Experiments on the formation of *p-n* junctions are described.

537.311.33:546.681.19 3380
Elastic Moduli of Single-Crystal Gallium Arsenide—T. B. Bateman, H. J. McSkimin and J. M. Whelan. (*J. Appl. Phys.*, vol. 30, pp. 544–545; April, 1959.)

537.311.33:546.681.19 3381
Polarity of Gallium Arsenide Single Crystals—J. G. White and W. C. Roth. (*J. Appl. Phys.*, vol. 30, pp. 946–947; June, 1959.) Polarity has been determined by X-ray diffraction and the result correlated with a simple etching procedure.

537.311.33:546.681.19 3382
On the Growth of Gallium Arsenide Crystals from the Melt—S. G. Ellis. (*J. Appl. Phys.*, vol. 30, pp. 947–948; June, 1959.) Comparison of the growth rates on opposing {111} faces.

537.311.33:546.681.19 3383
Infrared Absorption and Electron Effective Mass in *n*-Type Gallium Arsenide—W. G. Spitzer and J. M. Whelan. (*Phys. Rev.*, vol. 114, pp. 59–63; April 1, 1959.) Absorption between 0.85 and 25 μ has been measured as a function of carrier concentration. The value obtained for the mass is larger than previously reported values.

537.311.33:[546.682.19+546.682.86]:537.32 3384

InAs and InSb as Thermoelectric Materials—R. Bowers, R. W. Ure, Jr., J. E. Bauerle and A. J. Cornish. (*J. Appl. Phys.*, vol. 30, pp. 930–934; June, 1959.) The results indicate that *n*-InAs and InSb in series could be used to generate electrical power with an efficiency of about 6 per cent when used in conjunction with *p*-type materials of similar characteristics.

537.311.33:546.682.19:538.63 3385

Anomalous Temperature Characteristic of the Hall Coefficient of Weakly *p*-Doped InAs—H. Rupprecht. (*Z. Naturforsch.*, vol. 13a, pp. 1094–1096; December, 1958.) Investigations show that the anomaly is due to surface conditions.

537.311.33:546.682.86 3386

Distribution Coefficients and Carrier Mobilities in InSb—A. J. Strauss. (*J. Appl. Phys.*, vol. 30, pp. 559–563; April, 1959.) Impurity distribution coefficients have been measured and dependence of the coefficient for Zn on extraction and rotation rates studied. Experimental Hall mobility data are compared with mobilities calculated from scattering theory.

537.311.33:546.682.86 3387

On the Delay Time in Plastic Flow of Indium Antimonide—J. W. Allen. (*Phil. Mag.*, vol. 3, pp. 1297–1305; November, 1958. Plate) A study of delay time as a function of temperature and stress for InSb in uniaxial compression. See also 3157 of 1958.

537.311.33:546.682.86 3388

Plasma Pinch Effects in Indium Antimonide—M. Glicksman and M. C. Steele. (*Phys. Rev. Lett.*, vol. 2, pp. 461–463; June 1, 1959.) Evidence of a pinch effect obtained in a single crystal of 0.39 mm × 0.60 mm cross-section, at currents greater than 5 amps.

537.311.33:546.814.231 3389

Electrical Properties of Stannous Selenide—S. Asanabe. (*J. Phys. Soc. Japan*, vol. 14, pp. 281–296; March, 1959.) The electrical resistivity, Hall coefficient and thermoelectric power of pure and impurity-doped SnSe crystals were investigated over the temperature range 100°K–800°K. Anomalies in the Hall coefficient and resistivity have been studied and possible causes are discussed.

537.311.33:621.314.63 3390

Effect of Variations in Surface Potential on Junction Characteristics—J. H. Forster and H. S. Veloric. (*J. Appl. Phys.*, vol. 30, pp. 906–914; June, 1959.) A filamentary structure has been used to compare the electrical properties of a Ge surface with those of an adjacent *p-n* junction intersecting the same surface. The observed dependence of junction characteristics on surface potential indicated that, for many surface-sensitive devices, an optimum value of surface potential could be specified.

538.221 3391

Ferromagnetics and Antiferromagnetics near the Curie Point—K. P. Belov. (*Uspekhi Fiz. Nauk*, vol. 65, pp. 207–286; June, 1958.) A report on the general thermodynamic theory as applied to the investigation of ferromagnetics and antiferromagnetics with a discussion of experimental data. 94 references.

538.221 3392

The Magnetization of Cobalt-Aluminium, Cobalt-Silicon, Iron-Aluminium and Iron-Silicon Alloys—D. Parson, W. Sucksmith and J. E. Thompson. (*Phil. Mag.*, vol. 3, pp. 1174–1184; October, 1958.)

538.221:538.652 3393

Single-Crystal Anisotrophy and Magnetostriction Constants of Several Ferromagnetic Materials including Alloys of NiFe, SiFe, AlFe, CoNi, and CoFe—R. C. Hall. (*J. Appl. Phys.*, vol. 30, pp. 816–819; June, 1959.)

538.221:539.23 3394

Thin-Film Magnetization Analysis—K. Chu and J. R. Singer. (Proc. IRE, vol. 47, pp. 1237–1244; July, 1959.) Magnetic energy relations for three important field directions are readily established by a graphical method.

538.221:539.23 3395

New Experimental Results concerning the Conductivity of Very Thin Films of Nickel—A. Colombani and G. Goureaux. (*Compt. Rend. Acad. Sci., Paris.* vol. 248, pp. 380–383; January 19, 1959.) See also 3548 and 3907 of 1958, and 1279 of April (Goureaux, *et al*).

538.221:621.318.122 3396

The Influence of Annealing in a Magnetic Field on the Magnetic Characteristics and the Domain Structure of Polycrystalline Perminvar Ring Specimens—K. Motzke. (*Ann. Phys. (Lpz.)*, vol. 2, pp. 163–167; October 12, 1958.) The highest values of relative remanence and maximum permeability are obtained when the magnetic field is applied after the specimen has cooled to about 170°C below the Curie temperature.

538.221:621.318.124 3397

Square-Loop Ferrites and the Testing of their Storage Properties—C. Heck and H. Reiner. (*Nachrichtentech. Z.*, vol. 11, pp. 360–369; July, 1958.) The specifications of magnetic properties and shapes of ferrite cores suitable for storage applications are summarized, and methods of testing and core selection for matrix production are outlined.

538.221:621.318.134 3398

Magnetic Spectra of Ferrites—L. A. Fomenko. (*Uspekhi Fiz. Nauk*, vol. 64, pp. 669–731; April, 1958.) A survey including 220 references, of basic investigations into magnetic spectroscopy of ferrites.

538.221:621.318.134 3399

Study of Oriented Polycrystalline Ferrites with Cubic Anisotropy—C. F. Kooi, R. W. Moss and D. C. Stinson. (*J. Appl. Phys.*, vol. 30, pp. 895–898; June, 1959.) One easy axis of each crystallite could be aligned by pressing the sample in a magnetic field.

538.221:621.318.134 3400

Hysteresis Losses of Ferrites with Anomalous Magnetization Loops—M. Kornetzki. (*Z. angew. Phys.*, vol. 10, pp. 368–371; August, 1958.) An interpretation is given of the characteristics of perminvar-type and isoperm-type ferrites.

538.221:621.318.134 3401

Effective Exchange Constant in Yttrium Iron Garnet—D. T. Edmonds and R. G. Petersen. (*Phys. Rev. Lett.*, vol. 2, pp. 499–500; June 15, 1959.)

538.221:621.318.134:538.569.4 3402

Ferrimagnetic Resonance Line Widths and *g*-Factors in Ferrites—R. L. White. (*Phys. Rev. Lett.*, vol. 2, pp. 465–466; June 1, 1959.) Outlines an extension of the theory of *g*-factors and line widths in rare-earth garnets to spinel-type ferrites.

538.221:621.318.134:538.632 3403

Ordinary Hall Effect in Fe_3O_4 and (NiO) 0.75 (FeO) 0.25 (Fe_2O_3) at Room Temperature—J. M. Lavine. (*Phys. Rev.*, vol. 114, pp. 482–488; April 15, 1959.) Measurements are reported on synthetic crystals. The relation between mobility and conductivity is discussed.

538.221:621.318.134:621.318.57 3404

Study of the Switching Time of a Square-Loop Ferrite—R. Vautier and A. Marais. (*Compt. Rend. Acad. Sci., Paris.*, vol. 248, pp. 560–663; February 2, 1959.) An experimental study in which it is shown that for asymmetric pulses the threshold field and the coercive field vary with the magnetic state preceding reversal.

538.221:621.318.57 3405

Flux Reversal in Soft Ferromagnetics—F. B. Humphrey and E. M. Gyorgy. (*J. Appl. Phys.*, vol. 30, pp. 939–939; June, 1959.) It is demonstrated that at least three mechanisms must be considered in describing flux reversal.

538.221:621.375.9:621.3.011.3 3406

Parametric Amplifiers and Cancellation of Ferromagnetic Hysteresis by Orthogonal Polarization—Angel and Boutry. (See 3244.)

538.221:621.385.833 3407

The Use of Magnetic Models in the Interpretation of Domain Effects on an Electron Beam—M. Blackman and N. D. Lisgarten. (*Phil. Mag.*, vol. 3, pp. 1069–1073; October, 1958. Plates.) "A magnetic unit has been constructed which consists of an array of oppositely magnetized steel wires and its effect on an electron beam has been studied. The phenomena observed are similar to the domain effects found when magnetic crystals are examined in the same apparatus, and are used in interpreting some of the features of the domain effects."

538.569.4:538.222 3408

Electron Free Precession in Paramagnetic Free Radicals—D. E. Kaplan and M. E. Browne. (*Phys. Rev. Lett.*, vol. 2, pp. 454–455; June 1, 1959.) Transverse relaxation times between 30 and 100 mμsec were measured directly by a pulse technique, and agreed well with cw results.

538.569.4:538.222 3409

Observation of the Co^{59} Nuclear Magnetic Resonance in Paramagnetic Salts—R. G. Shulman. (*Phys. Rev. Lett.*, vol. 2, pp. 459–460; June 1, 1959.)

538.569.4:538.222 3410

Some Weak Lines in the Paramagnetic Resonance Spectrum of Impure MgO Crystals—J. H. E. Griffiths and J. W. Orton. (*Proc. Phys. Soc. London*, vol. 73, pp. 948–950; June 1, 1959.)

538.569.4:538.222 3411

Paramagnetic Resonance of Impurities in CaF_2—J. M. Baker, W. Hayes and D. A. Jones. (*Proc. Phys. Soc., London*, vol. 73, pp. 942–945; June 1, 1959.)

538.569.4:539.2 3412

Dynamic Nuclear Polarization—M. Abraham, M. A. H. McCausland and F. N. H. Robinsin. (*Phys. Rev. Lett.*, vol. 2, pp. 449–451; June 1, 1959.) Study of the enhanced nuclear polarization produced in several materials by saturation of the microwave electron paramagnetic resonance of impurities.

539.2:548.4:537.311.33 3413

Calculation of the Lattice Oscillations in Crystals with Zinc Blende Structure—L. Merten. (*Z. Naturforsch.*, vol. 11a, pp. 662–679, 1067–1080; August and December, 1958.)

Part 1: Lattice Oscillations without allowing for Coulomb Forces.

Part 2: Influence of Coulomb Forces on Lattice Oscillations.

621.315.616:678.742.2 3414

On the Dielectric Loss of Oxidized High-Density Polyethylene—S. Okamoto and K. Takeuchi. (*J. Phys. Soc. Japan*, vol. 14, p. 378; March, 1959.)

621.315.618.9 3415

Properties of Octafluorocyclobutane, a Dielectric Gas—F. W. Blodgett. (*Commun. and*

Electronics, no. 41, pp. 63–66; March, 1959.) The electrical, physical and chemical properties of C_4F_8 are described.

MATHEMATICS

512.99:519.241 3416

Mean Absolute Value and Standard Deviation of the Phase of a Constant Vector plus a Rayleigh-Distributed Vector—J. R. Johler and L. C. Walters. (*J. Res. Nat. Bur. Stand.*, vol. 62, pp. 183–186; May, 1959.) An evaluation of the first- and second-moment integrals of the probability distribution for various values of average relative intensity of the Rayleigh-distributed component. The results of a quadrature evaluation of the integrals are tabulated over a wide range of values of average relative intensity ($K^2=0.010$ to 1000).

519.271.3:621.391 3417

A Discussion of Sampling Theorems—D. A. Linden. (Proc. IRE, vol. 47, pp. 1219–1226; July, 1959.) An extended application of the convolution theorem of Fourier analysis to sampling problems.

MEASUREMENTS AND TEST GEAR

531.76:621.374.32 3418

High-Accuracy Time Interval Measurements—H. D. Tanzman. (*Electronic Ind.*, vol. 18, pp. 62–67; January, 1959.) Time intervals in the range 0.9–1000 μsec may be measured to an accuracy of 0.01 μsec by a vernier technique using two pulsed crystal ringing oscillators, a coincidence detector and associated circuits. The equipment and its mode of operation are described in detail.

621.317.328.087.9 3419

The Statistical Evaluation of Field-Strength Measurements—F. von Rautenfeld. (*Rundfunktech. Mitt.*, vol. 2, pp. 178–180; August, 1958.) Preliminary note on a method of storing field-strength measurements on magnetic tape for subsequent evaluation by special equipment which is also briefly described.

621.317.33:621.316.993 3420

Ground Resistivity Measurement for the Telecommunication Engineer—H. Pech. (*Câbles and Trans.*, vol. 12, pp. 235–243; July, 1958.)

621.317.336:621.372.413 3421

Simple Method of Shunt Impedance Measurement—J. Dekleva and K. W. Robinson. (*Rev. Sci. Instr.*, vol. 30, pp. 470–471; June, 1959.) The shunt impedance of a resonant cavity is calculated from measurements of the amplitude of the signal transmitted through the cavity while a small aluminium sphere is moved inside the cavity.

621.317.34:621.396.822:621.385.2.029.63 3422

A Noise Diode for Ultra-High Frequencies Groandijk.—(See 3533.)

621.317.361.029.62:621.396.933 3423

Equipment for the Direct Measurement of the Transmission Frequency of Aircraft in Flight (Metre Waves)—J. Marique. (*Rev. HF, Brussels*, vol. 4, no. 3, pp. 53–57; 1958.) A general outline of equipment in which the output from a spectrum analyzer is displayed on the calibrated timebase of a long-persistence cro together with a reference signal from a frequency standard. Frequencies in the region of 120 mc may be measured with an accuracy within ±1 kc.

621.317.39:536.52:621.383 3424

An Experimental Pyrometer using a Phototransistor and designed for Radio-Tube Inspection—F. H. R. Alner and P. G. Van Zanten. (*Philips Tech. Rev.*, vol. 20, pp. 89–93; October 30, 1958.)

621.317.7:621.314.7 3425

Contribution on the Measurement of the Capacitance of Junction Transistors—S. Vojtášek. (*Slab. Obz., Praha*, vol. 19, pp. 351–354; June, 1958.) Two methods of measuring emitter and collector capacitances are described: 1) using the properties of a differential transformer; 2) based on the principle of parallel resonance. The differential transformer method can be applied for the measurement of transistor resistances which with the capacitances determine transistor performance at high frequencies.

621.317.7:621.314.7 3426

Transistor Current Gain—R. W. Smith and F. J. Hyde. (*Electronic Radio Engr.*, vol. 36, pp. 249–252; July, 1959.) "Apparatus used for a twin-channel comparator method of measuring complex values of current-gain α at frequencies between 1 and 210 Mc/s, and for a single-channel null method at frequencies up to 25 Mc/s are described. Results obtained by both methods in the overlap frequency range are shown, and the effects of stray capacitance on the measurements are illustrated."

621.317.7:621.314.7 3427

Automatic Transistor Alpha Measuring Set—D. E. Thomas and J. M. Klein. (*Rev. Sci. Instr.*, vol. 30, pp. 458–462; June, 1959.) This set is designed for α curve tracing with provision for a direct reading of α at discrete operating points; it operates on a null basis. It can be adapted to card punching of data or automatic selection of transistors.

621.317.729.087.4:681.142 3428

Note on a Semi-automatic Tracer of Curves of Equal Electrical Gradient—P. Coet. (*Rev. HF Brussels*, vol. 4, no. 4, pp. 85–88; 1958.) A detailed description of a two-probe servosystem for use with an electrolytic tank. The system error is ≤0.1 per cent.

621.317.733 3429

Differential Bridge for the Measurement of Impedance and Impedance Differences between 10 kc/s and 10 Mc/s—I. Eyraud and P. Bonijoly. (*Câbles & Trans.*, vol. 12, pp. 148–161; April, 1958.) Description of a direct-reading bridge with an error of 0.01 Ω. Applications to measurements on concentric-pair cables are given.

621.317.733:621.314.7 3430

Measurement of Transistor Parameters Using Bridge Techniques—M. J. Gay. (*Brit. Commun. Electronics*, vol. 6, pp. 430–432; June, 1959.) Practical details are given of the measurement of transistor *y* parameters at frequencies up to 30 mc using a 3-terminal bridge.

621.317.742:621.315.212 3431

Study of a Coaxial Reflectometer for Metre Waves—J. P. Bruynseels and R. Gonze. (*Rev. HF, Brussels*, vol. 4, no. 2, pp. 33–37.) Ideal coaxial reflectometers are discussed and a model with a characteristic impedance of 50 Ω is briefly described. Over the frequency range 100–900 mc directivity is >33 db and the residual voltage swr <1.05:1.

621.317.79 3432

An Instantaneous Microwave Polarimeter—P. J. Allen and R. D. Tompkins. (Proc. IRE, vol. 47, pp. 1231–1237; July, 1959.) A trimode turnstile junction with crystals in the rectangular arms is used as a linear mixer of local-oscillator and signal frequencies. The original phase and amplitude of the orthogonal components of the signal input are thereby preserved in the IF outputs, which are applied to the orthogonal plates of a cathode-ray tube. The result is an instantaneous display of the signal polarization.

621.317.794.029.64 3433

A Practical Microwave Power Standard—S. Fuse, Y. Takahashi and A. Furukawa. (*Rep. Elect. Commun. Lab., Japan*, vol. 7, pp. 1–13; January, 1959.) Details of a wattmeter for the measurement of power in the range 0.01–10 mw with an error of ±2 per cent at frequencies below 12 kmc. The measuring element is a fine-wire barretter operating in a Wien bridge circuit. Mount efficiencies of 95 per cent have been obtained.

OTHER APPLICATIONS OF RADIO AND ELECTRONICS

535.376.07 3434

Electroluminescent Panels for Automatic Displays—R. C. Lyman and C. I. Jones. (*Electronics*, vol. 32, pp. 44–47; July 10, 1959.) The phosphor is excited by an alternating voltage applied to a symbol of the required shape formed by an evaporated metal layer. The shaping of any figure from 0 to 9 by exciting suitable parts of a 9-segment outline is illustrated, and means of exciting those parts remotely by a coded pulse system is described.

535.824.3:621.314.63 3435

Millimicrosecond Light Source—A. Whetstone. (*Rev. Sci. Instr.*, vol. 30, pp. 447–450; June, 1959.) Reverse current was passed through a Si *p-n* junction containing diffused phosphorus. The light pulse emitted had a rise time shorter than 2 mμsec.

621.362:621.385.1 3436

Conversion of Heat to Electricity by Thermionic Emission—V. C. Wilson. (*J. Appl. Phys.*, vol. 30, pp. 475–481; April, 1959.) Electrons freed from a hot cathode are collected by a colder surface with a low work function, and their excess potential energy is available to do work in an electrical circuit in returning to the cathode. An experimental converter tube containing Cs vapour is described which gives an efficiency of about 10 per cent. See also 3437 and 3438 below.

621.362:621.385.1 3437

Theoretical Efficiency of the Thermionic Energy Converter—J. M. Houston. (*J. Appl. Phys.*, vol. 30, pp. 481–487; April, 1959.) Analysis of the converter described by Wilson (3436 above) indicates that efficiencies of 30 per cent or more are attainable.

621.362:621.385.1 3438

Calculation of the Performance of a High-Vacuum Thermionic Energy Converter—H. F. Webster. (*J. Appl. Phys.*, vol. 30, pp. 488–492; April, 1959.) Calculation of output *V/I* characteristics, and application to practical converter design.

621.365.5:621.387. 3439

An Experimental Induction-Heating Generator using Hydrogen Thyratrons—H. L. van der Horst and P. H. G. van Vlodrop. (*Philips Tech. Rev.*, vol. 20, pp. 101–107; October 30, 1958.) A circuit, similar in principle to the spark-gap oscillator but in which the spark gap is replaced by a hydrogen thyratron, is used in the generation of damped oscillations at frequencies up to 10 kc. Compared with generators employing transmitting tubes, this type is easier to operate, has a higher efficiency and needs no forced cooling.

621.384.6 3440

Cyclic Acceleration of Particles in High-Frequency Fields—V. I. Veksler and L. M. Kovrizhnykh. (*Zh. eksp. teor. Fiz.*, vol. 35, pp. 1116–1118; November, 1958.) The use of high-frequency fields for controlling the motion of particles in cyclic accelerators is outlined.

621.385.833 3441

Field Superposition to Increase the Sensitivity of the Electron-Optical Shadow Method—C. Schwink and H. Murrmann. (*Z. angew. Phys.*, vol. 10, pp. 376–379; August, 1958.) See also 1515 of 1956 (Schwink).

621.385.833 3442
New Electron Diffraction Apparatus for Continuous Recording—R. Thun. (*Rev. Sci. Instr.*, vol. 30, pp. 399–407; June, 1959.)

621.385.833 3443
Unusually Highly Biased Gun on the Electron Microscope—F. W. Bishop. (*Rev. Sci. Instr.*, vol. 30, pp. 468–469; June, 1959.) The advantages include reduction of contamination rate and of specimen heating.

621.385.833 3444
Contributions to Quantitative Electron Microscopy—E. Zeitler and G. F. Bahr. (*J. Appl. Phys.*, vol. 30, pp. 940–944; June, 1959.) The effect of various parameters on contrast in an electron-microscope image is analyzed.

621.385.833 3445
Theoretical and Experimental Investigation of the Field of a Series of Cylindrical Magnetic Lenses—P. Gautier and C. Latour. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 1637–1640; March 16, 1959.)

621.387.4:621.374.3 3446
Nucleonic Instrumentation—(*Nucleonics*, vol. 17, pp. 63–94; June, 1959.) A report comprising the following papers:
1) **Transistorization**—F. S. Goulding (pp. 64–71).
2) **Discriminators**—K. A. McCollom (pp. 72–77).
3) **Scalers**—H. F. Stoddart (pp. 78–81).
4) **An International Survey of Photomultipliers for Scintillation Counting**—J. Sharpe (pp. 82–85).
5) **G-M Counters**—K. van Duuren, A. J. M. Jaspers and J. Hermsen (pp. 86–94).

621.396.969:533.6.011.7 3447
A Method of Detection of Non-ionizing Shock Waves and the Use of Radio Waves for their Investigation—R. der Agobian. (*Compt. Rend. Acad. Sci., Paris*, vol. 248, pp. 1308–1311; March 2, 1959.) The discharge of a condenser through a tube of gas under conditions which do not normally give rise to ionization may be observed by means of a plasma produced by a secondary discharge.

621.398 3448
Radio Telemetry: Part 2—Techniques—P. Rohan. (*Proc. IRE, Aust.*, vol. 20, pp. 20–31; January, 1959.) A description of the instruments and equipment carried by the missile and the ground equipment for receiving and processing the information. Part 1: 3084 of September (Shimmins).

621.56:537.322.1 3449
Theory of Thermoelectric Cooling—R. Dahlberg. (*Z. angew. Phys.*, vol. 10, pp. 361–368; August, 1958.) The relation between applied voltage, cooling temperature obtained, and thermal loading of the cooling element is determined for static conditions, taking account of Thomson effect and contact resistance. The use of cascade circuits without thermal shunting by connecting leads, and the reduction of thermal conductivity of the thermoelement materials are discussed.

PROPAGATION OF WAVES

621.396.11 3450
Remarks on the Fading of Scattered Radio Waves—R. A. Silverman. (IRE Trans. on Antennas and Propagation, vol. AP-6, pp. 378–380; October, 1958.) An extension of a previous paper (3274 of 1957) by the inclusion of contributions to fading due to appreciable scattering angle variation and mean wind shear within the scattering volume.

621.396.11:551.510.52 3451
The Rotation of Polarization of Electric Waves in Inhomogeneous Isotropic Dielectrics with particular reference to the Troposphere—G. Eckart. (*Z. angew. Phys.*, vol. 10, pp. 393–396; August, 1958.) The effect observed by Chisholm, *et al* (238 of 1956) is explained on the basis of Maxwell's equations taking account of tropospheric inhomogeneities. See also 3634 of 1957.

621.396.11:551.510.52 3452
Tropospheric Scatter Propagation and Equipment—J. Fieguth. (*Proc. IRE, Aust.*, vol. 20, pp. 12–19; January, 1959.) A general discussion of theoretical aspects and of systems and equipment. Tests made in Australia over a 112-mile path at a frequency of 900 mc are briefly described.

621.396.11:551.510.535 3453
Propagation of Long-Distance H.F. Signals—K. Miya and M. Kawai. (*Electronic Radio Eng.*, vol. 36, pp. 263–271; July, 1959.) Description of the characteristics of HF pulse and cw signals propagated between the United Kingdom and Japan. Two propagation modes, including forward ground scattering and sporadic-E-layer reflection, are postulated to explain deviations from the classical m.u.f. and also from the great-circle path.

621.396.11:551.594.6 3454
Very-Low-Frequency Spectra of Atmospherics Propagated through the Ionosphere—Obayashi. (See 3313.)

621.396.11.029.45:551.510.535 3455
The Numerical Solution of Differential Equations governing the Reflexion of Long Radio Waves from the Ionosphere: Part 3—D. W. Barron and K. G. Budden. (*Proc. Roy. Soc. London A*, vol. 249, pp. 387–401; January 13, 1959.) The differential equations given in Parts 1 and 2 (2402 and 3390 of 1955) are cast in a simpler form by using a generalized matrix admittance function **A**, and the corresponding equations are obtained in a form suitable for numerical integration by a step-by-step process from properly chosen initial solutions. From the value of **A** so obtained for a point below the ionosphere the reflection coefficient matrix **R** is obtained. A theorem on the equivalence of certain directions of propagation is stated and proved.

621.396.11.029.6 3456
Investigations of Propagation over Non-uniform Terrain in the Frequency Bands I, II, and III—U. Kühn. (*Tech. Mitt. BRF, Berlin*, vol. 2, pp. 1–7, 41–43; February and May, 1958.) Results are given of field-strength measurements made by a mobile recording unit to assess coverage of television and vhf FM transmissions in East Germany. A statistical analysis was made to obtain probability distribution curves for field strength in various types of terrain for the three bands.

621.396.11.029.6 3457
Foreground Terrain Effects on Overland Trans U.H.F. Transmissions—L. G. Trolese and L. J. Anderson. (IRE Trans. on Antennas and Propagation, vol. AP-6, pp. 330–337; October, 1958. Abstract, Proc. IRE, vol. 46, p. 1975; December, 1958.)

621.396.11.029.6 3458
Effect of Mountains with Smooth Crests on Wave Propagation—I. P. Shkarofsky, H. E. J. Neugebauer and M. P. Bachynski. (IRE Trans. on Antennas and Propagation, vol. AP-6, pp. 341–348; October, 1958, Abstract, Proc. IRE, vol. 46, p. 1976; December, 1958.)

621.396.11.029.63 3459
Measurements of the Bandwidth of Radio Waves Propagated by the Troposphere Beyond the Horizon—J. H. Chisholm, L. P. Rainville, J. F. Roche and H. G. Root. (IRE Trans. on Antennas and Propagation, vol. AP-6, pp. 377–378; October, 1958.) A summary of experimental results obtained over a 188-mile path, using a frequency-modulated 2290-mc klystron transmitter and a wide-band travelling-wave-tube receiver. In both time and frequency domains results are consistent and suggest that multipath delays are less than predicted by Rice (1473 of 1953).

621.396.11.029.64:621.396.677.31 3460
A Rapid Beam-Swinging Experiment in Transhorizon Propagation—Waterman. (See 3198.)

621.396.812:551.510.535 3461
Equatorial Sunset Effect—A. M. Humby. (*Wireless World*, vol. 65, pp. 343–345; July /August, 1959.) The effect described is related to difficulties of communication with stations situated in equatorial regions. The salient feature is a considerable azimuthal scattering of signals for about two hours near the time of local sunset at the equatorial station, notably during equinox months of high solar activity. Performance figures for the Singapore/London and Colombo/London tele-printer circuits, operating in the frequency range 4–22 mc illustrate the effect. Osborne (2320 of 1952) has described a similar phenomenon.

RECEPTION

621.396.62 3462
High-Frequency Radio Receiver RX.5C—L. J. Heaton-Armstrong and J. D. Holland. (*Elect. Commun.*, vol. 35, no. 3, pp. 193–201; 1958.) Description of a dual space-diversity-type receiver designed primarily for the reception of frequency-shift telegraphy in the band 2–30 mc.

621.396.82:621.327 3463
Radio Interference from Fluorescent Lamps—H. J. J. van Boort, M. Klerk and A. A. Kruithof. (*Philips Tech. Rev.*, vol. 20, pp. 135–144; November 22, 1958.) From extensive tests it is concluded that the principle interference is caused by oscillations of positive ions in the potential minimum in front of the cathode. The effect is most troublesome on frequencies below 1500 kc. It can be largely suppressed by a delta filter in the supply leads.

621.396.822.1 3464
The Interference Protection Ratios Required for Common Channels and Adjacent Channels in A.M. Sound Broadcasting—E. Belger and F. von Rautenfeld. (*Rundfunktech. Mitt.*, vol. 2, pp. 172–177; August, 1958.) Listener tests were carried out, similar to those described in 1863 of 1956, to determine the reduction of interference resulting from a limitation in radiated bandwidth, and the improvement in reception by a small frequency-offset of common-channel carriers.

STATIONS AND COMMUNICATION SYSTEMS

621.376 3465
Investigation of Asymmetrical Modulation—G. Bronzi. (*Alta Frequenza*, vol. 27, pp. 486–509; October, 1958.) The distortion is calculated by harmonic analysis of the function of an asymmetrically modulated carrier; Ssb and dsb modulation systems are compared.

621.391 3466
Application of Information Theory to Amplitude Compression—J. A. Ville. (*Câbles & Transm.*, vol. 12, pp. 144–147; April, 1958.) Simple cases of signals of limited amplitude with super-imposed noise are studied from which the form of approach to the general problem is indicated.

621.391 3467
Variable-Length Binary Encodings—E. N. Gilbert and E. F. Moore. (*Bell Sys. Tech. J.*, vol. 38, pp. 933–967; July, 1959.) Theoretical

treatment of letter-to-binary-digit encodings which could be used for the storage and transmission of information.

621.391 3468
Recurrent Codes: Easily Mechanized, Burst-Correcting, Binary Codes—D. W. Hagelbarger. (*Bell Sys. Tech. J.*, vol. 38, pp. 969–984; July, 1959.) A class of code capable of correcting successive digital errors is discussed. It is shown that with a code redundancy $1/b$ errors of Kb or fewer digits will be corrected. (K, b are integers).

621.396.1:629.19 3469
Communicating in Space—L. P. Yeh. (*Electronic Ind.*, vol. 18, pp. 54–58, 159, 94–99; February and March, 1959.) Experimental data obtained from moon echo experiments and from satellite observations are analyzed and it is concluded that a fading margin of 20 db should be allowed for in a communication system. Problems associated with radio transmission and reception in space and between satellites and the earth are examined.

621.396.2:621.376.5 3470
Frequency Shifts Improve Pulse Communications—J. L. Hollis. (*Electronics*, vol. 32, pp. 66–69; June 19, 1959.) In the system described, the transmitter frequency and receiver tuning are shifted in synchronism as soon as a pulse is received. Thus signal energy arriving by paths other than the shortest is rejected by the receiver.

621.396.24:621.396.826 3471
The Effect of Echo on the Operation of High-Frequency Communication Circuits—D. K. Bailey. (IRE TRANS. ON ANTENNAS AND PROPAGATION, vol. AP-6, pp. 325–329; October, 1958. Abstract, PROC. IRE, vol. 46, p. 1975; December, 1958.)

621.396.41:621.376.3:621.396.813 3472
Echo-Distortion in Frequency Modulation—R. G. Medhurst. (*Electronic Radio Eng.*, vol. 36, pp. 253–259; July, 1959.) Consideration of the distortion in a frequency-division-multiplex FM trunk radio system caused by a small echo. Numerical results are presented graphically for 240-, 600-, and 900-channel systems with characteristics recommmended by the CCIR. The effect of pre-emphasis is discussed.

621.396.44:621.372.823 3473
Effects of Amplitude and Phase Distortion on a Signal Carried by a H.F. Wave. Application to Waveguides—M. Jouguet. (*Câbles & Transm.*, vol. 12, pp. 206–213; July, 1958.) The method of echoes (1177 of 1953) is applied to the study of distortion in a $\cos^2$ pulse transmitted in a circular waveguide.

621.396.5:534.76 3474
A Compatible Stereophonic Sound System—F. K. Becker. (*Audio*, vol. 43, pp. 17–18; May, 1959.) An outline is given of a system based on the "precedence effect" in the localization of sound sources. Channels are interconnected via delay lines, each channel carrying a direct signal and, in order to achieve compatibility, delayed signals derived from other channels of the system. Optimum parameters of the delayed signal in a two-channel system are discussed.

621.396.65:621.396.812.3 3475
Fading in Microwave Relays—K. Morita and K. Kakita. (*Rep. Elect. Commun. Lab., Japan*, vol. 6, pp. 325–370; September, 1958.) A statistical analysis of thermal and carrier-interference noise variations caused by fading on long-distance radio relay circuits.

621.396.73 3476
Pagemaster Receiver and Modulation Equipment—J. W. Young. (*Commun. and Electronics*, no. 41, pp. 45–52; March, 1959.) The general design of a selective paging system operating at 35 mc is described.

621.396.73 3477
City-Wide Personal Signalling at Allentown-Bethlehem, Pa—C. R. Kraus. (*Commun. and Electronics*, no. 41, pp. 52–55; March, 1959.) A general description is given of the operation and the equipment of a selective paging system.

621.396.73 3478
Systems Engineering of Personal Radio Signalling System—W. Strack. (*Commun. and Electronics*, no. 41, pp. 55–59; March, 1959.) Outlines the planning and design considerations for systems operating at 35 or 150 mc.

621.396.822.1:534.86 3479
Subjective Assessment of Shared-Channel Interference in Broadcasting—J. Köhler. (*Tech. Mitt. BRF, Berlin*, vol. 2, pp. 52–55; July, 1958.) Report on a series of subjective tests for various signal/interference ratios and levels using music or single-frequency tones as interference.

621.396.932.029.62 3480
The Application of V.H.F. Radio Communication to Port and Harbour Control—W. F. Bonage. (*Point to Point Telecommun.*, vol. 3, pp. 31–44; June, 1959.) A discussion of some of the problems involved in achieving a satisfactory service.

SUBSIDIARY APPARATUS

621.311.6.024:621.314.7 3481
Transistor Rectifier gives D.C. of Either Polarity—R. R. Bockemuehl. (*Electronics*, vol. 32, p. 76; June 19, 1959.) A diagonally symmetric power-transistor circuit which permits smooth load-current variation over a range of several amperes.

621.314.58:621.314.7 3482
Transistorized Three-Phase Power Supplies—W. Brannian. (*Electronic Ind., Electronic Operations Section*, vol. 18, pp. 02–05; January, 1959.) A description of the circuits required to convert a 28-volt dc supply to a 3-phase 400-cps 115-volt ac supply.

621.314.63+621.314.57 3483
Transistor Invertors and Rectifier-Filter Units—F. Butler. (*Electronic Engrg.*, vol. 31, pp. 412–418; July, 1959.) Details are given of 1) a symmetrical push-pull circuit with improved output waveform, and 2) a bridge connection for very high output power. Both operate from a 12-volt supply.

621.314.63:546.28:621.317.614 3484
Measuring Load Characteristics of Silicon Cells—R. D. Lynch. (*Electronic Equipm. Engrg.*, vol. 7, pp. 33–34; January, 1959.) A simple test circuit is described which can be inserted in any type of rectifier circuit.

621.314.63:546.289 3485
***P-N* Junctions in Germanium**—L. Černý, V. Husa, J. Kříž and J. Ladnar. (*Elektrotech. u. Maschinenb.*, vol. 76, pp. 233–236; June 1, 1959.) The improvement of the physical characteristics of junctions is discussed with particular reference to power rectifiers.

621.314.63:621.318.57:621.396.67 3486
Switching V.H.F. Power with Silicon Diodes—R. H. Mattson and S. H. Liu. (*Electronics*, vol. 32, pp. 58–59; June 19, 1959.) A coaxial type of switch designed for aircraft antenna systems.

621.318.5 3487
Ball Relays—(*Radio Rev. TV*, vol. 10, pp. 267–270; April, 1958.) A relay is described in which a ball, or a number of balls, of a permeable material, attracted to an energized core, simultaneously completes both the magnetic and the output circuits. Various forms of the basic design and their applications are given. The advantages are long contact life, low control power and simple construction.

TELEVISION AND PHOTOTELEGRAPHY

621.397.24 3488
The Equalization of Television Signals in Video-Frequency Transmission by Cable—H. Dobesch, F. Weide and H. Sulanke. (*Tech. Mitt. BRF, Berlin*, vol. 2, pp. 45–51; July, 1958.) The signal distortion arising in cable links between television camera and transmitter equipment is investigated and the design of equalizers for 200-meter and 400-meter cable links is outlined.

621.397.5:621.391 3489
A Method of Coding Television Signals based on Edge Detection—B. Julesz. (*Bell Sys. Tech. J.*, vol. 38, pp. 1001–1020; July, 1959.) "A method is described for transmitting digitalized video signals to reduce channel capacity from that needed for standard p.c.m. This method takes advantage of the inability of the human eye to notice the exact amplitude and shape of short brightness transients. The transmitted information consists of the amplitudes and times of occurrence of the 'edge' points of video signals. These selected samples are coarsely quantized if they belong to high-frequency regions, and the receiver then interpolates straight lines between the samples. The system was simulated on the IBM 704 computer. The processed pictures and obtained channel-capacity savings are presented."

621.397.6 3490
Delay Equalization for Residual-Sideband Transmission in Television—H. Hopf. (*Rundfunktech. Mitt.*, vol. 2, pp. 180–183; August, 1958.) The method of reducing delay distortion in German television transmitters is described.

621.397.6:621.372.57 3491
Application and Circuit Technique of Limiters in Television Equipment—W. Dillenburger. (*Frequenz*, vol. 12, pp. 255–261; August, 1958.) The use of limiter circuits at various points in the transmission chain is discussed and several such circuits are described.

621.397.61 3492
Television Transmitter Design—H. A. Teunissen. (*Philips Telecommun. Rev.*, vol. 20, pp. 89–95; March, 1959.) Discusses the general lines to be followed. Some examples of completed transmitters are mentioned.

621.397.61 3493
A New Range of Television Transmitters—P. W. L. van Iterson. (*Philips Telecommun. Rev.*, vol. 20, pp. 96–116; March, 1959.) The design of the modulator RF drive and power stages of a 5-kw and a 10-kw transmitter is given. Tetrodes are used in the power stages, and this, together with a special feedback circuit, leads to a smaller and more stable design of modulator.

621.397.61:621.372.54 3494
Vestigial Sideband Filters and Diplexers for Television Transmitters—J. A. van der Vorm Lucardie. (*Philips Telecommun. Rev.*, vol. 20, pp. 126–138; March, 1959.) Two types of diplexer and a vestigial-sideband filter are discussed. A "filterplexer" combining the properties of these and having a much more constant input impedance is described. Equalization problems are discussed.

621.397.61:621.372.55 3495
Video Correction Equipment—P. W. L. van Iterson. (*Philips Telecommun. Rev.*, vol. 20, pp. 117–125; March, 1959.) Equipment for stabilizing levels and providing gain and phase-error correction is described.

621.397.611.2 3496

A Single-Lens Reflex Television Camera—P. Lindner and E. Kosche. (*Tech. Mitt. BRF, Berlin*, vol. 2, pp. 7–10; February, 1958.) An optical-electronic view- and range-finder is described combining the advantages of both systems.

621.397.611.2 3497

Recent Investigations into the Operation of Image Orthicon Camera Tubes—R. Theile. (*J. Telev. Soc.*, vol. 9, pp. 45–59; April–June, 1959.) Characteristic defects in the picture quality of high-contrast scenes are considered *e.g.*, white border lines, loss of definition in horizontal white-to-black transients and geometric distortion in picture detail. A new method of operating the tube, line storage operation, is described whereby quality is improved at the expense of sensitivity. See 3994 of 1958 (Theile and Pilz).

621.397.611.2:535.81 3498

Modern Optics in Relation to Television—G. H. Cook. (*J. Telev. Soc.*, vol. 9, pp. 61–69; April–June, 1959.) The over-all optical performance of a television system is expressed by the peak-modulation/frequency graph derived by the use of a sine-wave test card. This representation is superior to a criterion based on lens resolving power. Characteristics of image-orthicon and vidicon lenses, and the design of zoom lenses are discussed.

621.397.621.2:621.373.444.1 3499

Frame Multivibrator and Diode Separator Circuit—M. C. Gander and P. L. Mothersole. (*Mullard Tech. Commun.*, vol. 4, pp. 260–264; May, 1959.) The higher sensitivity of multivibrator circuits to trigger pulses may be used to produce simplified frame-synchronizing-pulse separator circuits, in place of the blocking-oscillator types.

621.397.621.2:621.373.444.1 3500

Multi-triode Flywheel Synchronizing Circuit—M. C. Gander and P. L. Mothersole. (*Mullard Tech. Commun.*, vol. 4, pp. 265–272; May, 1959.) The development of a new circuit using a triode phase detector to control an anode-coupled multivibrator is described. The phase detector produces positive and negative control potentials for the oscillator stage.

621.397.7 3501

A Method of Designing Optimum Transmitter Networks in Bands IV and V—E. Eden. (*Rundfunktech. Mitt.*, vol. 2, pp. 187–196; August, 1958.) The relation between number of channels and transmitter distribution density is investigated. The effect of the variety of television standards on network planning in Europe is discussed in detail.

621.397.813 3502

A Simple Test Pattern for Assessing the Quality of Monochrome Television Installations—F. Below. (*Rundfunktech. Mitt.*, vol. 2, pp. 184–186; August, 1958.) The pattern is based on that developed by Kroebel (1393 of April).

TRANSMISSION

621.396.61 3503

On the Rating of S.S.B. and I.S.B. Transmitters—(*Point to Point Telecommun.*, vol. 3, pp. 24–30; June, 1959.) Methods of measurement are reviewed; it is necessary to know which method has been used when comparing performance figures.

621.396.61:621.314.7 3504

A 4-Watt 500-kc/s Transistor Transmitter—J. R. Nowicki. (*Mullard Tech. Commun.*, vol. 4, pp. 250–259; May, 1959.) Detailed description of the design of a light-weight unit with an estimated range of 50 miles over the sea. Full details are given of an automatic keying circuit based on a sequence of transistor multivibrators.

TUBES AND THERMIONICS

621.314.63:621.375.51 3505

Tunnel Diodes as High-Frequency Devices—H. S. Sommers, Jr. (Proc. IRE, vol. 47, pp. 1201–1206; July, 1959.) The behavior and uses of a particular type of heavily doped semiconductor junction diode [1784 of 1958 (Esaki)] are described. A voltage-controlled negative resistance, due to quantum-mechanical tunneling of carriers through the junction, enables the diode to be used as an oscillator at frequencies above 1 kmc, as a self-excited converter, and as a high-speed low-power bistable switch (switching time 2 mμsec).

621.314.63:621.396.822 3506

The Current Noise of Semiconductors, particularly of Germanium Point-Contact Diodes in the Frequency Range 1 kc/s–10 Mc/s—H. Bley. (*Nachrichtentech. Z.*, vol. 11, pp. 349–359; July, 1958.) The dependence of current noise on frequency and direct current flow is discussed with reference to measurements on commercial-type components for the case of 1) film-type resistors, and 2) semiconductor diodes, considering current flow in forward and reverse directions separately.

621.314.63:621.396.822 3507

Testing Diodes for R.F. Noise—J. C. Senn. (*Electronic Equipm. Engrg.*, vol. 7, pp. 28–32; January, 1959.) Measurements of reverse recovery transients in 197 Si and Ge diodes have been made in an attempt to correlate this effect with RF noise in the diode. If the recovery characteristic is stable with aging it can be used as a measure of noise tendency. Generally, the reverse transient effect is attributed to some uncontrolled factor in diode manufacture.

621.314.63+621.314.7].012.6:681.142 3508

Analogue Computer Measurements on Saturation Currents, Admittances and Transfer Efficiencies of Semiconductor Junction Diodes and Transistors—A. H. Frei and M. J. O. Strutt. (Proc. IRE, vol. 47, pp. 1245–1252; July, 1959.) In theoretical analyses of these devices a one-dimensional structure is normally assumed instead of the true three-dimensional structure with rotational symmetry. Solutions have been obtained for the latter structure using an analog computer; these are in fair agreement with experiment and yield some suggestions for better transistor dimensions.

621.314.7 3509

Modern High-Frequency Transistors—Yu. S. Ryabinkin. (*Uspekhi Fiz. Nauk*, vol. 65, pp. 689–719; August, 1958.) A description of several *p-n-i-p* and *n-p-i-n* types of Ge and Si transistor is included and their characteristics are tabulated. 87 references.

621.314.7 3510

Some Properties of Experimental 1000-Mc/s Transistors—R. F. Rutz and D. F. Singer. (*IBM J. Res. & Dev.*, vol. 3, pp. 230–236; July, 1959.) A diffused-base drift transistor and the coaxial oscillator circuit in which it has been tested are described in detail. An analysis of measurements showed that short-circuit current gains greater than unity can be obtained when a resonant condition exists in the passive circuit of the transistor. This condition is attributed to the low bulk resistance.

621.314.7:[061.3+061.4 3511

International Transistor Convention and Exhibition—(*Wireless World*, vol. 65, pp. 348–356; July/August, 1959.) Report of papers presented and devices and equipment exhibited at the IEE Convention held in London, May 21–27, 1959.

621.314.7:621.317.7 3512

Contribution on the Measurement of the Capacitance of Junction Transistors—Vojtášek. (See 3425).

621.314.7:621.317.7 3513

Transistor Current Gain—Smith and Hyde. (See 3426.)

621.314.7:621.317.7 3514

Automatic Transistor Alpha Measuring Set—Thomas and Klein. (See 3427.)

621.314.7:621.317.733 3515

Measurement of Transistor Parameters using Bridge Techniques—Gay. (See 3430.)

621.314.7:621.385.4 3516

High-Frequency Tetrode Transistors—J. T. Kendall and J. S. Walker. (*Brit. Commun. Electronics*, vol. 6, pp. 438–442; June, 1959.) Theory of operation, characteristics and applications of two types of grown-junction Si transistor are given.

621.314.7(083) 3517

1959 Transistor Interchangeability Chart—(*Electronic Ind.*, vol. 18, pp. 145, 173; March, 1959.) A tabulated list of 500 transistors, giving direct-replacement or nearest equivalent type with a note of intended application and dimension diagrams.

621.314.7.012.8 3518

The Junction Transistor as a Network Element at Low Frequencies: Part 3—Stabilization of the Operating Point, in particular with regard to Temperature Changes—J. P. Beijersbergen, M. Beun and J. te Winkel. (*Philips Tech. Rev.*, vol. 20, pp. 122–134; November 22, 1958.) Part 2: 2781 of August.

621.314.7.015.5 3519

The Breakdown Voltages of Germanium Transistors—R. E. Aitchison. (*Proc. IRE, Aust.*, vol. 20, pp. 32–34; January, 1959.) A representative range of Ge transistors has been tested and breakdown voltages which vary from 4 volts to over 100 volts, are tabulated. Reasons for these variations are discussed and a graph is given for determining the limiting value of breakdown voltage as a function of base-layer resistivity with base-layer thickness as a parameter.

621.383:535.215 3520

Photoresistors, Photodiodes and Phototransistors: General Remarks—P. Görlich, A. Krohs and W. Lang. (*Arch. tech. Messen*, no. 272, pp. 189–192; September, 1958.) The principal photoconductive materials are tabulated. 119 references.

621.383.2:535.371.07 3521

Image Fluctuations in Cascade Intensifiers—L. Mandel. (*Brit. J. Appl. Phys.*, vol. 10, pp. 233–234; May, 1959.) Random fluctuations of the number of photons making up the final image are evaluated in terms of the system parameters of the intensifiers. These fluctuations become appreciable only at very low light levels.

621.383.4 3522

An Image Converter Equipment for the Observation of Self-Luminous Gaseous Discharges—K. G. Beauchamp and A. J. Tyrrell. (*Electronic Eng.*, vol. 31, pp. 384–389; July, 1959.) "A high speed camera unit using a pulsed image convertor tube is described. Methods of gating this tube at potentials up to 20 kV are detailed and the advantages of deflecting the electronic image across the tube face are discussed. A considerable reduction in the bulk and complexity of an image convertor equipment is made possible by using a permanent magnet for focusing the tube and the equipment described includes a focusing unit of this type."

621.385.002.2 3523
Technological Problems in the Manufacture of Valves—W. Meier. (*Bull. schweiz. elektrotech. Ver.*, vol. 50, pp. 9–16; January 3, 1959.) The problems briefly discussed are 1) the production of vacuum-tight seals, 2) the preparation of highly emissive cathodes, 3) the maintenance of a high vacuum, and 4) the dissipation of heat.

621.385.029.6 3524
International Convention on Microwave Valves—(*Proc. IEE*, Part B, vol. 105, supplement no. 12, pp. 813–1040; 1958.) The text is given of the following papers which were among those read at the IEE Convention held in London, May 19–23, 1958. For titles of papers included in supplement no. 11, see 3159 of September.

Velocity-Modulation Tubes—I:

1) **A 5% Bandwidth 2.5 MW S-Band Klystron**—P. G. R. King (pp. 813–820).

2) **An Experimental High-Perveance Klystron Amplifier**—W. L. Beaver (pp. 821–823).

3) **Long-Transit-Time Multipactoring at Ultra-High Frequencies, and the Effect of Surface Emitting Layers**—D. K. Aitken (pp. 824–829).

4) **Phase Relationships in a Stagger-Tuned Klystrom Amplifier**—N. E. Dixon (pp. 830–832).

5) **A Three-Cavity L-Band Pulsed Klystron Amplifier**—A. H. W. Beck and P. E. Deering (pp. 833–838).

6) **The Choice of Beam and Coupling Parameters for Broad-Band Klystrons**—M. O. Bryant (pp. 839–842).

7) **An Experimental Investigation of Velocity-Modulated Electron Beams**—A. L. Cullen and I. M. Stephenson (pp. 843–854).

8) **Factors Influencing the Design of Multi-Cavity Klystrons**—H. J. Curnow (pp. 855–859).

9) **Studies in the Development of High-Power Klystrons**—K. H. Kreuchen and N. J. Diserens (pp. 860–869).

10) **A Theoretical and Experimental Study of a High-Power 400 Mc/s Klystron**—R. Latham, E. D. Dracott, M. Flinn, D. A. Gray, G. Nassibian, J. C. Vokes and G. Wexler (pp. 870–877).

Discussion (p. 878).

Measurements

11) **Measuring Modulation Noise from a High-Power C. W. Klystron Amplifier**—D. E. T. F. Ashby and R. B. Dyott (pp. 879–882).

12) **Analysis of a Thin High-Current-Density Electron Beam by Means of a Mechanical Scanner**—T. Fujii and A. Saburi (pp. 883–885).

13) **The Coaxial-Line Diode: A Rectifier of Microwaves**—P. O. Hawkins, H. J. Curnow and R. Redstone (pp. 886–889).

14) **Measurements of Field Patterns for a Comb-Type Slow-Wave Structure**—K. Kamiryo, H. Hozumi, Y. Shibata and Y. Fukushima (pp. 890–892).

15) **Electron Beam Technique for Measuring Circuit Velocity and Impedance**—H. R. Johnson and R. R. Bagi (pp. 893–896).

16) **Measuring Techniques for M-Type S- and X-Band Backward-Wave Oscillators**—D. D. Silvester (pp. 896–897).

17) **Technology of Carcinotrons for Short and Long Wavelengths**—G. Boucher (pp. 897–899).

Discussion (pp. 899–900).

Electron Optics—I

18) **The Electron Optics of High-Current-Density Electrostatic Electron Guns**—M. R. Barber and K. F. Sander (pp. 901–906).

19) **Non-Laminar Electron Beams in High Magnetic Fields**—T. W. Johnston (pp. 907–917).

20) **Electron-Ray Tracing in Electric and Magnetic Fidlds taking into account Space Charge**—J. R. Hechtel and K. R. Johne (p. 917).

21) **The Design and Performance of High-Current-Density Electron Guns**—P. C. Ruggles (pp. 918–927).

22) **The Effect of Anode-Aperture on a Dense Beam**—G. A. Stuart and B. Meltzer (pp. 928–931).

23) **A Method for the Calculation of Electron-Beam Contours in Axially Symmetrical Magnetic Fields, including the Gun Region**—W. Veith (pp. 932–938).

24) **High-Density Electron Gun with Magnetic Field**—C. Zlotykamin (pp. 939–940).

Discussion (pp. 941–943).

Velocity-Modulation Tubes—II

25) **Frequency Synchronization of an X-Band Reflex Klystron**—C. S. Aitchison (pp. 944–951).

26) **The Periodically-Loaded Travelling-Wave Multiple-Beam Klystron**—V. A. Heathcote, P. A. Lindsay, J. Barraclough and J. R. Newby (pp. 952–965).

27) **A New Development of the Monotron Oscillator**—J. R. Pickin and D. H. Trevena (pp. 966–968).

28) **An Experimental Annular Reflex Klystron**—D. J. Wootton, J. A. Lucken and R. C. Bannerman (pp. 969–977).

29) **The Tuning of Coupled-Cavity Reflex Klystrons**—P. E. V. Allin (pp. 978–984).

30) **An Investigation into the Design of a Gridless Low-Voltage Reflex Klystron**—R. C. Bannerman, J. A. Lucken and D. J. Wootton (pp. 985–992).

31) **Constant-Reflector-Voltage Klystron**—M. Kenmoku (pp. 993–1000).

32) **The Effect of Reactive Loads on the Frequency-Modulation Characteristics of Reflex Klystrons**—B. J. Mayo (pp. 1001–1005).

34) **A New Analysis of Electronic-Hysteresis and Secondary-Emission Phenomena in Local-Oscillator Reflex Klystrons**—R. P. Musson-Genon and C. Audoin (pp. 1006–1007).

35) **Swept-Frequency Klystron Operation at Constant Power Levels**—Th. Tamir and J. O. Spector (pp. 1008–1014).

Discussion (p. 1015).

Electron Optics—II (Crossed Field)

36) **Periodic Electrostatic Beam Focusing**—H. A. C. Hogg (pp. 1016–1020).

37) **Study of the Electron Trajectories in Guns for M-Type Tubes**—A. F. Leblond (pp. 1021–1023).

38) **An Electron-Trajectory Calculator and its Component Poisson Cell**—J. E. Rowe and R. J. Martin (pp. 1024–1032).

39) **Production and Focusing of High-Density Sheet Beams**—P. A. Sturrock (p. 1032).

40) **A Method of Focusing Electron Beams**—B. T. Murphy (pp. 1033–1039).

Discussion (p. 1040).

621.385.029.6 3525
Wide-Range Tuning Cavities for Plug-in Reflex Klystrons—R. Mather and J. Sharpe. (*Electronic Engrg.*, vol. 31, pp. 390–393; July, 1959.) Tuning ranges of 35 per cent have been obtained in the frequency band 5–12 km using cavities of modest dimensions. Test results are given.

621.385.029.6:621.3.032.26 3526
Axially Symmetric Electron Beams of Uniform Axial Velocity—E. J. Cook. (*J. Appl. Phys.*, vol. 30, pp. 860–865; June, 1959.) An analysis of the general class of long, dense electron beams with a uniform velocity profile.

621.385.029.6:621.3.032.269.1 3527
Toroidal Electron Guns for Hollow Beams—L. A. Harris. (*J. Appl. Phys.*, vol. 30, pp. 826–836; June, 1959.) A theory is given of convergent, high-efficiency electron guns for the production of dense hollow beams. Experimental measurements on a series of toroidal guns tend to confirm the space-charge theory. They also show that the annular slit in the anode seriously affects the uniformity of cathode emission and diverges the beam more than can be accounted for by present theories.

621.385.029.6:621.316.726 3528
Voltage Control of Magnetron Frequency—C. L. Cuccia. (*Electronics*, vol. 32, pp. 56–58; July 10, 1959.) "New type of crossed-field device uses magnetron structure to provide reflex action. Magnetron frequency is a function of magnetron current and may be varied by changing relative electrode voltages. Up to 18 Mc/s frequency shift at 800 Mc/s is obtainable."

621.385.029.65 3529
Voltage-Tunable Millimetre-Wave Oscillators—D. J. Blattner and F. Sterzer. (*Electronics*, vol. 32, pp. 62–63; June 19, 1959.) The characteristics and performance are given of two backward-wave oscillators in which frequency variations of 29–49 kmc and 48–74 kmc respectively are obtained by a single voltage adjustment.

621.385.032.213.13 3530
Carbonization of Thoria Cathodes—K. Hong. (*J. Appl. Phys.*, vol. 30, pp. 945–946; June, 1959.) Describes techniques for the manufacture of cathodes used in high-power magnetrons.

621.385.032.76 3531
About Forms of Electronic and Ionic Devices with Thermionic Cathodes—A. I. Vishnievsky, S. Sampath and C. S. Upadhyay. (*J. Indian Inst. Sci.*, Section B, vol. 40, pp. 150–164; October, 1958.) A design is proposed in which the tube envelope has a thin rectangular shape. Electrodes suitable for such an envelope are suggested. Advantages of this form of construction compared with cylindrical forms are reduced size and improved vacuum properties.

621.385.1:621.362 3532
Conversion of Heat to Electricity by Thermionic Emission—Wilson. (See 3436.)

621.385.2.029.63:621.317.34:621.396.822 3533
A Noise Diode for Ultra-High Frequencies—H. Groendijk. (*Philips Tech. Rev.*, vol. 20, pp. 108–110; October 30, 1958.) An experimental tube for use at frequencies up to 1000 mc.

621.385.3 3534
The Thimble Tubes—(*Electronic Ind.*, vol. 18, pp. 64–65, 97; April, 1959.) The method of construction and characteristics of the Type-7077 UFH grounded-grid triode and the "Nuvistor" low-power triode are described. Both valves are designed for high performance and reliability under extreme operating conditions.

621.387:621.396.96 3535
Factors Leading to the Development of Ceramic Hydrogen Thyratrons—H. N. Price and A. W. Coolidge. (*Commun. and Electronics*, no. 41, pp. 76–80; March, 1959.) The application in high-power radar equipment is considered.

THE INSTITUTE OF RADIO ENGINEERS

(Incorporated)

CONSTITUTION

(*Effective April 20, 1959*)

ARTICLE I

NAME AND OBJECT

SEC. 1—The name of this organization shall be The Institute of Radio Engineers, Incorporated, hereinafter called the IRE.

SEC. 2—Its objects shall be scientific, literary, and educational. Its aims shall include the advancement of the theory and practice of electronics, radio, allied branches of engineering, and of the related arts and sciences, their application to human needs, and the maintenance of high professional standards among its members. Among the means to this end shall be the holding of meetings for the reading and discussion of professional papers and the publication of papers, discussions, communications, and such other matters as may be appropriate for the fulfillment of its objects.

SEC. 3—The territory in which its operations are to be conducted is the United States of America and other countries.

ARTICLE II

MEMBERSHIP

SEC. 1—The membership of the IRE shall consist of members qualified to hold professional membership grades, who shall be designated as voting members, and members admitted to non-professional grades, who shall be designated as non-voting members.

SEC. 2—The grades of membership and the member qualifications for each and the requirements for admission, transfer, and severance of members shall be specified in the Bylaws.

ARTICLE III

DEFINITIONS

SEC. 1—The term "member" when printed without an initial capital, where used in this Constitution and by Bylaws, includes all grades of membership. The term "voting member" where used in this Constitution and the Bylaws means a member entitled to vote on IRE matters.

SEC. 2—The term "allied" as used in this Constitution and the Bylaws refers to electronics, electrical communication, radio, and such other technical fields as are directly contributory to, or derived from, electrical science.

SEC. 3—The term "Bylaw" as used in this Constitution refers to any Bylaws adopted by the Board of Directors as herein set forth and amended as hereinafter provided.

ARTICLE IV

DUES AND FEES

SEC. 1—Dues and fees shall be specified in the Bylaws.

SEC. 2—Under exceptional circumstances, the payment of dues and fees may be deferred or waived in whole or in part by the Board of Directors.

ARTICLE V

ANNUAL ASSEMBLY, BOARD OF DIRECTORS, AND PRESIDENT

SEC. 1—There shall be an annual Assembly composed of Delegates elected by the voting members. The annual Assembly shall receive reports and perform such other functions as required by law or specified in the Bylaws.

SEC. 2—There shall be a Board of Directors which shall be the governing body of the IRE, shall manage its affairs, and shall consist of Directors elected by the voting members and Directors elected by the annual Assembly. These Directors shall include the President and the two most recent Past Presidents.

SEC. 3—The election by the voting members of each Delegate shall signify the election of the same individual as a Director.

SEC. 4—The president shall be elected by the voting members, by virtue of which election he shall also be a Delegate-at-large and a Director-at-large.

ARTICLE VI

DELEGATES, DIRECTORS, AND OFFICERS

SEC. 1—The number of Directors elected by the voting members together with the number of Directors elected by the annual Assembly shall be not less than nine nor more than fifty, but in no case below the minimum prescribed by law. The number of Directors elected by the voting members shall not be less than two-thirds of the total number.

SEC. 2—All of the Directors shall be of full age and at least one shall be a citizen of the United States and a resident of the State of New York.

SEC. 3—The term of each Delegate elected by the voting members shall run concurrently with the term of his office as Director and shall be not less than two nor more than five years, except that the term of office of the President as a Delegate-at-large and Director-at-large shall be three years, and such terms of office of such Delegates and Directors shall begin with the first annual Assembly after their election and acceptance.

The terms of office of Directors elected by the annual Assembly shall be one year.

SEC. 4—The United States and Canada and other areas, at the discretion of the Board of Directors, shall be divided into geographical districts known as Regions, which shall be specified in the Bylaws. The voting members of each Region shall elect a Delegate to the annual Assembly designated as its Regional Delegate who shall also by virtue of such election be a Director designated as its Regional Director. The election of Regional Delegates and Regional Directors under this paragraph shall be conducted as specified in the Bylaws.

The voting members of the IRE shall elect Delegates to the annual Assembly designated as Delegates-at-large who shall also by virtue of such election be Directors-at-large. The number of Directors-at-large elected by the voting members shall not be less than the number of Regional Directors.

The number and method of election of Delegates-at-large and Directors-at-large shall be specified in the Bylaws.

SEC. 5—If the Directors shall not be elected on the day designated by law or fixed in the Bylaws, the corporation shall not for that reason be dissolved; but every Director shall continue to hold his office and discharge his duties until his successor has been elected.

SEC. 6—The Corporate Officers of the IRE shall be the President, one or more Vice Presidents as specified in the Bylaws, the Secretary, the Treasurer, and the Editor.

SEC. 7—The terms of office for all Corporate Officers shall begin with the annual meeting of the Board of Directors and shall terminate at the beginning of the following annual meeting of the Board of Directors or at such time as their successors are elected and accept.

SEC. 8—The terms of office of the Vice Presidents, Secretary, Treasurer, and Editor, as Directors shall be one year.

SEC. 9—No Corporate Officer or Director shall receive, directly or indirectly, any salary, traveling expenses, compensation, or emolument from the IRE either as such Officer or Director or any other capacity, unless authorized by the Bylaws or by the concurring vote of two-thirds of all the Directors at a regularly constituted meeting.

No Director or other Officer shall be interested, directly or indirectly, in any contract relating to the operations of the IRE, nor in any contract for furnishing supplies thereto, unless authorized by the Bylaws or by the concurring vote of two-thirds of the Directors at a regularly constituted meeting.

ARTICLE VII

MANAGEMENT

SEC. 1—The President shall be the regular presiding officer at meetings of the Board of Directors and the annual Assembly. He shall be an ex-officio member of each committee.

A Vice President shall assume the duties of the President in the absence or incapacity of the President, as specified in the Bylaws. Duties of Vice Presidents and the order of priority regarding assumption of presidential duties shall be specified in the Bylaws.

In the event of the absence or incapacity of both the President and all of the Vice Presidents, the Board of Directors shall elect a chairman from its membership who shall perform the presidential duties during such absence or incapacity of the President and all of the Vice Presidents. The tenure of such temporary chairman shall be at the discretion of the Board of Directors provided, however, that said temporary chairman shall not serve longer than the unexpired term of the incumbent President.

SEC. 2—The Secretary shall be responsible for the preparation of all meetings of the Board of Directors, the annual Assembly, and all principal meetings of the IRE. He shall be responsible for the keeping of full records of all the meetings and activities of the IRE, the records of the membership, such other records as are required by law, and shall be responsible for the provision of such information from them as is requested by the Board of Directors.

SEC. 3—The Treasurer, under the control of the Board of Directors, shall have general supervision of the fiscal affairs of the IRE and shall be responsible for the keeping of records thereof.

SEC. 4—All funds received by the IRE shall be deposited in an account requiring the signatures of at least two of the Corporate Officers for withdrawal. The Board of Directors shall designate those Corporate Officers authorized to execute such withdrawals.

SEC. 5—The number of Directors required to constitute a quorum shall be stated in the Bylaws but shall not be less than one-third of the total number of Directors.

SEC. 6—The number of Delegates required to constitute a quorum at the annual Assembly shall be stated in the Bylaws but shall not be less than one-third of the total number of Delegates or if one-third be nine or more, not less than nine.

SEC. 7—The chairman of the Board of Directors shall have no vote on the Board unless the votes of the other Directors present are equally divided.

SEC. 8—The chairman of the annual Assembly shall have no vote on the Assembly unless the votes of the other Delegates present are equally divided.

SEC. 9—The Board of Directors may make, amend, or revoke the Bylaws made by the IRE. Proposed Bylaw changes and reasons therefore shall be mailed to all IRE Directors at least twenty days before the stipulated meeting of the Board of Directors at which the vote shall be taken. Two-thirds of all votes cast at that meeting shall be required to approve any new Bylaw, amendment, or revocation.

SEC. 10—All committees, unless otherwise provided for in the Bylaws, shall be appointed by the Board of Directors or in such manner as the Board of Directors may designate.

SEC. 11—The fiscal year of the IRE shall be defined in the Bylaws.

ARTICLE VIII

NOMINATION AND ELECTION OF OFFICERS, DIRECTORS AND DELEGATES

SEC. 1—On or before July first of each year, the Board of Directors shall submit to all voting members a list of nominees for Delegates, Directors, the President, and such Officers as may be specified in the Bylaws to be elected by the voting members for the coming term, in accordance with the Bylaws. Submission may be by publication.

SEC. 2—Nominations by petition may be made by letter to the Board of Directors setting forth the name of the proposed candidate and the office for which the candidate is desired to be nominated, provided such letter is received at the general offices of the IRE no later than 12 o'clock noon on the Friday prior to August fifteenth. Such a petition shall be signed by at least one per cent of the total number of voting members as listed in the official membership records of the IRE at the end of the previous year, but in no case shall the number be less than one hundred.

SEC. 3—On or before September first, the Board of Directors shall submit to all voting members as of August fifteenth, a ballot listing all nominees to be voted upon by the voting members, in accordance with the Bylaws.

SEC. 4—The annual Assembly of the IRE shall elect additional Directors and such of the Corporate Officers as shall also serve as Directors, as provided by the Bylaws. Other Corporate Officers as specified in the Bylaws shall be elected by the Board of Directors.

SEC. 5—The Board of Directors is authorized to fill a vacancy occurring in the Board of Directors or among the Corporate Officers for a period not to exceed the unexpired term of the office becoming vacant, provided, that if a vacancy shall occur in the office of a Regional Director, then the geographical Region which elected such Director shall have the right to elect a Director to serve the remainder of such term. If such geographical Region shall fail to elect a Director within sixty days after notice of such vacancy, as provided in the Bylaws, then the same may be filled by the Board of Directors.

SEC. 6—The existence of vacancies in the Board of Directors or among the Corporate Officers shall be determined by the Board of Directors in accordance with the Bylaws.

ARTICLE IX

MEETINGS

SEC. 1—There shall be an annual Assembly of Delegates-at-large and Regional Delegates of the IRE during January of each year. The President of the IRE shall preside. The annual Assembly may be held elsewhere than in the State of New York.

SEC. 2—There shall be an annual meeting of the Board of Directors during January of each year following the annual Assembly. This annual meeting may be held elsewhere than in the State of New York.

SEC. 3—Additional meetings and special meetings of the Board of Directors shall be held in accordance with the Bylaws.

SEC. 4—The Board of Directors may authorize the establishment of groups of members for promoting the objectives and interests of the IRE in accordance with the Bylaws. The

Board of Directors may, at its discretion, terminate the existence of any such group.

ARTICLE X

AMENDMENTS

SEC. 1—Amendments to this Constitution shall be made by ballot of the voting members. Amendments may be proposed by a resolution adopted by vote of at least two-thirds of those present at a regularly constituted meeting of the Board of Directors or by a petition signed by at least one per cent of the total number of voting members as listed in the official membership records of the IRE at the end of the previous year, but in no case shall the number be less than one hundred. A copy of such proposed amendment or amendments, if lawful, shall be mailed with a ballot to each voting member at least sixty days before the date designated for counting the ballots. Balloting shall be in accordance with the Bylaws.

SEC. 2—If the total ballot vote be at least twenty per cent of the total voting membership and if at least sixty-seven per cent of all ballot votes cast shall be favorable, the proposed amendment or amendments shall be adopted. Voting members shall be notified as soon as practicable after the adoption by publication.

SEC. 3—Amendments to this Constitution shall take effect thirty days after adoption, but if by amendment Officers and Officers-elect are changed in status or the number of Directors is reduced, each Officer and each Director shall continue to serve until his term expires.

BYLAWS

(*Effective August 19, 1959*)

TABLE OF CONTENTS

100. DEFINITIONS

1. The term "all the rights and privileges" refers to right to participate fully in the activities of the IRE in furtherance of its objects as defined in the Constitution.
2. The term "group of members" refers to any classification of members whether on a geographical territorial basis or on a particular technical or professional interest basis, that has been authorized by the IRE.
3. The terms "meetings of the IRE" or "meetings of IRE members" refers to meetings authorized by the IRE, with or without geographical or other limitation, held to further the objects of IRE.
4. The term "IRE Headquarters" refers to the principal office of the IRE at which the Executive Secretary and his staff are located.
5. The term "Constitution" refers to the IRE Constitution.
6. The term "Executive Committee" refers to the Executive Committee of the IRE.
7. The term "school of recognized standings" refers only to schools which award degrees for not less than four years of full-time academic study in electronics, radio, allied branches of engineering or the related arts and sciences, and are designated as such by the Executive Committee.
8. The word "junior" when used in the terms "junior past President," "junior past Regional Director," "junior past Chairman" or other titles of IRE official positions refers to the immediate past holder of the office referred to.
9. The word "senior" when used in the term "senior past President" or other titles of IRE official positions refers to the holder of the office referred to next preceding the immediate past holder of the office.

200. MEMBERSHIP

201. *Grades*

1. The grades of IRE membership are:
   a. Fellows
   b. Senior Members
   c. Members
   d. Associates
   e. Students
   f. The designation "Life Member" is applicable to all members who have attained the age of 65 years and who have been members of the IRE for at least 35 years.
2. The designation "Group Affiliate" refers to persons entitled to participate in IRE Professional Group activities under provisions established by the Executive Committee under the Bylaws.

202. *Rights and Privileges*

1. Fellows shall be entitled to all rights and privileges of the IRE.
2. Senior Members shall be entitled to all rights and privileges of the IRE except the right to serve on the Awards and Fellow Committees.
3. Members shall be entitled to all rights and privileges of the IRE except the right to hold any corporate office and the office of Director, and to serve on the Admissions, Awards and Fellow Committees.
4. Associates shall be entitled to attend meetings of IRE members, to vote only on matters presented to groups of members, to be appointed on committees when specifically approved by the Board of Directors or the Executive Committee, but shall not have the right to hold any office.
5. All voting members shall have the right to use the IRE emblem (dark background) on letter heads and business cards to indicate their professional membership status in the IRE, provided that a sample of each letterhead or business card is first submitted for approval to the IRE Executive Secretary, who shall have the discretion to refuse such permission if the intended use of the emblem does not meet ethical and professional standards.
6. The emblem of the IRE is protected by trademark registrations and shall be reproduced only in connection with official business of the IRE or to indicate professional membership status in the IRE.
7. Emblems purchasable from IRE Headquarters may be worn by members.
8. Assertion of IRE membership by members is authorized, including the right to use the following abbreviations:
   Fellow IRE
   Sen. Mem. IRE
   Mem. IRE
   Assoc. IRE
9. Every member of the IRE shall receive the PROCEEDINGS OF THE IRE.
10. Every Student member shall receive the IRE STUDENT QUARTERLY.

203. *Qualifications*

1. Fellow: The grade of Fellow is one of unusual professional distinction and shall be conferred only by invitation of the Board of Directors upon a person of outstanding and extraordinary qualifications and experience in the fields of electronics, radio, allied branches of engineering or the related arts and sciences, who meets the requirements for Senior Member as stated in these Bylaws and who has been a member in any grade for a period of five years preceding the year of nomination, except that the five-year provision in any individual case may be waived for cause by the Board of Directors.
2. Senior Member is the highest professional grade for which application may be made and shall require experience or attainment reflecting professional maturity. For admission or transfer to the grade of Senior Member, a candidate shall be an engineer, scientist, technical executive, or teacher, in the fields of electronics, radio, allied branches of engineering or the related arts and sciences; as such, he shall have been in the active practice of his profession for at least ten years and shall have attained distinction as measured by above-average performance over a period of at least five of these years, such performance including one or more of the following:
   a. Publication of important original engineering or scientific papers, books or inventions, or
   b. Technical direction with evidence of accomplishment of important scientific or engineering work, or
   c. Creative contributions to the welfare of the scientific or engineering profession, or
   d. Establishment or furtherance of important scientific or engineering courses in a school of recognized standing, or
   e. Contributions equivalent to those of "a" to "d" above in such areas as technical editing, patent prosecution or patent law, provided these contributions serve to advance progress substantially in the fields of electronics, radio, allied branches of engineering or the related arts and sciences.
3. Member is a professional grade limited to those who have demonstrated professional competence in the fields of electronics, radio, allied branches of engineering or the related arts and sciences. For admission or transfer to the grade of Member, a candidate shall be either:
   a. An engineer or scientist in the fields of electronics, radio, allied branches of engineering or the related arts and sciences, who shall have had at least three years of professional experience.
   b. A teacher of electronics, radio, allied branches of engineering or the related arts and sciences, for at least three years, who shall have held the rank of instructor or higher, and shall have participated in planning and conducting courses.
   c. A person regularly employed in the fields of electronics, radio, allied branches of engineering or the related arts and sciences, for at least six years who, by experience, has demonstrated competence in work of a professional character.
   d. An executive who, for at least six years, has had under his direction important technical, engineering or research work in the fields of electronics, radio, allied branches of engineering or the related arts and sciences.
4. Associate: For admission or transfer to the grade of Associate, a candidate shall be interested in the theory or practice of electronics, radio, allied branches of engineering or the related arts and sciences.
5. Student: For admission to the grade of Student, a candidate shall be:
   a. Devoting a major portion of his time as a registered student in a regular course of study in engineering or science in a school of recognized standing, or
   b. Devoting a major portion of his time as a resident and registered student in a regular course of study in electronics, radio, allied branches of engineering or the related arts and sciences, in a technical institute or other school approved by the Executive Committee for Student membership purposes only.

6. Membership in the Student grade shall not extend more than six months beyond the termination of his student status as described in these Bylaws.
7. Graduation from a course of study of at least four years duration in electronics, radio, allied branches of engineering or the related arts and sciences, in a school of recognized standing, may be accepted as equivalent to three years' professional experience in those fields.
8. Graduation from a course of study of at least two years' duration in electronics, radio, allied branches of engineering or the related arts and sciences, in a technical institute approved by the Executive Committee, may be accepted as equivalent to one year's professional experience in those fields.
9. Full-time graduate work, or part-time graduate work with teaching or research, in the fields of electronics, radio, allied branches of engineering or the related arts and sciences, in a school of recognized standing, may be accepted as equivalent to professional experience.
10. The time requirements for admission or transfer to any grade of membership may be satisfied by applying *pro rata* the experience of the candidate under the various alternative requirements for the grade.

204. *Applications*

1. Applicants for membership shall furnish names of references as follows:
   For Senior Member—Five Fellows or Senior Members.
   For Member—Four Fellows, Senior Members or Members.
   For Associate—Three Fellows, Senior Members, Members, Associates or other responsible individuals.
   For Student—The IRE Representative at his school, if such a Representative exists; otherwise, a member of the faculty of his school.
2. When an applicant for Senior Member or Member shows that his work or location is such as to make impracticable compliance with the reference requirements of these Bylaws, the names of persons who are familiar with his professional work may be substituted. The Admissions Committee shall exercise discretion in determining suitability of substitute references.
3. Applications for admission or transfer to any grade of membership, except Fellow, shall be addressed to the Board of Directors and submitted to IRE Headquarters. Recommendation of election or transfer of an applicant to any grade, except Fellow, shall be by two-thirds affirmative vote of the Admissions Committee.
4. A re-application for admission or transfer to a particular grade may be made after the expiration of one year from the date of a rejection.
5. The Executive Committee is authorized to act for the Board of Directors in electing members, re-admitting members or transferring their membership grades, except where involving Fellow grade, and may waive application and reference requirements in the transfer of members from Student to Associate or Member grade.

205. *Proposals*

1. Each year, the Fellow Committee shall recommend to the Board of Directors, nominees for Fellow grade not more than seventy-five in number. A citation summarizing the accomplishments of the nominee shall be a part of each recommendation.
2. Admission or transfer to any grade except Fellow may be proposed by any member acting as sponsor by supplying to the Admissions Committee sufficient information and testimonials from the required number of references to satisfy the Admissions Committee as to qualifications. Such proposals shall be acted upon by the Admissions Committee and if approved, transmitted to the Executive Committee. If approved by the Executive Committee, an invitation blank shall be sent to the proposed member inviting him to accept the grade of membership proposed, which membership shall become effective automatically and immediately upon his supplying the biographical and professional information required and paying the necessary dues and fees. The name of an invitee shall be placed on the mailing list for the PROCEEDINGS OF THE IRE immediately upon receipt of dues and fees.

206. *Admissions and Transfers*

1. The entrance fee for all grades shall be three dollars ($3) except that there shall be no entrance fee for the Student grade.
2. No entrance fee shall be charged any non-IRE member of a Joint Student Branch or Joint Student Associate Branch, who is a student member of another national engineering society which is a sponsor with IRE of the Joint Student Branch or Joint Student Associate Branch, and who applies for admission to membership in the IRE within six months after the termination of his student status.
3. There shall be no transfer fees. A notice of approval of a transfer shall be sent to the member. A member whose grade of membership is changed during his annual dues period shall be billed *pro rata* for the remainder of that period.
4. The name of every person newly-admitted or transferred to the grade of Senior Member, Member or Associate shall be published in the PROCEEDINGS OF THE IRE.
5. When an applicant for membership is elected, the membership period shall be dated as of the first day of the first month following receipt of application, if received on or before the nineteenth day of the month, or as of the first day of the second month following receipt of application, if received after the nineteenth day of the month.
6. The entrance fee and dues are payable on notification of election and if not received within six months from notification, the election shall be considered void.
7. A notice that he has been elected shall be sent to every newly-admitted member together with a bill for entrance fee and dues, if not previously paid, dues being computed for the annual period concurrent with the membership period. Entrance fee or dues remaining unpaid, additional bills shall be sent the newly-elected member two months and four months after notification of election, in the last instance accompanied by a warning that the election will be considered void if the entrance fee and dues are not received within six months of notification of election.

207. *Dues*

1. Every member's annual dues period and the period during which he shall receive at least one IRE publication shall run concurrently with his membership period. The annual dues shall be payable in advance at the beginning of the annual dues period.
2. For members elected prior to July 1, 1955, the annual dues for all grades shall be fifteen dollars ($15) except for Student which shall be five dollars ($5) and

except for Associate which shall be ten dollars ($10) for each year which is within the first five years of membership in any grade other than Student. Beginning with the sixth such year, the annual Associate dues shall be fifteen dollars ($15).

3. For members elected July 1, 1955 or thereafter, the annual dues shall be ten dollars ($10) for the first three years of IRE membership in any grade other than Student, and fifteen dollars ($15) for each succeeding year. The annual dues for Student grade shall be five dollars ($5).
4. Dues shall be waived for all IRE members who have both attained the age of sixty-five (65) years and who have been IRE members for at least thirty-five (35) years.
5. Under exceptional circumstances, such as inability of a member to remit dues due to wartime conditions, the Board of Directors may, if it waives dues, also declare that during the period of such waiver, the member has maintained continuous membership.

208. *Billing, Termination, and Reinstatement*

1. A bill covering his dues for the following year shall be mailed to every member not later than one month prior to the beginning of his annual dues period. A second bill shall be mailed after a lapse of two months to every member whose dues remain unpaid. There shall be mailed with the second bill a notice that, if the member's dues remain unpaid for four months, it shall be the duty of the Executive Secretary, in accordance with the Bylaws of the IRE, to notify the member that his membership has terminated.
2. After a member has been in arrears four months, his membership shall terminate and his name shall be removed from the roll of membership. Every such person shall be mailed a notice to the effect that, according to the Bylaws of the IRE, his membership in the IRE has in fact terminated. Accompanying the notice of termination shall be a final bill, with the suggestion that the former member pay the bill and resume his membership.
3. Membership so terminated may be resumed on payment of all dues in arrears, or on payment of a new entrance fee and current dues.
4. On resuming membership and paying dues in arrears, a member may receive available copies of the PROCEEDINGS OF THE IRE during the period covered by the back dues. A rebate of 25 cents per copy shall be made in lieu of copies not available.
5. The mailing of bills or statements to the last known address of a member shall be considered a valid notice of indebtedness.

209. *Resignations*

1. A member in good standing may resign by submitting a written resignation to the Secretary.
2. Subject to the approval of the Executive Committee, a resigned member may resume his membership upon payment of current dues.

210. *Severances*

1. To initiate action toward expulsion of a member, a written complaint must be submitted to the Board of Directors which, if it deems the reason sufficient, shall notify the accused by letter of the charges against him and of the place and date of a hearing to consider such charges which shall be held not less than thirty days after mailing of the notice. The accused may present his defense in person, in writing, or by an authorized representative. There shall be a majority of the members of the Board present at the hearing and the votes cast must be unanimous in order to expel. The action of the Board of Directors shall be final and conclusive.

300. GROUPS OF MEMBERS

301. *Regions*

1. The United States shall be divided into seven geographical Regions which shall be given consecutive numbers for identification. There shall be an additional Region which shall consist of all of Canada.
2. When new Regions are established, or when changes are made in Regional boundaries, the changes shall be made effective so that no Regional Delegate-Regional Director shall have his term shortened by such changes.
3. Each Region shall have a Regional Committee which shall consist of at least the Regional Director, the junior past Regional Director, the Chairman of the Regional Education Committee and two representatives of each Section in the Region, one of these to be the Chairman of the Section and the other to be appointed by the Section's Executive Committee from among a group comprised of the junior past Chairman and the present elected officers of the Section. Each member of the Regional Committee shall be, ex-officio, a member of the Executive Committee of his own Section. In case of necessity, a Section may be represented at a Regional Committee meeting by alternates appointed by the Section Executive Committee.
4. Each Region shall establish a Regional Education Committee to be responsible to the Regional Committee. It shall be concerned with the handling of all educational problems of the Region including Student Member and Student Branch operations, general supervision of student papers contests and similar activities when held on a Regional basis and the fostering and coordination of the educational activities of the Sections. The Regional Education Committee shall be composed of all IRE Representatives in colleges and technical institutes in the Region and all Education Committee members residing in the Region, ex-officio. The Chairman of this Committee is to be appointed from the Region membership by the Regional Director; his term of office shall correspond with that of the Regional Director; he shall be a member of the Education Committee, ex-officio, and a member of the Regional Committee, ex-officio.
5. Pending installation of the first Regional Director of a Region, the Chairman of the largest Section, numerically, in the Region shall act as Chairman pro-tem of the Regional Committee, and the Chairman pro-tem shall appoint a Secretary-Treasurer pro-tem.
6. The Regional Committee may include additional voting members not exceeding in number one-fourth of the above mandatory number of members representing the Sections in the Region. The purpose for such additional members shall be determined by the Regional Committee based upon each Region's individual requirements. Such additional members shall be appointed on a year-to-year basis by the Regional Director, subject to approval of the Regional Committee.
7. Each member of the Regional Committee shall be of Member, Senior Member or Fellow grade in the IRE. The term of office of each member other than the Regional Director and the junior past Regional

Director shall be one year concurrent with the term of office of the Section officers. Vacancies, except for Regional Director, shall be filled by appointment of the Section Executive Committee in which the vacancy exists.

8. The Regional Director shall be Chairman of the Regional Committee and shall appoint one or more Vice Chairmen from among the membership of the Committee. The Chairman of the Regional Committee shall appoint a Secretary-Treasurer from among the membership of the Region for a two-year term concurrent with that of the Chairman. The duties of the Secretary-Treasurer shall include correspondence, the keeping of the minutes of the Committee meetings, mailing notices, handling of funds, the keeping of financial records, the submission of a report at the end of each year and such other duties as are assigned to him by the Chairman. The Secretary-Treasurer shall not be a voting member of the Committee unless he shall be appointed a voting member for a term of one year by the Regional Director, subject to the approval of the Regional Committee.
9. The Regional Committee shall hold at least one meeting each year, within or without the Region. A quorum for a Regional Committee meeting shall consist of at least one-third of the members of the committee or their alternates and shall include representatives from at least half of the Sections in the Region. The Regional Director or, in the event of his unavailability, a Regional Committee Vice Chairman, shall be responsible for calling the necessary number of Regional Committee meetings.
10. The duties of the Regional Committee shall include the making of at least one nomination for Regional Delegate-Regional Director from its Region during election years, participation in the development of the IRE by means of suggestions on national and international matters to the Board of Directors and on Regional and Sectional matters to the Sections comprising the Region.
11. Regions failing to maintain reasonable activity may be dissolved and the Sections therein may be absorbed into other Regions the boundaries of which are to be correspondingly altered.

302. *Sections*

1. There shall be established throughout the territory in which IRE conducts operations groups of members residing in specified geographical districts designated as Sections.
2. A petition for the formation of a Section within the IRE Regional organization shall be signed by not fewer than twenty-five members other than Students, having mailing addresses within the territorial limits proposed in the petition. The territorial limits shall be specifically delineated.
3. A petition for the formation of a Section outside of the IRE Regional organization shall be signed by not fewer than ten members, other than Students, having mailing addresses within the territorial limits proposed in the petition. The territorial limits shall be specifically delineated.
4. The Executive Committee, upon receipt of a petition for the creation of a Section, may authorize its formation. Each Section so authorized shall within not more than six months thereafter adopt and conform to a Section Constitution approved by the Executive Committee, which Section Constitution shall provide for a Chairman, Secretary-Treasurer and a Section Executive Committee. After the Section is organized and a Section Constitution has been adopted, the Executive Committee may give final approval to its establishment.
5. Each Section shall conduct its activities within the Constitution and Bylaws of the IRE and other rules which by law affect the membership and activities of the IRE.
6. Sections outside the IRE Regional organization shall be represented on the Board of Directors by the senior past President in his capacity as a Director.
7. No Section or any officer or representative thereof shall have any authority to contract debts for, pledge the credit of, or in any way bind, the IRE.
8. All Sections shall be exclusively those of the IRE. No Section joint with any other organization, society, or group shall be recognized. However, IRE Sections may cooperate with other organizations in the holding of joint meetings and may invite members of such organizations and the public to their meetings.
9. Section Secretaries shall forward to the Executive Secretary a report of each meeting held by the Section and its Subsections, if any, for the presentation or discussion of papers and at the end of each fiscal year, a financial statement for that fiscal year.
10. Failure of a Section to maintain the required activities, which shall include the holding of at least five meetings each year and also the maintenance of a minimum membership of twenty-five members, other than Students in a Section within the IRE Regional organization, and a minimum membership of ten members, other than Students, in a Section outside of the IRE Regional organization, shall place the Section on probation. All members of the Section shall be informed of the probation by the Executive Secretary who shall also call their attention to the requirements for maintaining the Section. If the delinquency continues for a second year, a second notification to the Section membership shall be made by the Executive Secretary, and the Board of Directors shall be informed of the probationary status of the Section. If the delinquency continues for a third year, the Section shall, thereupon, automatically be dissolved. The Executive Secretary shall so report to the Board of Directors and so inform the members previously constituting that Section. The Board of Directors may waive the provisions herein regarding dissolution of a Section in case such Section becomes affected by war or any other force majeure, for the period of the duration thereof and for at least six months thereafter.
11. For the maintenance of a Section the IRE shall pay to the Section for each calendar year:
    a. One dollar and ten cents ($1.10) for each member, except Students, up to a total of seven hundred members plus one dollar and twenty-five cents ($1.25) for each member, except Students, in excess of a total of seven hundred and;
    b. Fifty cents ($.50) for each Affiliate of each and every Professional Group having a mailing address within the territory of the Section as of December 31 of the calendar year for which payment is made and;
    c. Ten dollars ($10) per meeting for not more than ten meetings of the Section within the calendar year.
12. Local noncompulsory financial contributions may be accepted by Sections.
13. The Board of Directors may dissolve any Section for any reason deemed sufficient by the Board of Directors.
14. Any Section that publishes a periodical shall inform the Executive Committee of the establishment there-

of and shall provide IRE Headquarters with copies thereof, as published.

303. *Subsections*

1. In conformance with the Constitution for Sections, a Section may establish Subsections which shall be geographical subdivisions of the Section, subject to the same limitations and control as the Section itself, pursuant to the following conditions:
   a. In the case of Subsections to be formed in Sections within the IRE Regional organization, a petition shall be submitted to the parent Section Executive Committee signed by twenty or more members, except Students, with mailing addresses within the proposed territory of the Subsection;
   b. In the case of Subsections to be formed in Sections outside the IRE Regional organization, a petition shall be submitted to the parent Section Executive Committee signed by eight or more members, except Students, with mailing addresses within the proposed territory of the Subsection;
   c. Approval of the petition by the parent Section Executive Committee;
   d. Notification of the Executive Secretary of actions by a Section regarding Subsections with a copy of petition or petitions;
   e. Approval by the Executive Committee.
2. For each authorized Subsection, the IRE shall pay to the Section a meeting allowance of ten dollars ($10) per meeting up to a maximum of ten meetings per year, during the time the Subsection remains in good standing according to the Section Constitution.

304. *Professional Groups*

1. There shall be established within the scope and objectives of the IRE, groups of members interested in specific technical fields or related subjects, designated as Professional Groups.
2. A petition for the formation of a Professional Group shall be signed by not fewer than twenty-five members, other than Students, shall state the proposed field of interest and shall be forwarded to the Executive Committee after review thereof by the Professional Groups Committee.
3. The Executive Committee, upon receipt of a petition to form a Professional Group, may authorize its formation. After the Group is organized and a Constitution for it is approved by the Executive Committee, the Executive Committee may give final approval to its establishment.
4. Each Professional Group authorized by the Executive Committee shall within not more than six months thereafter adopt and conform to a Professional Group Constitution which shall have been submitted to and approved by the Executive Committee. A Group Constitution may be amended by the Group subject to the approval of the Executive Committee. Such Group Constitution shall provide for a Group Chairman and a Secretary-Treasurer.
5. All Professional Groups shall consist exclusively of IRE members and such affiliates as may be recognized by the Executive Committee. No Professional Group joint with any other organization, society or group outside of the IRE shall be recognized. However, Professional Groups may co-operate with other organizations in the holding of joint meetings and may invite members of such organizations and the public to their meetings.
6. A Professional Group may not charge dues. A Professional Group may raise revenues by fees or other means approved by the Executive Committee.
7. For the maintenance of a Professional Group, the IRE shall pay a subsidy equal to one-third of the cost to the Group of printing and mailing all TRANSACTIONS, Newsletters and any other publications which are distributed to all paid members of a Group free-of-charge; and in addition, there shall be provided for each new Professional Group, matching funds of one dollar ($1) for each dollar of income derived by the Group from membership fees during the first two years.
8. All IRE members may become members of Professional Groups in conformance with the respective Group Constitutions.
9. All Professional Group meetings shall be open to all members of the IRE on an equal basis with Group members. Any IRE member or any Group Affiliate, upon payment of charges judged to be equitable by the Executive Committee, shall receive any notices of meetings of any Group.
10. A Professional Group may hold or join in holding conferences or conventions and may charge for registration; but a Group may not charge for registration at a meeting, conference or convention when it operates as part of an IRE Sectional, Regional or national meeting, conference or convention.
11. All publications of Professional Groups, other than programs, notices, and the like, shall be subject to prior authorization of the Executive Committee and shall be made available to all IRE members and Group Affiliates on equitable bases approved by the Executive Committee.
12. Each Professional Group shall forward to the Executive Secretary a copy of each meeting notice of the Group, an annual report and a financial statement for each fiscal year prepared immediately following the end thereof.
13. No Professional Group or any officer or representative thereof shall have any authority to contract debts for, pledge the credit of, or in any way bind, the IRE.
14. Each Professional Group shall conduct its activities within the Constitution, Bylaws and other rules which by law affect the membership and activities of the IRE.
15. The Board of Directors may dissolve any Professional Group for any reason deemed sufficient by the Board of Directors.
16. Group Affiliates are non-IRE members who have been admitted by a Professional Group to some of the rights and privileges of Group membership. In general, a Group Affiliate must be a person whose professional life does not justify full membership in the IRE, but which may be helped by participation in Group activities or who may contribute benefits to Group members by his participation in their activities. The qualifications for Group Affiliation and the limits in rights and privileges shall be established by the Executive Committee.

305. *Professional Group Chapters*

1. A Professional Group Chapter may be established in a Section to function in the manner of a committee of a Section.
2. A petition to establish a Chapter must contain the following:
   a. Name of the Section.
   b. Name of the Professional Group.
   c. Name of the organizer (who becomes interim Chairman pending election of a regular Chairman at a later organization meeting).
   d. Signatures of at least ten IRE members, other than Students, who must indicate they are either

members of the Professional Group involved, or are prepared to become members if the petition is granted.

3. The petition for a Chapter shall be submitted to the Section Executive Committee for written approval and forwarded with this written approval to the Executive Secretary. The Executive Secretary shall endeavor to obtain written approval from the Executive Committee. Upon receiving approval, the existence of the new Chapter will be recorded at IRE Headquarters and the Section Executive Committee will be informed thereof.
4. For each Chapter, the IRE shall pay ten dollars ($10) to a Section for each meeting of the Section promoted by a Chapter up to five meetings per Chapter per year, provided such meeting is attended by ten or more Section members and a report of such meeting, including a statement of the number of members attending, signed by a Chapter officer, is is forwarded to IRE Headquarters through the Section Secretary-Treasurer. A meeting of the Section promoted by a Chapter may be counted among the above five, or may be counted among the ten meetings of the Section as provided in these Bylaws, but not among both as such would result in duplicate payments for the same meeting.

306. *Student Branches*

1. There shall be established throughout the territory in which IRE conducts operations, groups of Student members residing in specified geographical localities, designated as Student Branches.
2. There shall be established IRE Representatives who shall be IRE members and teachers of electronics, radio, allied branches of engineering or the related arts and sciences, at schools or technical institutes eligible for IRE Student members, and who shall be appointed and released by the Executive Committee. Each such IRE Representative is charged with promoting the welfare of the IRE at his his school, particularly in matters relating to the Student membership.
3. Upon receipt of a petition signed by the IRE Representative, or in the event one has not been appointed, by a representative of the faculty of the school involved, and by fifteen or more IRE members who are students in a school of recognized standing, the Executive Committee may authorize the establishment of an IRE Student Branch in that school.
4. Upon receipt of a petition signed by the IRE Representative, or in the event one has not been appointed, by a representative of the faculty of the institute involved, and by fifteen or more IRE members who are students in a technical institute approved by the Executive Committee, the Executive Committee may authorize the establishment of an IRE Student Associate Branch in that institute.
5. Each IRE Student Branch and each IRE Student Associate Branch shall adopt and conform to a Student Branch Constitution which shall have been submitted to and approved by the Executive Committee and which shall provide for a Chairman and a Corresponding Secretary.
6. The IRE, by action of the Executive Committee, may join with another national engineering or technical society to co-sponsor a Joint Student Branch in a school of recognized standing or a Joint Student Associate Branch in a technical institute approved by the Executive Committee. The petition for the establishment of the Branch must be signed by the IRE Representative, or in the event one has not been appointed, by a representative of the faculty of the school or institute involved, and by ten or more Student members of either society, at least twenty-five per cent of which petitioners must be IRE Student members.
7. Each Joint Student Branch and each Joint Student Associate Branch shall adopt and conform to a Joint Student Branch Constitution which shall have been submitted to and approved by, both the Executive Committee and the other co-sponsoring society and which shall provide for a Chairman and a Corresponding Secretary.
8. No IRE Student Branch, IRE Student Associate Branch, Joint Student Branch or Joint Student Associate Branch, or any officer or representative thereof shall have any authority to contract debts for, pledge the credit of, or in any way bind the IRE.
9. Each IRE Student Branch, IRE Student Associate Branch, Joint Student Branch and Joint Student Associate Branch Corresponding Secretary shall forward to the Executive Secretary a report of each meeting held by the Branch for the presentation or discussion of papers, and during June of each year, a financial statement for the preceding year.
10. An IRE Student Branch, IRE Student Associate Branch, Joint Student Branch or Joint Student Associate Branch may collect dues from its members as provided in its constitution and, in addition, may accept local non-compulsory financial contributions.
11. Failure of an IRE Student Branch or IRE Student Associate Branch to maintain a minimum of ten IRE Student members and to hold at least three meetings each year shall place the Branch on probation. All officers of the Branch and the IRE Representative, or in the event one has not been appointed, a representative of the faculty of the school or institute involved, shall be informed of the probation by the Executive Secretary who shall also call to their attention the requirements for maintaining the Branch. If the delinquency continues for a second year, the IRE sponsorship shall be withdrawn automatically. The Executive Secretary shall so report to the Executive Committee and the Board of Directors and so inform the former officers of the dissolved Branch and the IRE Representative.
12. Failure of a Joint Student Branch or Joint Student Associate Branch to maintain a minimum of ten Student members, of which at least twenty-five per cent must be IRE Student members, and to hold at least three meetings each year shall be reported by the Executive Secretary to the Executive Committee.
13. For maintenance of an IRE or Joint Student Branch upon certification by the IRE Representative or, in the event one has not been appointed, by a representative of the faculty of the school where the Branch is located, the IRE shall pay to each such Branch upon request the sum of twenty-five dollars ($25) for any one school year; and shall also pay to each such Branch the sum of one dollar ($1) for each IRE Student member of the Branch as of November fifteenth of each year and the sum of fifty cents ($.50) for each additional IRE Student member of the Branch as of the following February first.
14. For maintenance of an IRE or Joint Student Associate Branch, upon certification by the IRE Representative or, in the event one has not been appointed, by a representative of the faculty of the school where the Branch is located, the IRE shall pay to each such Branch upon request the sum of fifteen dollars

($15) for any one school year; and shall also pay to each such Branch the sum of fifty cents ($.50) for each IRE Student member of the Branch as of November fifteenth of each year and the sum of twenty-five cents ($.25) for each additional IRE Student member of the Branch as of the following February first.

15. The Executive Committee may dissolve an IRE Student Branch or IRE Student Associate Branch, or withdraw its sponsorship and support of a Joint Student Branch or Joint Student Associate Branch for any reason deemed sufficient by the Executive Committee.

400. ANNUAL ASSEMBLY

401. *How Constituted*

1. The annual Assembly shall consist of nine Delegates-at-large elected by the membership-at-large and eight Regional Delegates elected respectively by the members residing in the eight Regions of the IRE, making a total of seventeen.
2. The nine Delegates-at-large shall comprise the President, the senior past President, the junior past President, and six additional Delegates-at-large. The six additional Delegates-at-large shall serve three-year terms. Two of the six shall be elected each year so that the terms of two of the six additional Delegates-at-large shall expire each year.
3. The Delegates elected by Regions shall be members of and reside in the Regions electing them and shall have terms of office of two years, the Delegates from even-numbered Regions being chosen and elected in even-numbered and those from odd-numbered Regions in odd-numbered years. A vacancy shall occur in the office of Regional Delegate at such time as the incumbent ceases to reside in the Region that elected him.
4. All Delegates, both Regional Delegates and Delegates-at-large, shall be elected Directors by virtue of their election as Delegates and shall serve terms as Directors concurrent with their terms as Delegates.
5. The presiding officer of the annual Assembly, as provided by the Constitution, shall be the newly-elected President.
6. Prior to the election of the Secretary or in the event of his absence, the Executive Secretary shall perform the secretarial functions.

402. *Meetings*

1. The annual Assembly of the IRE, required to be held during January of each year, shall be held as early in that month as feasible, at IRE Headquarters. Notice of time and place shall be mailed not less than thirty days in advance. In the case of extraordinary circumstances, a special meeting of the Assembly or the designation of a place of meeting for the annual Assembly other than IRE Headquarters may be authorized by a two-thirds vote of those present at a regularly constituted meeting of the Board of Directors held at least thirty days in advance of the date of such specially designated meeting.
2. A meeting of the Assembly may be held without notice if waivers of notice signed by all of the Delegates are filed with the Secretary with notation thereof entered in the minutes of the meeting.
3. Ten Delegates shall constitute a quorum at meetings of the Assembly.
4. Any notices of meetings or other official business required by the Constitution or these Bylaws to be mailed to Delegates shall be sent by airmail if the addresses are located at places beyond the continental United States or Canada, or at any place where ordinary mail might be unusually slow and airmail more expeditious.
5. Reimbursement of necessary travel expenses incurred by Delegates for the purpose of attending any special or extraordinary meeting of the Assembly not associated with a meeting of the Board of Directors held contiguously with it, is authorized.

403. *Functions*

1. The annual Assembly shall elect six additional Directors for the required term of office of one year. Three of these Directors shall also be elected Corporate Officers, respectively designated as the Secretary, the Treasurer and the Editor. The annual Assembly shall also elect such alternate Directors as may be necessary to insure the attaining of a total number of six Directors including the three Corporate Officers, in the event that consent to serve by any of the electees cannot be secured prior to the conclusion of that Assembly meeting.
2. The annual Assembly shall receive reports, verified by the President and Treasurer, or by the Board of Directors, showing the amount and status of real and personal property owned by the IRE, its fiscal condition, changes in membership, changes in groups of members and other vital statistics, all of which are to be as of the most recent date available from the records kept by the Officers. These reports shall be filed with the records of the IRE and abstracts thereof entered in the minutes of the proceedings of the Assembly.

500. MANAGEMENT

501. *Board of Directors*

1. The Board of Directors shall consist of eleven Directors-at-large elected by the membership-at-large, eight Regional Directors elected respectively by the members residing in the eight Regions of the IRE and six Directors elected by the annual Assembly, making a total of twenty-five.
2. The Directors-at-large shall comprise the President, the senior past President, the junior past President, the two Vice Presidents and six Directors elected for three year terms. Two of the six shall be elected each year so that the terms of two of the six Directors shall expire each year.
3. The Directors elected by Regions shall be members of and reside in the Regions electing them and shall have terms of office of two years, the Directors from even-numbered Regions being chosen and elected in even-numbered, and those from odd-numbered Regions in odd-numbered years. A vacancy shall occur in the office of Regional Director at such time as the incumbent ceases to reside in the Region that elected him.
4. The annual meeting of the Board of Directors, required to be held during January of each year, shall be held immediately after the adjournment of the annual Assembly and at the same place. Notices thereof shall be mailed at the same time that notices are mailed for the annual Assembly.
5. At the annual meeting, the Board shall determine the time and place of other meetings to be held throughout the year, which normally shall be scheduled for the first Wednesday of May, the first Wednesday after the second Monday of September and the first Wednesday after the second Monday of November.

6. The time and place of any scheduled Board meeting may be altered or the meeting cancelled only by consent of a majority of all Directors secured by or transmitted to the Secretary, not less than twenty days before the original date or the new date set for the meeting, whichever is the earlier. Notice of such approved change shall be mailed to all Directors not less than ten days before the original or the new date, whichever is the earlier.
7. Special meetings of the Board of Directors may be called by any five Directors on notice to all other Directors. Notice of such special meetings giving the the time and place of meeting, the purpose of the meeting and the names of the Directors calling the meeting shall be mailed to all Directors not less than twenty days before the date set for the special meeting. The place of such special meetings may be only at IRE Headquarters unless otherwise authorized by a majority vote of all Directors, such vote to be secured by or transmitted to the Secretary.
8. In the absence or incapacity of the Secretary, the Executive Secretary shall perform the secretarial functions.
9. A meeting of the Board of Directors may be held without notice if waivers of notice signed by all of the Directors are filed with the Secretary with notation thereof entered in the minutes of the meeting.
10. Notices of meetings and any other documents required to be sent to Directors pursuant to provisions of the Constitution and these Bylaws shall be sent by airmail if the destination is beyond the continental United States or Canada, or at any place where ordinary mail might be unusually slow and airmail more expeditious.
11. Twelve Directors shall constitute a quorum at meetings of the Board of Directors.
12. Reimbursement of necessary travel expenses incurred by Directors for the purpose of attending meetings of the Board of Directors, the Executive Committee and meetings taking place during the IRE National Convention is authorized.
13. The Board of Directors may appoint or direct the appointment of representatives of the IRE on joint committees, boards and other local, national and international bodies.

502. *Executive Committee*

1. The Board of Directors at its annual meeting shall appoint an Executive Committee composed of Officers and Directors, to exercise powers and assume duties of management as directed by the Board of Directors and as may be outlined in the Bylaws, subject to the following limitations:
   a. That the Board of Directors may at any meeting overrule any act or decision of the Executive Committee except insofar as any act has in fact been carried out, or
   b. Suspend at any meeting any power conferred upon the Executive Committee, such suspension to remain in effect pending repeal of any Bylaw conferring such power, or
   c. Direct any action or plan of the Executive Committee.
2. The Executive Committee shall consist of eight members and shall comprise the President, the Vice President residing in North America, the junior past President, the Treasurer, the Secretary, the Editor and two other Directors. The President shall be Chairman, the Vice President shall be Vice Chairman and the Secretary shall be Secretary thereof.
3. At the first meeting of a newly-appointed Executive Committee, it shall schedule other meetings throughout the year. Minutes of the meetings of the Executive Committee shall be mailed to all Directors.
4. The terms of office of members of the Executive Committee shall begin with the annual meeting of the Board of Directors and shall continue until the succeeding annual Assembly.
5. The Executive Committee may divide its duties and responsibilities among its members as it sees fit and conduct its business as it finds necessary. Such activities as the Executive Committee shall direct shall be administered by the Executive Secretary.
6. The Executive Committee shall be responsible for the management of IRE Headquarters. The Executive Committee shall prepare and present an annual budget to the Board of Directors for approval.
7. The Executive Committee shall be charged with broadly considering IRE policies and making appropriate recommendations to the Board of Directors on its own initiative.
8. The Executive Committee shall appoint, direct and coordinate the work of all Standing Administrative Committees except those appointed directly by the Board of Directors.
9. The Executive Committee shall appoint, direct and coordinate the work of all Standing Technical Committees.
10. The Executive Committee shall approve or disapprove the formation, dissolution and operation of groups of members.
11. The Executive Committee may authorize conventions, conferences and meetings of the IRE and approve their dates and locations.
12. The Executive Committee shall direct and manage IRE standardization activities, relations with other organizations, special activities, technical activities, advertising and publications.
13. The Executive Committee shall establish an office account, limited in size, the funds for which shall be made available from the funds received by IRE by authorization of at least two Corporate Officers, as provided in the Constitution. Funds shall be withdrawn from the office account only by authorized bonded employees of the IRE.

503. *Officers*

1. As provided in the Constitution, the Corporate Officers of the IRE shall be a President, two Vice Presidents, a Secretary, Treasurer and an Editor.
2. As provided in the Constitution, the President shall be elected by the voting members by virtue of which election he shall also be a Delegate-at-large and a Director-at-large. His term as President shall be one year but his term as Delegate-Director shall be three years.
3. There shall be two Vice Presidents of the IRE who shall be elected by the voting membership for terms of office of one year, by virtue of which election they shall also be Directors-at-large. One of these shall be a member residing in North America who, by reason of availability and qualifications, is able effectively to assist the President in the performance of his presidential duties, and the other shall be a member residing elsewhere than in North America. Both Vice Presidents shall serve terms as Directors concurrent with their terms as officers.
4. The Vice President residing in North America shall take precedence as regards assumption of presidential duties as provided in the Constitution.
5. The Secretary, Treasurer and Editor shall be elected by the annual Assembly by virtue of which election

they shall also be Directors. The terms of the Secretary, Treasurer and Editor shall be one year and their terms as Director shall also be one year.

6. Reimbursement of necessary travel expenses of the President, incurred for the performance of his IRE duties other than his duties as Delegate or Director, is authorized. Reimbursement of necessary travel expenses of the Vice President residing in North America, incurred for the performance of his IRE duties other than his duties as Director, is authorized.
7. Incapacity of the President to perform his duties and the length of time such incapacity may continue shall be determined by the Board of Directors, or by the Executive Committee, if the incapacity occurs at a time when it is not feasible to convene a meeting of the Board and there is need for action. The President, however, may declare a condition of incapacity and the duration thereof by written communication to the Secretary, in which event the Secretary shall notify the Vice President residing in North America, to assume the presidential duties and the Secretary shall take such other action as provided in the Constitution if that Vice President is unavailable or incapacitated.
8. A vacancy among the Corporate Officers and the duration thereof shall be determined by the Board of Directors which may fill such vacancy as provided by the Constitution. If a vacancy occurs among the Corporate Officers at a time when it is not feasible to convene a meeting of the Board of Directors and there is need for action, the Executive Committee shall make such determination and fill such vacancy if the approval of a majority of all the Directors is secured by mail or telephone, such approval to be recorded by the Secretary with the names of the approving Directors.
9. The President shall determine the extent of the assistance to be rendered him by the Vice President charged with the duty of assisting the President.
10. The Secretary shall prepare an Annual Report on membership, meetings and IRE activities, to be submitted to the Board of Directors.
11. The Treasurer shall cause to be prepared an annual audit of the affairs of the IRE by certified public accountants and a report thereof submitted to the Board of Directors. The Treasurer shall also make a report to the Executive Committee semi-annually concerning the status of the IRE investments. The Treasurer, before making changes in the IRE portfolios, shall report the proposed changes to the Executive Committee and obtain its approval.

504. *Committees*

1. The IRE shall have Standing Committees each of which shall normally consist of five or more persons. These shall be classified as Standing Administrative Committees and Standing Technical Committees. Each Standing Committee shall have the right to create subcommittees of its own selection.
2. The Standing Administrative Committee shall be:

Admissions	History
Awards	Nominations and Appointments
Constitution and Laws	Policy Study
Editorial Board	Professional Groups
Education	Tellers
Fellow	
Finance	

The Awards, Constitution and Laws, Fellow, Nominations and Appointments, Policy Study and Tellers Committees and the Editorial Board shall be appointed by and be advisory to the Board of Directors. The others shall be appointed by and be advisory to the Executive Committee. The appointments to these committees, except the Nominations and Appointments Committee, shall start with the first day following appointment and shall continue until the date the succeeding terms of appointments take effect.

3. The Standing Technical Committees shall be:
Antennas and Waveguides
Audio and Electroacoustics
Circuits
Electron Tubes
Electronic Computers
Facsimile
Feedback Control Systems
Industrial Electronics
Information Theory and Modulation Systems
Measurements and Instrumentation
Medical Electronics
Mobile Communication Systems
Navigation Aids
Nuclear Techniques
Piezoelectric and Ferroelectric Crystals
Radio Frequency Interference
Radio Transmitters
Radio Receivers
Recording and Reproducing
Solid State Devices
Standards
Symbols
Television Systems
Video Techniques
Wave Propagation
They shall be appointed by and be advisory to the Executive Committee; the appointments to them shall be made between January first and May first, and the terms of appointment shall be from May first of the year when the appointments are made through April thirtieth of the following year.
4. Additional appointments may be made to any standing committee to fill vacancies or to care for special cases as the need arises, but such appointments shall expire with the expiration of the term of the committee.
5. The functions and responsibilities of each Committee, together with rules for operation and guidance, shall be compiled by the Executive Secretary under the direction of the Executive Committee and contained in a document titled ". . . Committee Manual." Following appointment, each new member of a Committee shall be provided with a copy of the applicable Manual.
6. The Admissions Committee shall be responsible for advising the Executive Committee as to the admission and transfer of applicants to the professional grades of Senior Member and Member. It shall be concerned with determining whether the applicants meet the requirements of the Constitution and the Bylaws, functioning primarily as a jury and acting as a guardian of the membership standards of the IRE. As far as possible, its meetings shall be held just prior to the scheduled meetings of the Executive Committee in the interest of expediting the processing of applications. Members of the Admissions Committee shall be of Fellow or Senior Member grade.

   Applications for the grades of Associate and Student shall be handled by IRE Headquarters and periodic reports shall be made to the Executive Committee by the Executive Secretary.
7. The Awards Committee shall be responsible for making recommendations to the Board of Directors for the following IRE National awards:

a. Medal of Honor
b. Morris Liebmann Memorial Prize Award
c. Browder J. Thompson Memorial Prize Award
d. Harry Diamond Memorial Award
e. Vladimir K. Zworykin Television Price Award
f. Founders Award
g. W.R.G. Baker Award

The specifications for all of these awards shall be set forth in the Awards Committee Manual.

The Awards Committee shall also be responsible, in odd-numbered years, for making a recommendation to the Board of Directors on the Norman W. V. Hayes Memorial Award of The Institution of Radio Engineers, Australia.

The Board of Directors may assign other matters concerning IRE awards to the Committee.

The Awards Committee shall consist of twelve members chosen from among those of Fellow grade, eight being selected so that each of the eight Regions will have representation. When appointments are being make, retiring members of the Fellow Committee shall become nominees for consideration for such appointments. Six of the twelve members shall be appointed each year for terms of two years each. A Chairman shall be appointed from among the twelve members to serve for a term of one year.

8. The Constitution and Laws Committee shall be responsible for advising the Board of Directors and assisting the Executive Committee, on matters concerning changes and amendments to the Constitution or the Bylaws that are referred to it.

9. The Editorial Board shall be advisory to the Board of Directors concerning all matters of editorial policy, the publication of the PROCEEDINGS OF THE IRE, including policy determination of its editorial and technical content, and general publication policies for IRE publications, including the IRE DIRECTORY, the TRANSACTIONS of the Professional Groups, the IRE STUDENT QUARTERLY and the IRE NATIONAL CONVENTION RECORD.

The Editorial Board shall consist of the Editor as Chairman, a Vice Chairman, the Managing Editor if such position exists, and not less than three nor more than five other members.

10. The Education Committee shall be a policy advisory group for the IRE in the field of education and shall report to the Executive Committee at periodic intervals concerning the status of and proposals affecting that field. These reports shall be simultaneously circulated to the members of the Board of Directors.

The Committee shall study the needs of electronic engineering education in all areas, including the high schools, the universities and the technical institutes, determine whether the general area of electronics and basic electrical engineering is adequately covered in the colleges of the nation and consider a possible policy for accreditation if deficiencies are found.

The Committee shall consider the needs of the profession in graduate education, and shall guide IRE policy in recognizing and promoting graduate work in electronics. Means whereby the IRE may improve the educational status of its members in general shall also be suggested.

The Committee shall closely follow the use of electronics in education and shall make recommendations concerning means of promoting electronics as a teaching and learning medium in the schools and universities. The Committee shall consider and make recommendations concerning IRE policies and activities for the Student members and Student Branches.

The Committee shall establish appropriate standing subcommittees, such as Student Branch, Graduate Education, Accreditation and for other areas where appropriate to carry out its responsibilities more adequately.

The Committee shall seek contributions to the formulation of policy recommendations from the Professional Group on Education, encouraging that Group to provide for exchange of information through seminars and publications and for discussion of educational philosophy, problems and methods.

The membership of the Education Committee shall include, ex-officio, the Chairmen of the Regional Education Committees and members-at-large not to exceed the number of ex-officio members.

11. The Fellow Committee shall be responsible for making recommendations to the Board of Directors for nominees to be conferred the grade of Fellow. The Committee shall also provide a citation for each such nominee. The recommendations may not exceed a total of seventy-five in any one year, approximately five thereof to be allocated for nominees residing elsewhere than in North America. Sources from which nominations of candidates and other useful information shall be obtained shall be set forth in the Fellow Committee Manual.

The Fellow Committee shall consist of twelve members chosen from among those of Fellow grade, eight being selected so that each of the eight Regions will have representation. Six of the twelve members shall be appointed each year for terms of two years each. A Chairman shall be appointed from among the twelve members to serve for a term of one year.

12. The Finance Committee shall consist of not more than five members of the Board of Directors and shall include the Treasurer, ex-officio. The Chairman of the Finance Committee shall be chosen from among the members of the Executive Committee.

The Committee shall be responsible for periodically reviewing the fiscal affairs of the IRE concerning which it is to take the initiative for making recommendations to the Executive Committee when warranted. It shall review the annual budget in advance of its submission to the Executive Committee and the Board of Directors.

13. The History Committee shall be responsible for considering questions concerning historical matters in the fields covered by IRE activities, including assistance to the Editorial Board, assisting institutions of a public nature such as the Smithsonian Institution where helpful information is requested and can be secured, and for providing information and recommendations to the Executive Committee when pertinent.

14. The Nominations and Appointments Committee shall be appointed at the last meeting of the Board of Directors held prior to December tenth. Its membership shall be chosen from among those members of the Board of Directors whose terms continue through the following year and shall include two of the new Directors-elect and the President-elect.

At least fourteen days before the next annual Assembly, the Nominations and Appointments Committee shall mail to the Delegates whose terms continue through the following year and to the new Delegates-elect, a list of candidates for the offices of Secretary, Treasurer, Editor and Directors to be elected by the annual Assembly.

A list of candidates to fill vacancies that will exist among the chairmen and members of the Awards, Constitution and Laws, Executive, Fellow, Policy

Study and Tellers Committees and the Editorial Board shall be sent to all of the Directors whose terms continue through the following year, to the new Directors-elect and to those candidates recommended by the Nominations and Appointments Committee for election as Directors by the annual Assembly.

The Nominations and Appointments Committee shall also, subsequent to the termination of the fiscal year in which it was appointed, carry out the other provisions of these Bylaws. In addition, the Nominations and Appointments Committee shall recommend to the Board of Directors nominees for the Nominations and Appointments Committee to be appointed later in that year.

15. The Policy Study Committee shall be responsible for making studies and recommendations as the basis for the possible establishment of policies by the Board of Directors or for the taking of action by the Board of Directors for the good of the IRE. Such problems for study shall be referred to the Committee by the Board of Directors.

    The Policy Study Committee shall consist of the most recent past President, two Regional Directors serving the second year of their two-year terms and two Directors-at-large serving the second or third years of their three-year terms. In the event a Director is unable to serve out a term, the Board of Directors may appoint another Director who is serving the second or third year of his term. The Chairman of the Committee shall be the junior past President. Should the junior past President be unable to serve as Chairman, the Board of Directors shall appoint another member of the Committee as Chairman.
16. The Professional Groups Committee shall administer the formation and discontinuance of Groups. It shall also recommend to the Executive Committee the establishment or dissolution of a Group, provide a uniform Constitution, approve Bylaws for Groups involving specialized matters, provide a Manual of Instructions and develop the Group system to its fullest capacity consistent with a healthy growth, all within the limits of whatever framework the Executive Committee shall establish for the conduct of Groups.

    The Professional Groups Committee shall consist of the Chairman of each Professional Group, ex-officio, and members-at-large, the number of these to be one-half the number of Professional Group Chairmen.
17. The Tellers Committee shall be responsible to the Board of Directors for the supervision of the counting of ballots required to be obtained by the Constitution on questions submitted to the membership of IRE or to groups of members. The staff at IRE Headquarters will assist the Committee in the performance of its work. Reports by the Committee to the Board of Directors shall be signed by the Chairman of the Committee.
18. The Standards Committee shall be responsible for the coordination of the standardization activities of the Standing Technical Committees, the initiation of Standards activities, the establishment of the scope of activity of each Technical Committee and the rendering of final approval of all Standards evolved by IRE prior to publication.
19. Scopes of Technical Committee activities, organization and membership, procedures and methods of coordination between them, and with other standardization bodies shall be compiled in a Manual for Technical Committees (including the Standards Committee).
20. The Chairman of each Standing Technical Committee shall be a member of the Standards Committee, ex-officio.

505. *Executive Secretary*

1. The Executive Committee shall appoint an Executive Secretary and shall fix his tenure of office and determine his salary.
2. The Executive Secretary shall be in charge of IRE Headquarters, its staff and operations, for which he shall be responsible to the Executive Committee.

506. *Nominations and Elections*

1. The Secretary shall circularize all Section Chairmen and Secretaries before March first of each year requesting the submission of suitable names to be considered by the Nominations and Appointments Committee for the various elective offices.
2. The Nominations and Appointments Committee shall submit the name or names of one or more candidates for the office of President and for the office of the Vice President residing elsewhere than in North America and the names of more than one candidate for each of the other elective offices.
3. The Nominations and Appointments Committee shall transmit its list of proposed nominees to all Directors at least one week before the date at which the Board of Directors is expected to act upon them. Under normal conditions, this date is the first Wednesday of May.
4. Each Regional Committee shall submit to the Board of Directors by April thirtieth of the Regional election year, the names of at least one qualified nominee for the office of Regional Delegate-Regional Director. To qualify, each candidate shall indicate to the Regional Committee his acceptance of the nomination.
5. Nominations by petition for the office of Regional Delegate-Regional Director may be made by the submission, by the date stated and according to the procedure outlined in the Constitution, of a petition signed by at least one per cent of the total number of voting members residing within the Region, as listed in the official membership records of the IRE at the end of the previous year, but in no case shall the number be less than one hundred.
6. During the period between August fourteenth and September first of each year, the Executive Committee is authorized to take any necessary actions on any petitions and to approve the ballot.
7. Each proposed nominee named by the Board of Directors shall be consulted, and, if he so requests, his name shall be withdrawn.
8. The list of nominees required to be submitted to all voting members by the Constitution shall contain at least one name for the office of President, at least one name for each of the two offices of Vice President, and at least four names for the two offices of Delegate-at-large-Director-at-large. The ballot going to each of the Regions where an election of a Regional Delegate-Regional Director is to take place shall also contain the names of all nominees for the office of Regional Delegate-Regional Director for that particular Region.
9. The ballot required by the Constitution shall contain the number of names for each office to be voted upon, and the names of the nominees for Regional Delegate-Regional Director shall be suitably identified by Region since these will be voted on only by the voting members of each Region involved. The names of the nominees for each office shall be arranged in alphabetical order. The ballots shall carry

a statement to the effect that the order of the names is alphabetical for convenience only and indicates no preference.

10. Voting members shall vote for the candidates whose names appear on the list of nominees, by written ballots in plain sealed envelopes, enclosed within mailing envelopes marked "Ballot" and bearing the member's written signature. No ballots within unsigned outer envelopes shall be counted. No votes by proxy shall be counted. Only ballots arriving at IRE Headquarters before twelve o'clock noon on the last working weekday prior to October twenty-fifth shall be counted. Ballots shall be checked, opened and counted under the supervision of the Tellers Committee between October twenty-fifth and the first Wednesday in November. The result of the count shall be reported to the Board of Directors at its next succeeding meeting and the nominee for each office for which the election is being held, receiving the greatest number of qualified votes shall be declared elected. In the event of a tie vote the Board of Directors shall choose between the nominees involved.

507. *Amendments*

1. Before an amendment to the Constitution is submitted to the voting members, it shall be reviewed by legal counsel designated by the Executive Committee and a written opinion secured that such amendment is in accordance with the laws under which the IRE is incorporated and operated.
2. Ballots for Constitutional amendments shall carry a statement of the time limit for return to IRE Headquarters. The ballots after marking shall be placed in plain sealed envelopes, enclosed within mailing envelopes marked "Ballot" and bearing the member's written signature. Only ballots with signed outer envelopes shall be counted. No votes by proxy shall be counted. Only ballots arriving at IRE Headquarters prior to the stated time limit shall be counted. The votes will be counted under the supervision of the Tellers Committee and the results reported to the Board of Directors at its next meeting.
3. A complete history of amendments shall be kept in the files of the IRE.

508. *Fiscal Year*

The fiscal year of the IRE shall be the calendar year.

600. PUBLICATIONS

601. *Periodicals*

1. The IRE shall publish a periodical named PROCEEDINGS OF THE IRE.
2. The IRE shall publish periodicals of the various Professional Groups named IRE TRANSACTIONS.
3. The IRE shall publish a periodical named IRE STUDENT QUARTERLY.

602. *Standards*

1. The IRE shall publish from time to time, in the PROCEEDINGS OF THE IRE, Standards concerning technologies encompassed within the scope of the objects of the IRE, which have been issued by its Standards Committee.

700. CONVENTIONS

701. *Annual Convention*

1. There shall be held annually in New York City in the month of March or as near thereto as possible, a convention named the IRE NATIONAL CONVENTION. The IRE shall be the sole sponsor.