VOLUME 1 NUMBER 2 FEBRUARY 1964

IEEE spectrum

features

departments

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

IEEE spectrum

JOHN D. RYDER — *Editor*
ELWOOD K. GANNETT — *Managing Editor*
RONALD K. JURGEN — *Executive Editor*
ROBERT E. WHITLOCK — *Technical Editor*
SAMUEL WALTERS — *Technical Editor*
GORDON D. FRIEDLANDER — *Technical Editor*
ELEANOR F. PECK — *Copy Editor*
GLADYS CHARITZ — *IEEE News Editor*
STELLA GRAZDA — *Editorial Assistant*

HERBERT TAYLOR — *Art Director*
RUTH M. EDMISTON — *Production Editor*

EDITORIAL OPERATIONS

HELENE FRISCHAUER — *Administrative Editor*
W. R. CRONE — *Assistant Editor*

CHARLES S. RICH — *Publications Adviser*

ADVERTISING

WILLIAM C. COPP — *Advertising Manager*
LILLIAN PETRANEK — *Assistant Advertising Manager*
CARL MAIER — *Advertising Production Manager*

EDITORIAL BOARD

JOHN D. RYDER — *Chairman*
THOMAS F. JONES, JR. — *Vice Chairman*
E. K. GANNETT
C. T. HATCHER
T. A. HUNTER
S. HERWALD
W. MACADAM
K. F. WILLENBROCK
A. H. WAYNICK

the cover

The cover illustration shows the interior of a model of a basic human cell enlarged 250 000 times. The model was designed by Will Burtin and constructed for The Upjohn Company (Kalamazoo. Mich.). The complete cell model measures six feet across.

This model of a basic cell illustrates strikingly the complexity of the human system—a type of complexity that is involved in human intelligence. When complex humans attempt to design complex machines having artificial intelligences. interesting systems evolve. In Louis Fein's article "The Artificial Intelligentsia" beginning on page 74 of this issue. a comprehensive discussion of some of the most promising of these systems is presented in a novel but technically penetrating manner.

IEEE SPECTRUM is published monthly by The Institute of Electrical and Electronics Engineers, Inc. Headquarters address: Box A, Lenox Hill Station, New York, N.Y. 10021. Cable address: ITRIPLEE. Telephone: 212—PL 2-6800. Published at 20th and Northampton Sts., Easton, Pa. **Change of address** requires thirty days notice to IEEE Headquarters. **Annual subscription:** IEEE members, first subscription included in dues, one additional subscription $13.50. College and public libraries, $13.50 in U.S.A. and Canada, $14.50 elsewhere. Others, $18.00 in U.S.A. and Canada, $19.00 elsewhere. Single copies: members, one additional copy $1.00, college and public libraries, $1.50 in U.S.A. and Canada, $1.75 elsewhere. Others, $2.00 in U.S.A. and Canada, $2.25 elsewhere. **Editorial correspondence** should be addressed to IEEE SPECTRUM at IEEE Headquarters. **Advertising correspondence** should be addressed to IEEE Advertising Department, 72 W. 45 St., New York, N.Y. 10036. Telephone: 212—MU 2-6606.

"Hurrah for the Yankee experiment," cried Faraday when Joseph Henry showed him how to get sparks from a thermocouple circuit by inserting an inductor. This admiration of Henry by Faraday sent Jack Ramsay looking into Henry's achievements in electromagnetism. To his astonishment he found that Joseph Henry was the Faraday of America and just missed fame by delaying in publishing his discoveries.

Joseph Henry and Michael Faraday

It is not perhaps fully realized that Joseph Henry was of the same intellectual stature as Faraday, and that a historical accident prevented his sharing officially the right to be described as a founder of electromagnetism. Fortunately posterity awarded to both men the privilege of associating their names with units of equal basic importance. Yet in parallelling inductance with capacity we are, in fact, parallelling genius.

Joseph Henry was born at Albany, New York, of Scottish parentage, in December, 1799, eight years after the birth of Faraday in England. Both boys were apprentices, Faraday to a bookbinder, Henry to a watchmaker. Each discovered the thrill of scientific knowledge by reading and experimenting. At the age of 22 Faraday became a chemical laboratory assistant at the Royal Institution in London; at 24 Henry was a chemical laboratory assistant at Albany Academy. Faraday published his first paper when he was 25, Henry when he was 24.

The year 1832 was scientifically epochmaking. In April Faraday (aged 41) published his discovery of electromagnetic induction. In July Henry (aged 32) published details of his own independent (apparently prior) discovery, not only of electromagnetic induction, but also of self-induction; this latter phenomenon Faraday did not announce till 1834. The earlier publication of mutual induction gave Faraday the lasting fame. Henry's claim to self-induction, however, was not disputed. By 1832 Henry had also produced an electromechanical machine, line transmission, signaling, the relay, and the transformer. If Faraday made the first electric motor, Henry incorporated the first commutator and the first electromagnet in his magnetoelectric machine.

Henry demonstrated line signaling to his students years before Morse in America and Wheatstone in England started their systems. His use of the relay as a calling and control device led to its universal adoption in telegraphy and later telephony as well as control engineering generally. The invention and use by Henry of the transformer, his understanding of its operation, and of impedance matching, as we now call it, led later to A.C. power transmission, as developed for example by Tesla. Henry, however, was the first transmission line engineer.

Faraday's ideas on light as an electromagnetic wave motion (1846) were anticipated by Henry in 1844. For Henry had in 1842, long before Hertz's experiments in 1887, found that needles in a basement could be magnetized by sparks two floors above. "The diffusion of motion," he said, "is almost comparable with that of a flint and steel in the case of light." In 1851, Henry stated that the discharge of a Leyden jar traveled "wave-fashion," and extended to a surprising distance. (A half-mile or more may have been obtained in 1844.) Henry had discovered that the spark discharge was oscillatory and radiated; its energy could affect "a whole village," like "light from a candle," as he said in suggesting that electrical waves and light waves were probably similar kinds of "wave or vibration of the same medium." Here Henry provided the first intuition, based on experiment, of an electromagnetic theory of light. Yet Maxwell credited the conception to Faraday.

Faraday's achievements in chemistry, electrochemistry and on the action of a magnetic field on polarized light have no exact parallel in Henry's work. Faraday's studies of dielectrics and his discovery of diamagnetism, his researches on discharges in gases—these and other matters gave him the preeminent scientific position he must always hold. But Henry's scientific insight in one of Faraday's fields, and a natural engineering flair, enabled Henry to originate devices and systems, and to obtain results upon which electrical engineering rather than electrical science has developed.

Henry's academic career covered six years at Albany Academy and fourteen years as professor at Princeton. In 1846, however, he became the first Secretary of the Smithsonian Institution (founded by an Englishman, James Smithson), and exchanged the broader field of general science for his earlier specialized research interests. About the same time Faraday, still Professor at the Royal Institution (founded by an American, Count Rumford) was continuing his research on polarization and diamagnetism, and was a researcher throughout his life. Henry, however, built different edifices. The National Academy of Sciences, the American Association for the Advancement of Science, the Washington Philosophical Society, the U. S. Weather Bureau, and much indeed of contemporary American science, all owe a debt to the largehearted, broadminded, gifted man who turned down fame in order to realize Smithson's ideal of "The increase and diffusion of knowledge among men."

JOSEPH HENRY

MICHAEL FARADAY

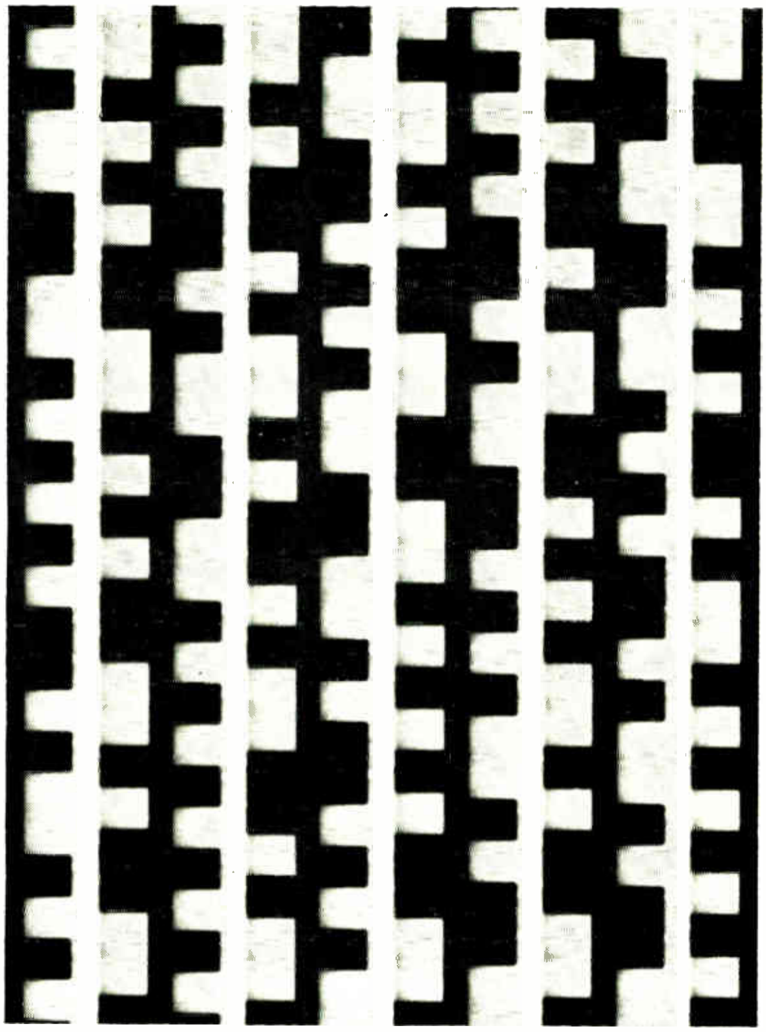
Acknowledgments are made to the Royal Institution, London, and to the Smithsonian Institution, Washington, for the likenesses of Faraday and Henry shown here; these portraits date from that period when both men were at the peak of their electromagnetic powers.

References:

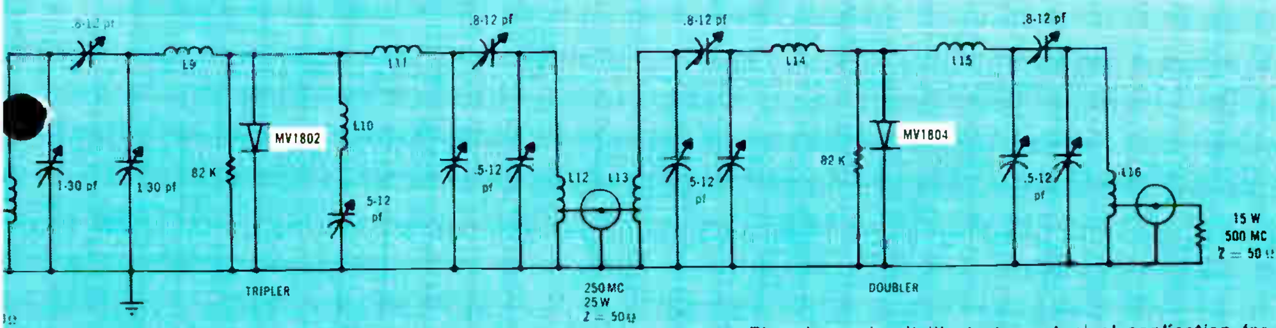
T. Coulson, "Joseph Henry," Princeton U. P., 1950

S. R. Reidman, "The Life of Joseph Henry," Rand McNally & Co., 1961.

The ART of engineering



This greatly enlarged reproduction shows what actually is an extremely small portion of a special computer memory. The memory in physical appearance is a circular band about 3/8 inch wide and 9-3/8 inches in diameter near the outer edge of a circular "photoscopic" disk. Each black or white area (shown reversed in this reproduction from that actually used in the memory) represents an entry in a "dictionary." The dictionary is the heart of an International Business Machines' system for automatic translation of foreign languages. The circular band, representing the dictionary, is capable of holding 60 to 70 million bits of information. In the language translating application, the band is used for the entire dictionary of several hundred thousand entries. Any one of these entries can be found automatically in a fiftieth of a second. Readout is accomplished with a cathode-ray-tube beam and phototube. The translating system was developed for the United States Air Force.



The above circuit illustrates a typical application for Motorola silicon RF transistors and varactor diodes, and is not intended to be a final design.

L9 = 5T #16, 3/8" I.D.
L10 = 8T #16, 1/4" I.D.
L11 = 6T #16, 1/4" I.D.
L12 = 4T #16, 3/8" I.D. 5/8" Tapped 1 T

L13 = 6T #16, 1/8" I.D.
L14 = 5T #16, 1/4" I.D.
L15 = 4T #16, 3/16" I.D.
L16 = 4T #16, 3/16" I.D. Tapped 1 T

POWER/FREQUENCY HORIZONS

'UT AT 500+MC!

nd POWER VARACTORS with efficiencies as high as 75%

Motorola silicon high-power varactor diodes are 100% tested for in-circuit efficiency at specified input power levels . . . designed for meeting high-frequency requirements up to and beyond 500mc where high power outputs are desired.

Look to Motorola for continued leadership in high-power, high-frequency RF silicon devices. These units are immediately available through your Motorola Semiconductor District Office or Distributor . . . available at production prices!

MOTOROLA SILICON HIGH-POWER VARACTORS

		PERFORMANCE SPECIFICATIONS AS TRIPLER			ELECTRICAL SPECIFICATIONS @ 25°C			ABSOLUTE MAXIMUM RATINGS	
Type	Package	Frequency	Input Power	Efficiency (%) min.	Breakdown Voltage (10μA) min	Q min (6 Vdc @ 50 mc)	C_J max (6 Vdc @ 50 mc)	T_J	R.F. P_{in}
MV1802	DO-4	50-150 mc	50 W	65	250 Vdc	75	50 pf	175°C	100 W
MV1804	DO-4	150-450 mc	30 W	50	150 Vdc	150	35 pf	175°C	40 W

Semiconductor Products Inc.

A SUBSIDIARY OF MOTOROLA INC.

Reflections

75 years ago

Multiplex Telegraphy. "I had constructed two ordinary direct current motors of the simplest form, an ordinary two pole field and single coil armature conveniently arranged for use as either direct or alternating current machines. The armature circuits were connected at one end to an ordinary two-part commutator and at the other to two insulated ring contacts. They are shown by a simple diagram in Fig. 1, where both are connected in a single circuit. The armature of the machine M^1 was normally open at the ring contacts, and could be closed either by connecting the ring or through an external circuit. Such a circuit was made to include the field coils of the machine M^2, and the armature of this machine was connected in the direct current circuit from the poles of the battery. Thus connected, the two machines moved as one; the machine M^2 following all the fluctuations of speed in the machine M^1, as if their armatures were carried by the same spindle, the speed being varied between wide limits. To convert this mechanism into a synchronizing device in such a way that the machines could be placed at a great distance from each other was another step. As there was evidently an alternating current in the external loop connected to the armature of the first machine, the reversals of current in this circuit must correspond in time to the half revolutions of the armature of this machine, and an ordinary polarized relay connected in the loop would vibrate in unison with the half revolutions of the armature in the machine M^1. The second machine could therefore be made operative at a distance by simply connecting another relay in an independent line circuit and causing the latter to reverse the current in the second machine. This arrangement, which constitutes the fundamental synchronizing device, is shown in Fig. 2. Here the first machine placed at station x has a polarized relay connected in the external armature circuit, and the relay vibrates in response to the reversals of current in the armature of this machine, which correspond accurately to the half revolution of the machine at x. A line circuit extends from the fixed point of the relay vibrator to the distant station z where it includes the coils of another relay and returns to the middle point of the line battery L B at x, the opposite poles of which are connected to the contact stop of the relay at this station. From these connections it results that any relays in the line circuit will be actuated by reversed currents in unison with the half revolutions of the machine at x. At any distant station, as the one shown at z, another machine has its field coils connected as follows: One terminal to the fixed point of the relay and the other to the middle point of a local battery, the opposite poles of which are connected to the contact stops of this relay, the armature of this machine being connected through ring contacts and brushes to an independent source of direct current. Connected in this manner, the two machines still move in perfect unison, even at variable speeds, though separated from each other by any line resistance through which a polarized relay can be made to act. (F. Jarvis Patten, *Trans. AIEE*, vol. VI, Nov. 1888–Nov. 1889, pp. 115–116.)

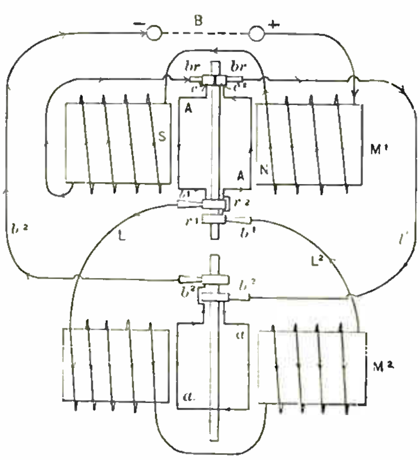
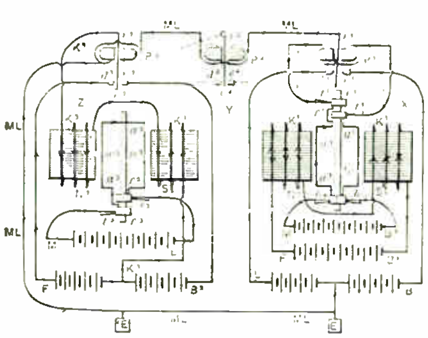


Fig. 1

Fig. 2

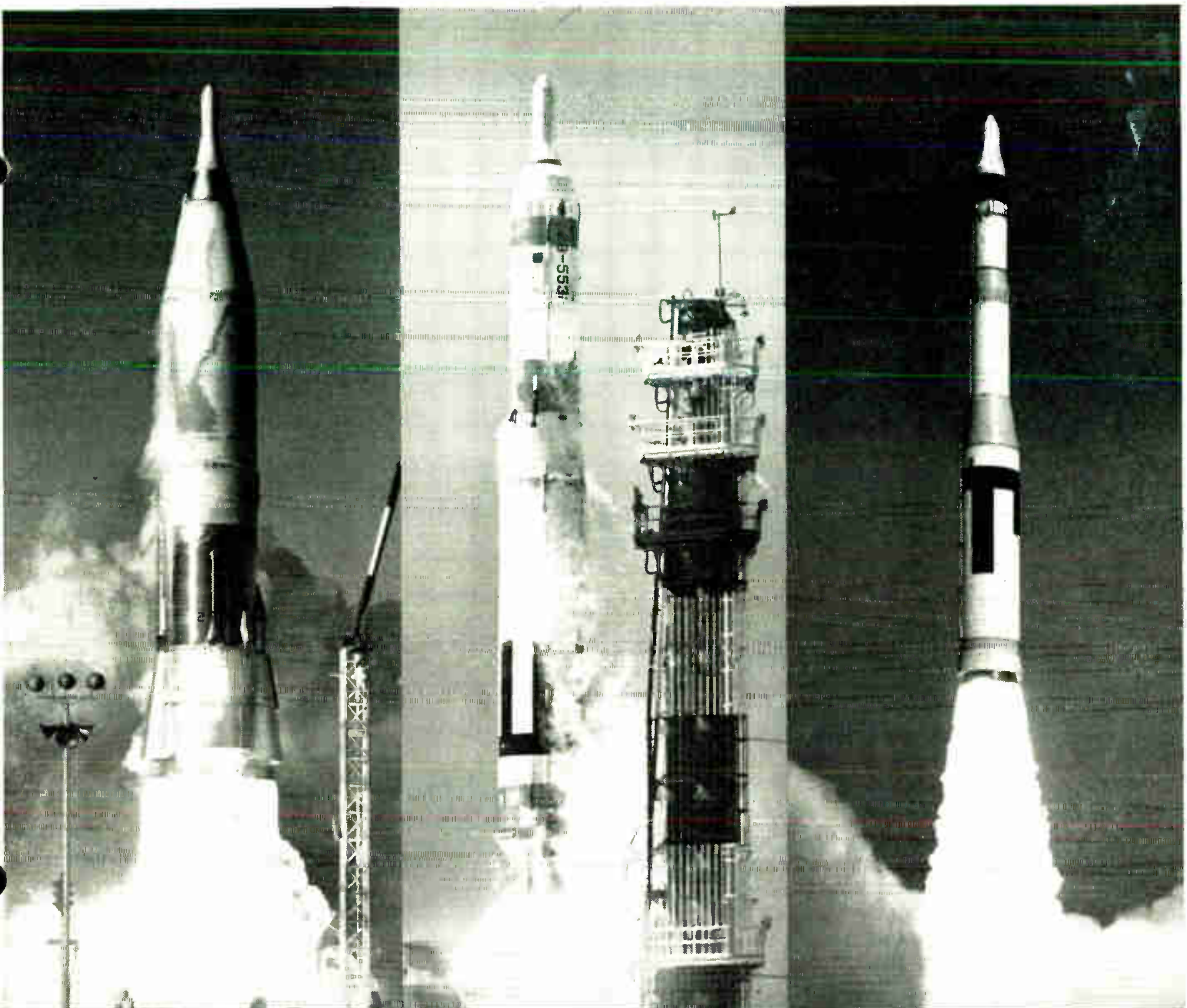


50 years ago

Distance Makes the Heart Grow Fonder. "Thruout this paper, whenever radio range variation is mentioned I shall refer to a variation in the distance over which messages could be transmitted and read on reception. This is a subject to which I have given considerable attention, having taken observations on it and on the strength of atmospheric disturbances intermittently over a period of about fourteen years. In the following paper I shall attempt to present a record of the most interesting and useful of these observations.

"The first records which we shall consider were made at the Manhattan Beach station of the United Wireless Telegraph Company. This station was located at Coney Island in New York City, and its call letter was 'DF' (American Morse code).

"At the Manhattan Beach station care was taken to cut down all losses in transmission and reception. Air insulation was used wherever possible in the transmitting circuit, and wherever solid insulators were employed they were made as long and narrow as practicable. A United Wireless 'Steamship Type' transmitter using about 2 kilowatts was installed in this station. The spark gap itself was not enclosed. For reception, inductively coupled tuners, equipped with Perikon detectors, were usually in use. The receiving apparatus was changed occasionally, but apparently the receivers used were so nearly equal in efficiency



ONLY KIN TEL CLOSED CIRCUIT TELEVISION WATCHES ALL THREE ICBMs

1. THE ATLAS 2. THE TITAN 3. THE MINUTEMAN

There are several reasons why KIN TEL Closed Circuit TV systems are chosen over all others to watch ICBMs and to perform hundreds of other important jobs. One or more of these reasons may well make KIN TEL the TV system you should use, too. To find out for sure, write for technical information or a demonstration. Representatives in all major cities.
5725 Kearny Villa Road, San Diego, California 92112 • Phone (714) 277-6700

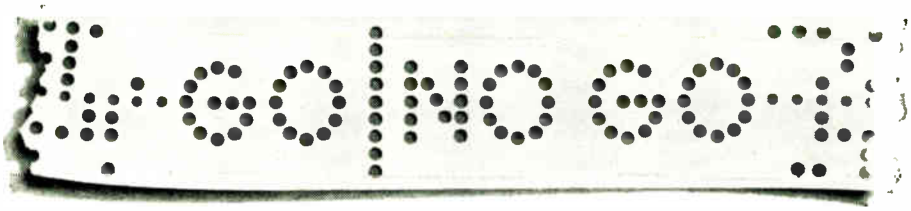
COHU ELECTRONICS, INC
KIN TEL DIVISION

that no material change in the range of reception was thus produced.

"The vessels worked with were mainly equipped with what were known at that time as the United Wireless 1 kilowatt sets. The evolution of this type of apparatus may be of interest. It was first known as American De Forest Wireless Telegraph Company equipment; later, with a few changes introduced, as United Wireless Telegraph Company equipment; and at present these sets, used at a looser coupling and frequently including non-synchronous rotary gaps, are known as American Marconi equipments. A test on 28 of these sets showed transformer inputs between 350 and 1,850 watts, the average transformer input being about 1.1 kilowatts. The average antenna current on the ships at the time the tests shown on Chart 3 were made, was probably about 5 amperes. Characteristic wave lengths (369 to 508 meters) and couplings employed are indicated on Chart 4. The tuners used on board ship were equipped with carborundum detectors. They were, as a rule, inductively coupled.

"Referring to Chart 3, Figure 1, there is shown the maximum daily range between 5 P.M. and 1 A.M. for the period between October 12, 1908, and October 15, 1909. The ranges, indicated by dots and connected by lines, refer to the longest distances (100 to 1500 miles) over which messages were sent and received on the respective days. The distances were checked by reference to the message records kept on board the ship. Practically no records of distance were made except where the distance in question was actually attained in handling business either sent or received, and acknowledged. Messages handled during this time included approximately 115,600 words, of which about 82,800 were received and about 32,800 were transmitted, not counting checks, repetition, acknowledgments or conversation. The operators who made these records were selected for their reliability and were given a bonus, the amount of which depended on the number of words handled and the distance; that is, ten words sent or received over a distance of 1,000 miles paid a considerably higher bonus than ten words sent or received over a distance of 100 miles . . . .

"Before I started the operating scheme with bonus for these tests at Manhattan Beach, the operators for that station were not particularly carefully chosen and they were given no particular incentive to handle considerable amounts of business, except that they were repri-



DATA PROCESSING SYSTEMS FOR SPACE

Advanced STL digital telemetry units, decoders, and command distribution assemblies are now being used on NASA's OGO and Pioneer, and the Air Force's Nuclear Test Detection spacecraft. STL hardware and experience with on-board data processing equipment is being applied in the development of new systems which will perform checkout and maintenance functions in space. This advanced technology requires circuit designers, logic designers, and digital systems engineers. For Southern California or Cape Canaveral opportunities, write Professional Placement, One Space Park, Dept. P-2, Redondo Beach, California, or P.O. Box 4277, Patrick AFB, Florida. STL is an equal opportunity employer.

TRW SPACE TECHNOLOGY LABORATORIES
THOMPSON RAMO WOOLDRIDGE INC.

manded from time to time if they did not carry on sufficient work. As a matter of fact, they handled very little business. When selected operators were in charge and were paid a bonus depending on the amount of business handled and the distance over which this business was handled the range was materially increased and the business handled was multiplied many times. In one instance the increase over a corresponding time was found to be 2,700 per cent.

"Apparently, in the winter time, the operator expects to receive long distance messages, and he therefore makes a greater effort to carry on such work, but during the time of summer atmospheric disturbances he is discouraged by the difficulty of communicating over long distances and therefore he does not even attempt long distance work. There are apparently considerable individual differences in the ability of persons to receive signals, and the acuteness of the individual in receiving signals also varies from time to time and from place to place. For example, among other peculiarities, it was noted that if barely audible signals were received, placing a cigar in the mouth rendered the signals inaudible. It is quite clear that a great amount of the variation in the strength of received signals which has been ascribed to ionization and consequent conductivity of the atmosphere is due to some other factors." (Robert H. Marriott, "Radio Range Variation," *Proc. IRE*, Mar. 1914, pp. 37–38, 45–47.)

High-Tension Transmission. "The different power companies operating on the Pacific Coast have done a great deal in the way of pioneer work in the transmission of power over long distances at high voltages. This was due, to a considerable extent, to the fact that there is an abundance of water available for the hydroelectric generation of power, and also to the fact that the cost of fuel is very high as compared with that further east.

"The first long-distance transmission to be put into successful operation on this coast, if not the first in the United States, was that of the San Antonio Light & Power Company in southern California, at what is known as their Pomona plant. This transmission was for a distance of 30 miles (48.2 km.) at 10,000 volts.

"Transformers were at that time not manufactured for such high voltages, and it was necessary to connect in series ten 100-volt transformers for stepping up. The low-tension windings were connected in multiple.

"The first polyphase installation was made in 1893, at Mill Creek, and furnished power to the city of Redlands.

"The success of these undertakings gave a great impetus to the hydroelectric industry, and inside the next five or six years, there were several different companies operating at voltages as high as 60,000." (P. M. Downing, "Problems of High-Tension Transmission Lines," Trans. AIEE, vol. XXXIII, part I, 1914, pp. 124–125.)

25 years ago

UFO? "'Some Contributions of Radio to Other Sciences' was the subject of a paper by J. H. Dellinger, chief of the radio section of the National Bureau of Standards. The application of radio principles to meteorological investigations and weather forecasting was treated in detail . . .

A description was given of radio waves which were reflected from some source far beyond the earth's atmosphere . . .

This meeting was held jointly with the Franklin Institute in its auditorium, December 8, 1938."

(*Proc. IRE*, Feb. 1939, p. 155.)

Engineering and Peace. "Engineering processes and engineering devices have been (and are being) abused by use in warfare, but not by choice of the leading minds of the engineering profession. The profession, with few exceptions, stands on the conviction that war is tragic, expensive, inefficient, and unnecessary. . . . It lies with us to hold so steadfastly to the duty of more fully developing the peaceful occupations of agriculture, industry, and commerce that the minds of our people may be turned away from contemplating aggressive expansion and warfare. Engineering favorably influences all of those peaceful occupations, and we engineers have an obligation to ourselves and our fellow citizens to encourage the highest practicable development of these ethical aspects of engineering.

". . . responsibility rests on critics of our status, who are failing in the call for agitation and education of the public to a realization of the relative fitness of things. This is a personal and group matter to achieve, and it cannot be achieved by partisanship or foul play." (Dugald C. Jackson, "The Social Significance of Engineering," *Electrical Engineering*, vol. 58, Feb. 1939, p. 63.)

Spectral lines

Education for Living. The educational fraternity hears many complaints about the inability of engineers to write, read, and spell. It is therefore refreshing to learn from a teacher of psychology, in a letter to the editor of *Science*, that students in the social sciences have parallel difficulties, demonstrated by a spelling such as "Cornepincus," probably translatable freely as Copernicus! Of course, science and engineering teachers often pass the buck to the high schools, whose teachers feel that the grade schools are ineffective, and at the latter level there is concern for the quality of work in the kindergarten.

Certainly, if I am to apply my own experiences, I would say that one can learn to write only by writing—and this is almost a lost art in secondary and higher education, due to the pressure of student numbers plus the ease of grading of multiple-choice tests. Consider the plight of an average high school English teacher in the United States, forced to handle classes totaling 150 different students during a week. To ask each student to write an essay that will require ten minutes of grading time each, adds 1500 minutes of grading time to that teacher's load in that week—and 1500 minutes is a mere 25 hours of extra work on top of an already heavy schedule. If practice is the means by which we can remedy those spelling and writing deficiencies, then we, as voters, are at fault in not supporting our schools so that teaching loads can be maintained at practicable levels.

There is more to the problem of broadened abilities for engineers than reading and writing. Our engineering colleges have expended much thought and effort in providing programs in the social sciences, humanities, and liberal arts for engineering undergraduates. That such programs were a necessity in engineering education was pointed out as far back as 1929, and augmented during the depression by calls from industry for such broadening in the education of engineers. These calls resulted from engineer shortcomings exposed during the twenties and thirties, as more engineers moved into management and supervision, into positions where they were measured against men with liberal arts or legal training, or into research where they rubbed shoulders with science-trained personnel. The engineer of that day, with only four years of rather narrow technology, was often found deficient when compared with the man with some degree of education in the humanities upon which was overlaid a solid business, legal, or science education.

As a result of these needs in industry the engineering colleges have appreciably broadened their curricula, and have replaced such courses as surveying with social science, lubrication with literature, psychometry with psychology, or hardware with humanities in general. Along with this trend has come in recent years the necessity to meet the technical conditions of our changing world, and thus to increase the mathematical emphasis, and to study engineering from the viewpoint of its science base rather than the temporal application of that science. As a result of these several changes, the psychological inquisitors tell us that the field of engineering is now attracting high school graduates of broader scholastic interests and higher intellectual abilities. The entering students tell us they chose engineering because they were good in mathematics and science—and it seemed that engineering offered opportunity to those with such interests. Others of equal ability avoid engineering because even its present programs show insufficient opportunity for further broadening of their educational objectives. Thus does the program attract and adjust the input to that program.

After graduation young engineers are now found to handle complex technical problems, and also are apt in presenting a paper to a Section, or in planning Section or Region activity; they can and do participate in civic and educational activities.

That such activity is going on is evidenced by the presence of our General Manager on his local school board as chairman; a vice president of the IEEE is an important member of his school board, and the president of his company serves on a university Board of Trustees; a past IRE president has improved the financial planning for his local school system.

But recently a report has reached us that one of our industries has banned such civic activities on the part of several of its engineers—that it has threatened termination of employment for such men if they are successful in a race for the local school board. Possible interference with the man's work was given as justification for such action. In argument, it has been suggested that good engineers are often good because they are willing and able to work at about 150 per cent of capacity—industry should be thankful that it has such outstanding men, and glad to concede the 50 per cent to local betterment and future gain.

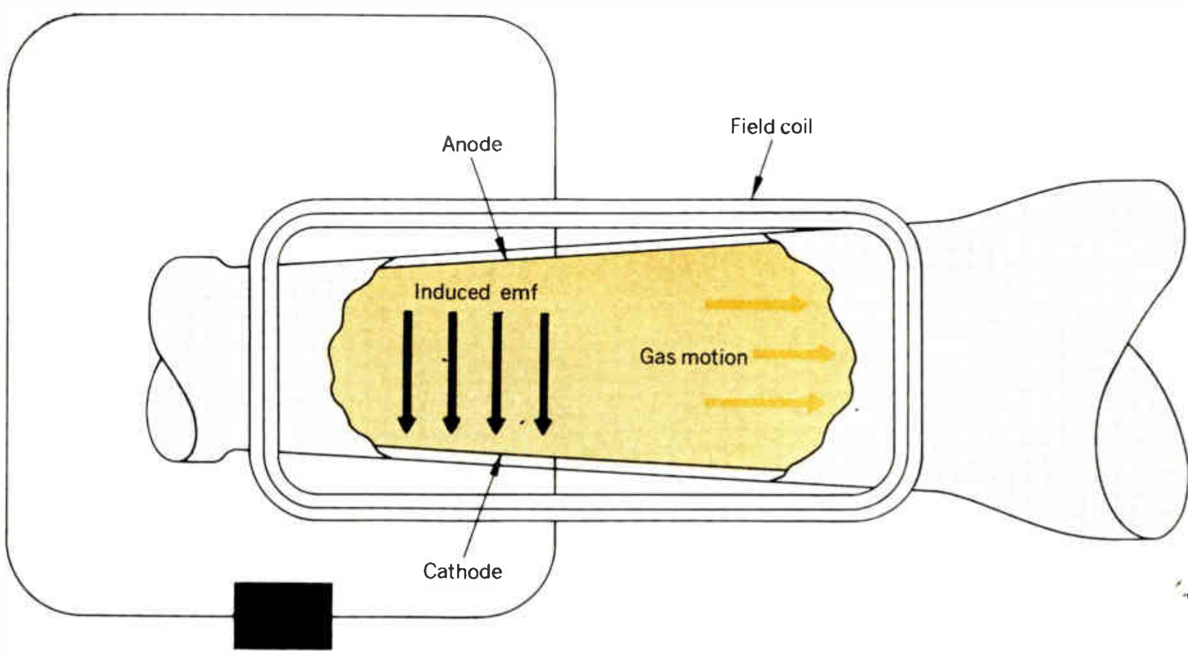
To educators the industry plea of the twenties and thirties for increased breadth in engineering education was rational and consistent with the general objectives of college education—to learn to live a life. If an industry now believes technical education should only teach to earn, then perhaps it should not be recruiting its technical personnel at the college level.

For the colleges and the profession can have only one answer here: an industry cannot simultaneously have its cake and eat it! *J. D. Ryder*

MHD power generation

The MHD generator appears to have wide applicability. Its development is proceeding both rapidly and encouragingly. Initial central station use is possible in the mid-1970s

Thomas R. Brogan *Avco-Everett Research Laboratory*



The potential of the magnetohydrodynamics (MHD) system for commercial power generation was first assessed in November 1959.[1] It was indicated then that, although an extensive research and development effort would be required, cycle efficiencies well above 50 per cent and compatible capital cost could be achieved with MHD. At about the same time, Avco and a group of electric utilities* entered into a cooperative arrangement to investigate the development of the MHD concept for commercial, central station power generation. The American Electric Power Service Corporation has acted as agent for the utility group in the program which is now in its fifth year.

Since the beginning of the Avco utility program, many other MHD large and small generator programs have been initiated in this country and abroad. Interest in the concept has multiplied. The various programs have included military and space applications as well as the commercial central station use. It now seems appropriate to make an evaluation of the current status of the MHD generator development, to estimate its future potential, and to assess the problems which remain to be solved before practical application is feasible. This paper is directed to that objective. The observations will be directed primarily to the development of the large combustion-fired MHD generator with which Avco has been particularly concerned.

Review of basic principles

The basic principles of MHD power generation have been described previously[2] but a brief review may be helpful, to show both the advantages and the problems connected with the MHD concept.

Figure 1 is a schematic representation of the elementary dc MHD generator. Gas from a high-pressure source flows from left to right through the generator which, in principle, is little more than a cylindrical pipe. A magnetic field is impressed perpendicularly to the direction of gas motion and an emf is induced in a third mutually perpendicular direction. If the gas is a conductor of electricity, current may be tapped from the gas flow to the load by means of electrodes as shown, in order to deliver power to the load. The action is identical with that of a conventional generator in which a copper conductor is rotated in a magnetic field. When current is drawn from a conventional generator, a torque is required to rotate the wire in the field. In the case of the MHD generator, a pressure difference is needed to force the gas through the field when a current is drawn. Thermodynamically, the operation of the generator is identical to that of a turbine. Useful work is extracted from a gas flow at the expense of pressure and enthalpy drop, and a "turbine efficiency" can be computed for any particular generator flow.

The advantages of an MHD generator arise directly from its conceptual simplicity. It has no highly stressed, hot moving parts of close tolerance. The walls can be cooled below the temperature of the working fluid. There appears to be no practical limit to the size of the duct, so that the generator is appropriate for high single-unit power level. Since the walls can be cooled, the temperature of the working fluid is not limited by generator construction but rather by the nature of the thermal source upstream of the generator. This leads to a high capability for peak cycle temperatures which must lead to high efficiency. It should be emphasized here that the flow in this type of device is very similar to the normal type of flow in a duct, with the exception of the body force action on the gas caused by MHD. The flow does not resemble the conditions in a conceptual fusion reactor where the gas is actually confined by the magnetic field.

Several authors[1,3,4] have presented both open and closed conceptual MHD power cycles. Fundamentally, all such power cycles consist of a regenerative gas-turbine cycle, topping a conventional steam cycle, with the exhaust of the MHD generator being used both for regeneration and steam generation. In the case of the open cycle, the oxidizer may be pure air,[1,4] or oxygen enrichment may be employed to raise the combustion temperature and increase the gas conductivity.[3] In the case of the open cycle, it seems likely that overall cycle efficiencies in excess of 50 per cent will be attainable. In the case of the closed cycle, with a nuclear heat source, the application appears to depend on fundamental improvements in the ability to have gases conduct electricity at the discharge temperatures appropriate to nuclear reactors.

It is important to emphasize that the MHD generator is only a component in a heat engine where the objective is to transform thermal energy into useful work at high efficiency. In practice we are limited to two thermal sources: first, the combustion energy of fossil fuels and, second, the nuclear reactor. In the case of fossil fuels, the peak cycle temperature is about 5000°F. In the case of the nuclear system the peak temperature capability is less certain, but at present it is hardly above 1500°F.

If the MHD concept is to be useful, the gas must conduct electricity, and herein lies the principal difficulty with the concept. Electrical conductivity in gases is a very steep function of temperature and is dependent on one or more species in the gas having a low "ionization potential," which is the energy necessary to strip an electron from one of the gas atoms or molecules. Upon examination, we would find that, even at the peak combustion temperatures of ± 5000°F, the normal products of combustion would have an electrical conductivity far too low for the MHD generator to be practical. To circumvent this difficulty, a material of low ionization potential, called "seed," is added to the gas with typical results shown in Fig. 2. Here the conductivity of stoichiometric JP4 (kerosene)–oxygen flame seeded with one per cent potassium by volume is shown as a function of temperature at atmospheric pressure. Temperatures between 2000° and 3200°K are appropriate for adequate conductivity in combustion gases in thermodynamic equilibrium. On the same scale, the conductivity of copper is 5×10^7, between 10^6 and 10^8 times the conductivity of combustion gases. It is apparent that no amount of ingenuity can circumvent such a difference and, consequently, the MHD generator must be a very different device from a conventional generator. These differences

* The present utility group includes: American Electric Power Company and its subsidiaries: Appalachian Power Co., Indiana & Michigan Electric Co., and Ohio Power Co.; Boston Edison Co.; The Dayton Power & Light Co.; Indianapolis Power & Light Co.; Kansas City Power & Light Co.; Louisville Gas & Electric Co.; Union Electric Co.; and The United Illuminating Co.

Fig. 1. Line diagram showing principal elements of the basic dc magnetohydrodynamic generator configuration.

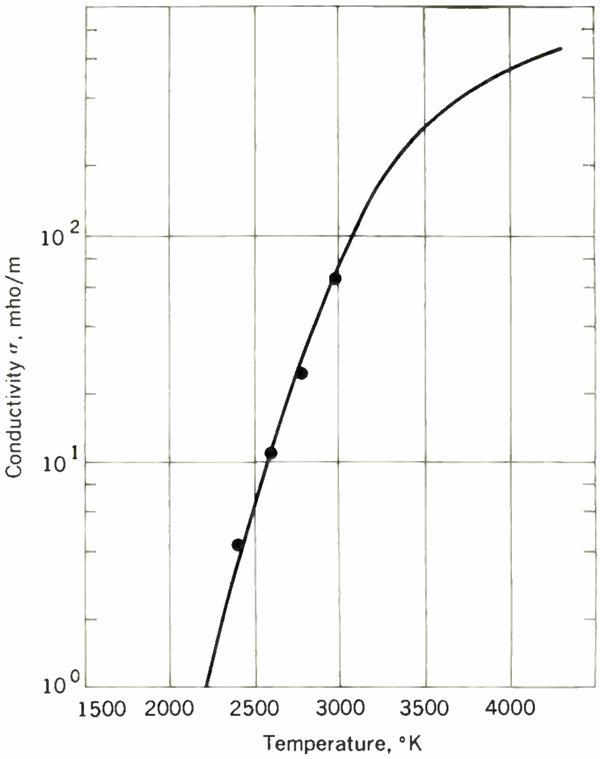


Fig. 2. Graph showing electrical conductivity of seeded JP4–oxygen flame plotted as a function of the temperature.

relate to the minimum useful size of an MHD generator. The power output is proportional to the electrical conductivity times the volume of the device. The relative losses, on the other hand, (heat transfer, friction to the walls and power to generate the magnetic field) decrease as the volume is increased. Therefore, a high volume-to-surface ratio is required. When the numbers are entered, it appears that for commercial application the MHD generator must be used with units of several hundred megawatts' capacity. Since central station units in this size are being constructed, the generator is ideally suited to this application. In other applications, notably the limited-duration generator discussed later, net outputs upwards of 20 MW or so would seem to be practical.

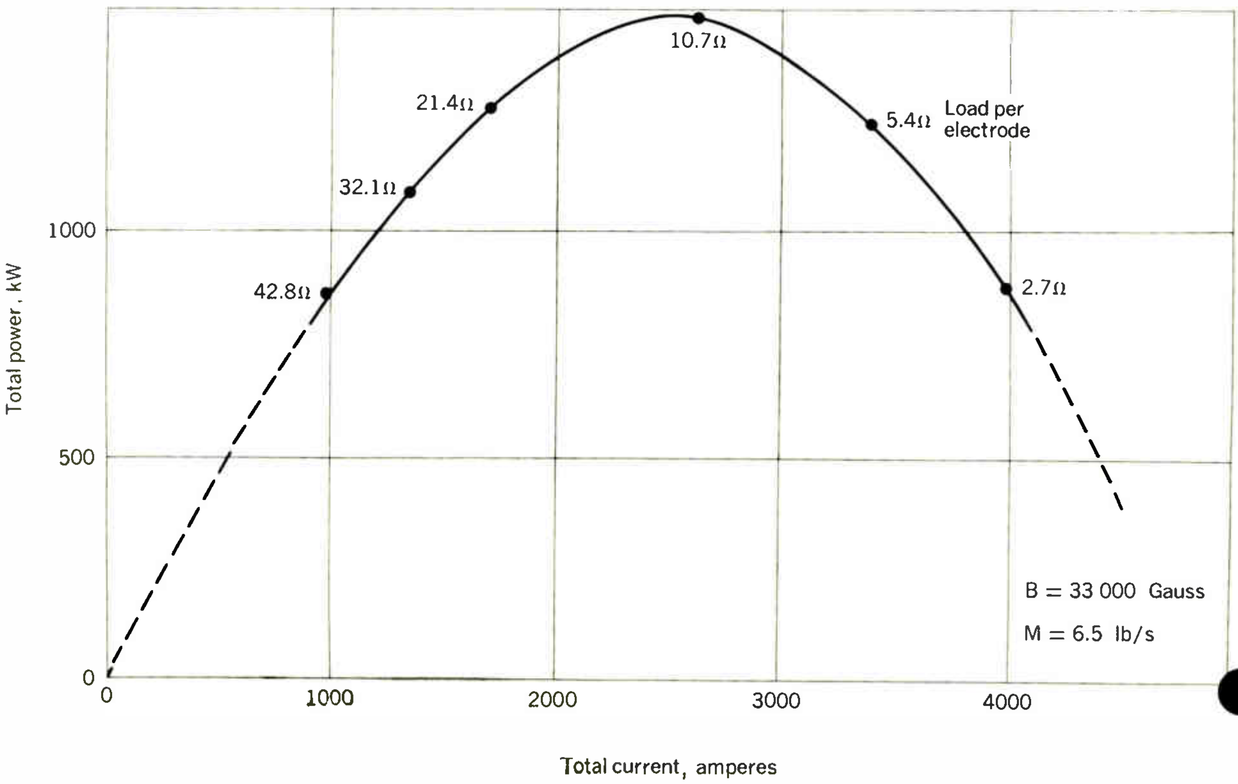
The advantages of the MHD generator stem from its conceptual simplicity. The problems arise because of the limited temperature range over which the gas can be an adequate electrical conductor.

Status of MHD power generation development

This section reviews very briefly the current status of MHD generator development. The areas of generator performance, fossile-fuel combustion, duct development, seed recovery and corrosion, and field coil development are discussed.

Performance. The performance of MHD generators has been studied extensively at the Avco-Everett Research Laboratory (AERL) using a large combustion-fired generator.[5] Typical data from this generator is shown in Fig. 3, where the total power output of the generator is plotted as a function of the total current. The peak output is 1.48 MW. The generator is operated for only short periods of about 10 seconds, and is designed solely for the study of generator performance. Under the operational

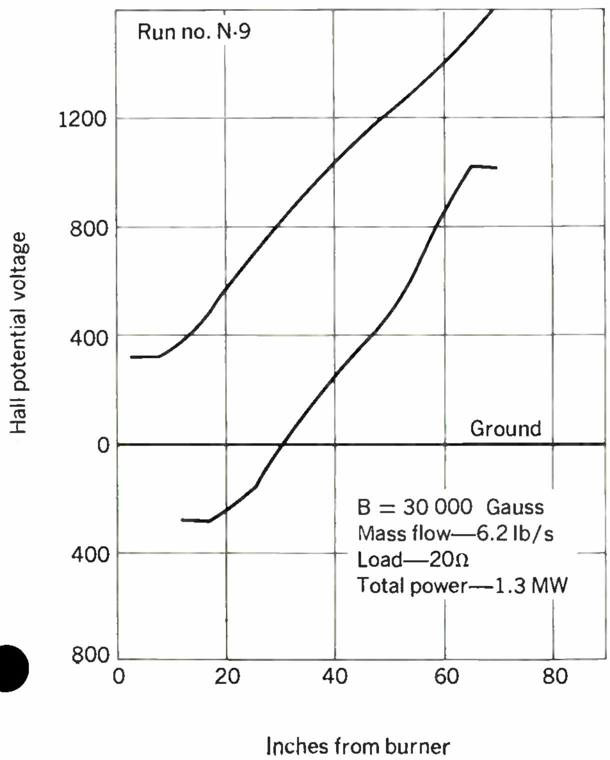
Fig. 3. Measured power output is shown plotted against total current for the experimental MHD (Mark II) generator.



conditions of the unit, the output can be accurately predicted; and we believe the ability to predict performance under these conditions is adequate for design purposes. At the peak output, the measured expansion-engine (turbine) efficiency is 46 per cent through a stagnation-pressure ratio of 3. Since, at this small size, wall losses are important, extrapolation of these results to generators of sizes appropriate for commercial stations would indicate that expansion-engine efficiencies in excess of 80 per cent may be expected for the MHD generator.

Figure 4 is a typical voltage distribution along the experimental channel for the generator just described. The distance between the top and bottom lines (anode and cathode) is equal to the voltage across the generator at any station along the duct. Note, moreover, that there is an axial field along the length of the duct. This is the so-called Hall field. It arises because the output current is caused by a flux of electrons drifting across the magnetic field. The drift velocity of the electrons, crossed into the magnetic field, induces the Hall potential along the axis. If a current flow from the Hall potential is allowed, the effect is to cause a back emf to oppose the induced emfs which are supplying the loads. To prevent the flow of axial Hall current, it is necessary to segment the generator axially;[2] that is, instead of the continuous electrodes indicated on Fig. 1, the anode and cathode must be broken up into many individually isolated segments, each feeding a separate load. This has been done successfully in the experimental generator. The Hall potential causes a considerable complication in the design of the device, as it leads to multiple outputs. In a practical commercial device, each of the outputs would be fed to an inverter, and the alternating output of the individual inverters combined in a transformer for transmission.

Fig. 4. The typical voltage distribution on the experimental MHD generator is shown in the graphic plot.



The Hall effect just described is of extreme importance for the design of MHD generators. Its importance is measured by the value of the "Hall coefficient," which is equal to 2π times the average number of gyro orbits sustained by an electron between collisions with gas atoms or molecules. The Hall coefficient is proportional, for a given gas, to the ratio of magnetic field strength to gas density. The power output per unit volume is proportional to the square of the magnetic field strength. Thus, one would like to employ the highest possible field strength. The Hall potential limits the field strength which can be employed in any given situation. With our present capability, successful generator operation may be obtained at Hall coefficients of approximately 2.5. Strong efforts are being made to increase this capability. As will be discussed later, the Hall effect is of very great significance in the nuclear application.

Combustion chamber development. In the fossil-fired commercial central station, combustion will take place in preheated or oxygen-enriched air at pressures of several atmospheres, and at combustion temperatures of approximately 5000°F. Very high heat release rates per unit volume, low wall loss, and low pressure drop are all required. Aside from the always difficult problem of establishing sufficient endurance, there appears to be little difficulty in firing heavy fuel oil or natural gas. Heavy oil is fired regularly at AERL in the long-duration tests of MHD generator duct components. There was serious question, however, as to whether coal of commercial pulverization could be fired at heat release rates high enough to avoid intolerable wall heat loss in the burner. To this end, in 1963, experiments on MHD coal combustion were conducted at AERL's Long-Duration MHD Test Facility. The techniques of ramjet or turbojet afterburner combustor design were employed in these tests, which demonstrated that combustion at the required heat release rates was possible. The long-term endurance of coal burners remains to be established.

MHD generator duct construction. The largest effort in MHD generator component development has been put into the evolution of the generator duct. The high-speed flow at elevated temperature induces large heat-transfer rates to the duct walls. The walls must be protected by a coolant—normally boiler feedwater in a commercial plant. Two walls of the generator must insulate electrically, both normal and parallel to the duct; the other two walls (the electrode walls) must make a good electrical contact with the gas in order to transfer power to the loads, while simultaneously insulating along the axis of the duct to avoid the short of the Hall potential. With regard to the insulating walls of the duct, AERL has developed the "peg wall" concept.[6] It consists of closely spaced, water-cooled metallic segments separated by a thin section of refractory which provides electrical insulation between the individual segments. The tests have established endurance times in excess of one week. With extensive development, this type of construction, should be appropriate for commercial MHD power plants.

In the case of MHD generator electrodes, proprietary considerations prohibit a full discussion; however in the combustion products of commercial fossil fuels, the endurance characteristics of electrodes are similar to those of the insulating side walls.

Seed recovery and corrosion of steam generator heat transfer surface. As previously mentioned, the working fluid in an MHD generator must be seeded to make it a conductor of electricity. The seed is commonly the hydroxide or the carbonate or sulphate salt of potassium. In some cases, cesium is the metallic element of a similar compound. The inexpensive chloride cannot be used, because the chlorine atom is highly electronegative and its presence in significant amounts can lead to a large reduction in conductivity. Without seed recovery, costs of about one mill/kWh would be sustained. This, of course, is uneconomical. In experiments at AERL, it has been demonstrated that recovery rates appropriate to total seed costs of 0.05 mills/kWh for coal or 0.1 mill/kWh for oil or natural gas can be achieved. The advantage of coal is in the presence of coal ash alkali, reducing the make-up required.

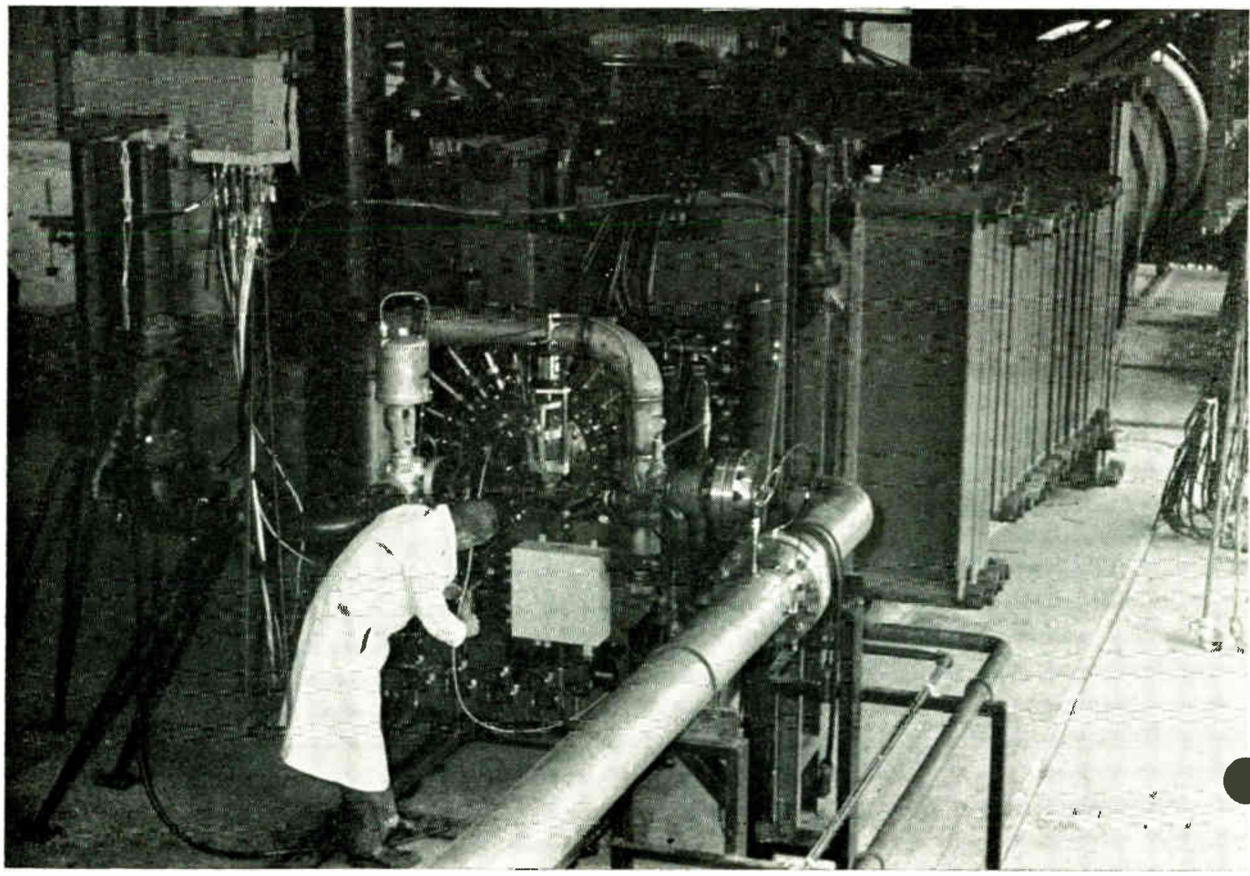
At AERL, the seed-recovery experiments just described have been combined with the study of heating surface corrosion under appropriate alkali loading of the flue gas. The normal corrosion mechanisms experienced in superheaters and reheaters in conventional plants have not appeared after several hundred hours of tests; but this, of course, is by no means conclusive. It appears, however, that the temperatures in a metal oxidizer preheater will be limited to 1530°F. or below, because of corrosion by alkali sulphates above this temperature.

MHD generator field coil development. In the early stages of the generator development, it had been expected that the magnetic field for the generator would be produced by copper field coils. Even for the MHD generator appropriate to a plant of several hundred megawatts' output, the power dissipation in these field coils was an appreciable fraction of the generator output, and consumed a considerable part of the advantage in efficiency. The dream of the superconducting field coil with identical zero-power dissipation was frustrated because the field strengths available from the then existing superconducting materials were inadequate for use with an MHD generator. This was changed in early 1961, by the discovery of the high-field-strength, hard superconductor,[7,8] an event followed by the intensive development of the superconducting field coil* for wide application—including that of the MHD generator. AERL is participating in this development.

In practice, the engineering application of the high-field-strength superconductor has not been easy.[9] The reason for this is that although the performance of a short sample of superconductor is accurately predictable, the performance of a long length of the same material, when wound into a coil, is not as reliable. Means must be employed to protect the coil against catastrophic failure on the appearance of nonsuperconducting regions. Since the energy stored per unit weight of coil material increases with coil size, the importance of this protection—and the consequences of failure to protect—become increasingly serious as the size increases. Higher fields have been produced with smaller coils. The achievement of larger current-carrying capacity per unit area, lower cost per

* See the article "Superconducting Magnets" in this issue, page 103.

Fig. 5. Prototype self-excited MHD (Mark V) generator that is now under test at the Avco-Everett Research Laboratory.



unit weight, and more bulk materials of larger cross-sectional area, would greatly facilitate the use of the superconducting coil with an MHD generator.

Military applications of high-power, limited-duration generators.* It appears that there are military applications requiring very large bursts of power for periods ranging from a few seconds to several hours. Because of its conceptual simplicity, the MHD generator can fulfill such applications at low capital cost. Since the operating duration of these devices is limited, the fuel cost is not a serious consideration. The generating system is simply composed of fuel and oxidizer storage, means for delivering them to a combustion chamber, an MHD generator channel, and the magnet driven by a portion of the generator output. At this writing (Dec. 1963), a prototype of such a limited-duration generator is ready for initial operation at AERL (see Fig. 5). The view shows the manifolding which conveys the oxygen, cooling water, and fuel into the combustion chamber, which is located at about the center of the picture, pointing into the magnet. The large steel beams contain the magnetic forces from the 33 000-gauss magnet. The MHD channel is inside the magnet, and the exhaust duct which conveys the spent combustion gases outside the test cell are shown to the rear. One-half of the generator's gross output of 40 MW is used to drive the magnet. It is believed that this will be the first instance of self-excitation of an MHD generator. The remainder of the generated power is delivered to water-cooled load resistors which are not shown. In sizes larger than the 20-MW (net) output of this device, the ratio of net output power to gross output power would more closely approach unity than the 0.5 figure appropriate to this Mark V generator.

The potential capability of these limited-duration generators is shown in Fig. 6, where the weight of the copper field coil is indicated as a function of gross electrical output in megawatts. Operation with oxygen and hydrocarbon fuel is assumed. The weights are shown for field coil dissipations of 3 and 30 per cent of the gross output at room temperature. If the coils are cooled by the liquid oxygen oxidizer, the dissipation would be reduced by a factor of five for equivalent weight, or alternatively, the weight could be reduced. The double line indicates the weights of an ideal pressure vessel to contain the magnetic stresses in the magnet; actual structural weights would be several times this value. Weights for superconducting coils are also indicated. Unless the weight is the first consideration, it is unlikely that superconductors would be used, since their capital cost could not be justified. The point indicated by "Mk. V" represents the design point of the 20-MW generator shown in Fig. 5. Note that it is of minimum size for this type of generator. If it were any smaller, no amount of copper could reduce coil dissipation to the point where the generator could self-excite. This application is useful for net outputs of 20 MW or over. In large sizes, very impressive improvements in cost per kilowatt may be achieved with this type of device.

Nuclear MHD generators. At first glance, it would appear that the nuclear heat source is extremely attractive for use with the MHD principle. Clean monatomic gases could be used, and the construction of the generator duct

* The work reported here has been sponsored by the Advanced Research Projects Agency and administered by the Aeronautical Systems Div. of the USAF, under Contract AF33(657)-8380.

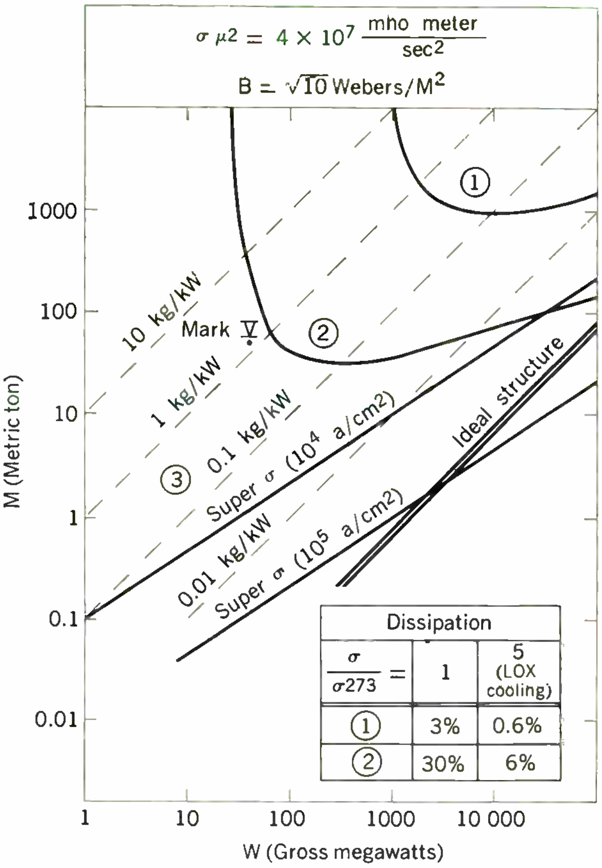


Fig. 6. Graph showing the essential weight-power characteristics of self-excited generators of the Mark V type.

would be greatly simplified. In practice, however, this is not the case because of the limited temperature capability of the nuclear reactor. The only reactor capable of producing gases at sufficient temperature to build an MHD generator with thermal ionization is the nuclear rocket which heats hydrogen gas for limited durations.

The limited temperature capability of the nuclear reactor has led to a search for methods to make the generator working fluids conduct electricity at lower temperatures than can be achieved with straight thermal ionization (where the electron density is in equilibrium with the temperature of the gas). This effort involves ionization of the seed atoms by high-speed electrons as they flow through the gas to the load, and is referred to as "nonequilibrium" or "high-field ionization."[10,11] To achieve such high-field ionization, MHD generators must be operated at Hall coefficients between 5 and 10 and under conditions of extreme gas uniformity with regard to conductivity and velocity. The problems with this type of operation have been considered by Rosa.[12] To date the search for high-field ionization in an induced electric field has not yielded positive results.

At AERL, apparatus known as the "inert gas heater" is used in these nuclear studies. In this device a container filled with graphite rods is heated electrically, and an inert gas, usually argon, is blown through the hot graphite and heated to temperatures up to 2000°K. It is then seeded with cesium and blown through an experimental MHD generator as shown in Fig. 6. Although positive results regarding high-field ionization have not been achieved,

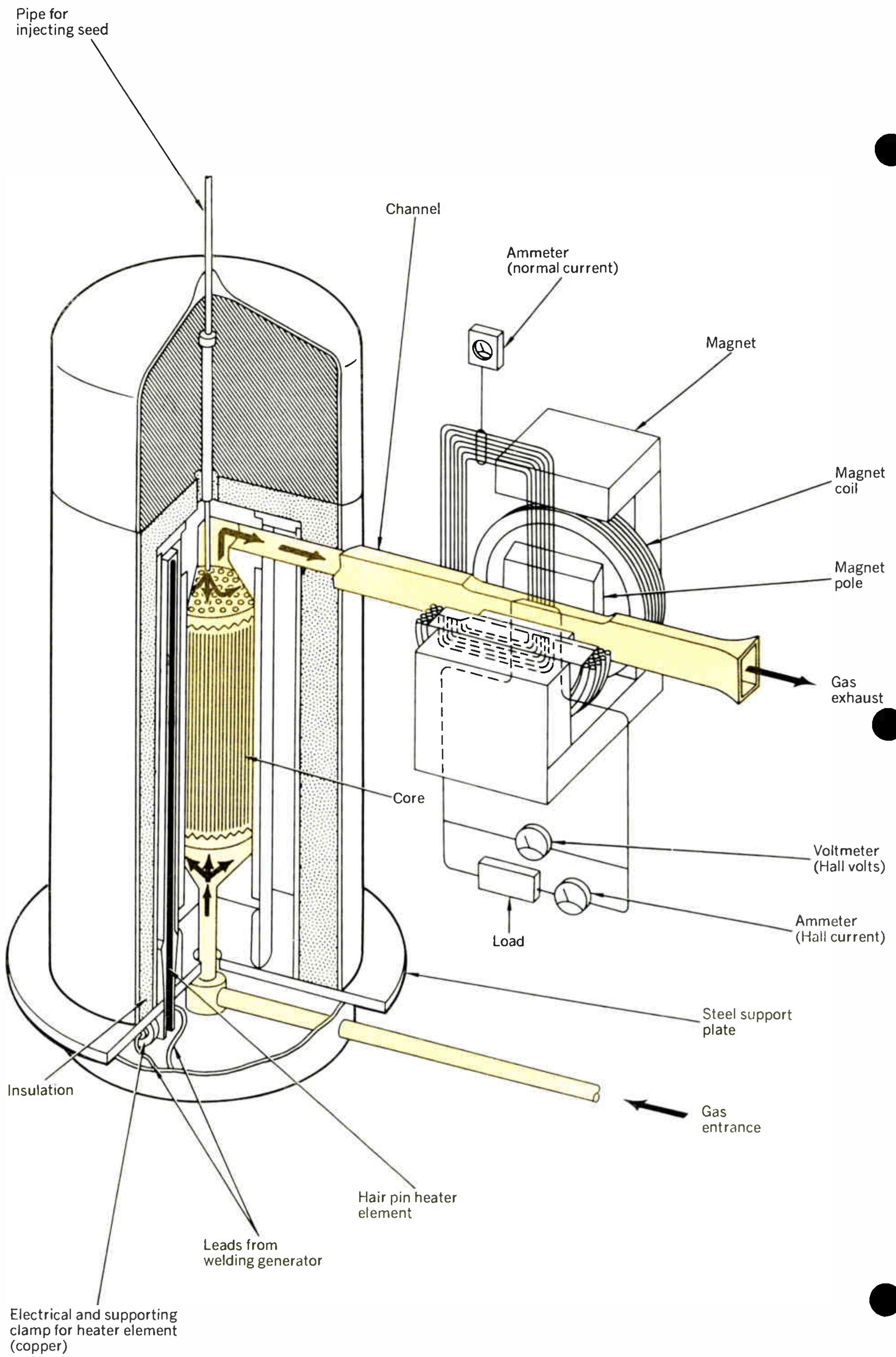

Pipe for
injecting seed
Channel
Ammeter
(normal current)
Magnet
Magnet
coil
Magnet
pole
Gas
exhaust
Core
Voltmeter
(Hall volts)
Load
Ammeter
(Hall current)
Steel support
plate
Insulation
Gas
entrance
Hair pin heater
element
Leads from
welding generator
Electrical and supporting
clamp for heater element
(copper)

Fig. 7. Graphite blowdown inert gas heater for the study of the nonequilibrium ionization in MHD generators.

recent work has been encouraging, and it is hoped that important progress will soon be reported in this area.

Future potential of MHD generators

Having concluded a description of the current "state of the art" of the MHD generator, we can make an estimate of its future potential.

Commercial central station power generation. In principle, both chemical and nuclear heat sources can be applied to the commercial central station application. However, the nuclear application is dependent on the achievement of nonthermal ionization in MHD generators or, alternatively, on substantial improvement in the temperature capability of nuclear reactors, so that the generator can operate on thermal ionization. We will therefore confine our remarks here to the application using the chemical energy of fossil fuels as the heat source. It is to this application that AERL has devoted a great deal of effort.

Based on the technical status outlined previously, we may expect that the MHD system can be developed to commercial fruition. Developments in generator performance, duct and coil construction, seed recovery, and systems are encouraging. We are confident that cycle efficiencies of 50 per cent or greater will be attained. Although an accurate assessment of the capital cost is not now possible, it appears that these costs will be compatible with economic power production.

The development program for the combustion-fired MHD generator for commercial central station application is in its fifth year. It has proceeded to the point where practicability has been established and we may consider the construction of an experimental MHD power plant of some 30-MW output. If all goes well with this test plant, the next step will be the construction of a commercial prototype of about 100-MW net output. This would be followed by a full-scale plant of somewhere between 500 and 1000-MW output per unit. The timetable for the application of the MHD concept is dependent upon the time necessary to carry out the three steps outlined in the previous paragraph. Thus, concept could come to fruition for commercial central station power generation sometime in the 1970s.

Military and space applications. MHD is most attractive for those military applications requiring a burst of power at high level for periods of several seconds to hours. A prototype generator of this type is under test at AERL, and the capability to supply this type of power may be said to exist. The power can be supplied for very low capital cost per kilowatt and weight per kilowatt in large sizes, and these equipments could be portable or even flyable.

The MHD generator could be used in space to provide thrust augmentation of a Rover type, nuclear rocket and to supply power for electrical propulsion apparatus in deep-space penetrations.

Use of the MHD generator to provide thrust augmentation of the nuclear rocket has been already proposed.[13] The rocket delivers hydrogen gas at a temperature of ± 2500°K, which is sufficient for thermal ionization. Since the methods of augmentation depend on thermal ionization, the questions of applicability are systemic, rather than of a fundamental nature. The rocket application is for limited-duration use only.

Some of the deeper space missions may employ an electric rocket engine in which a working fluid is accelerated to very high velocity by various electrical means, including the inverse of an MHD generator. The value of such applications depends largely on the achievement of a very light specific weight power supply, and there are indications that at high-power level the MHD power supply may have merit. The heat source is a nuclear reactor of the type which would necessitate the employment of nonequilibrium ionization. Here again the application is dependent, in the first instance, upon a demonstration that this type of ionization can be achieved efficiently. In some power cycles which have been considered, alkali metal vapor would be used as the generator working fluid in a Rankine cycle. The higher radiator temperatures here would permit low radiator weights. Other potential cycles would use the Brayton cycle, but these are accompanied by relatively low radiator temperature and the necessity to compress the working fluid in the gaseous state.

Conclusions

The MHD generator is ideally suited to the bulk generation of electric power in large unit sizes. With combustion heat sources, there are no fundamental problems remaining in generator operation. Generators for limited duration or demand-type power applications at high level can be made now.

REFERENCES

1. Kantrowitz, A. R., and P. Sporn, *Power*, Nov. 1959.
2. Rosa, R. J., "Physical Principles of Magnetohydrodynamic Power Generation," Avco-Everett Research Lab. Research Rept. 69, Jan. 1960; also, *Physics of Fluids*, vol. 4, Feb. 1961.
3. Brogan, T. R., A. R. Kantrowitz, R. J. Rosa, and Z. J. J. Stekly, "Progress in MHD Power Generation," Avco-Everett Research Lab. AMP 55, Mar. 1961; Mannal, C., and N. Mather, *Engineering Aspects of Magnetohydrodynamics*, New York: Columbia University Press, 1962.
4. Dibelius, N. R., E. A. Luebke, G. J. Mullaney, *Engineering Aspects of Magnetohydrodynamics*, C. Mannal and N. Mather, eds. New York: *Ibid.*, 1962, pp. 307–325.
5. Louis, J. F., J. Lothrop, and T. R. Brogan, "Studies of Fluid Mechanics Using a Large, Combustion-Driven MHD Generator," Avco-Everett Research Lab. Research Rept. 145, Apr. 1963; to be published in *Physics of Fluids*.
6. Novack, M. E. and T. R. Brogan, "Water Cooled Insulating Walls for MHD Generators," presented at 1963 ASME Winter Annual Meeting, Philadelphia, Pa., Nov. 1963.
7. Kunzler, J. E., E. Buehler, F. S. L. Han, and J. H. Wernick, "Superconductivity in Nb_3Sn at High-Current Density in a Magnetic Field," *Phys. Rev. Letters*, Feb. 1961.
8. Berlincourt, T. G., R. R. Hake, and D. H. Leslie, "Superconductivity at High Magnetic Fields and Current Densities in Some Nb–Zr Alloys," *Ibid.*, June 1961.
9. Stekly, Z. J. J., "The Feasibility of Large Superconducting Coils," Avco-Everett Research Lab. Research Rept. RR-160, Sept. 1963.
10. Rosa, R. J., "Shock Wave Spectroscopy—Part I," "Engineering Magnetohydrodynamics—Part II," Thesis, Cornell University, 1956.
11. Kerrebrock, J. L., "Conduction in Gases with Elevated Electron Temperature," *Engineering Aspects of Magnetohydrodynamics*, C. Mannal and N. Mather, eds. New York: Columbia University Press, 1962, p. 327.
12. Rosa, R. J., "The Hall and Ion Slip Effects on a Nonuniform Gas," *Physics of Fluids*, Sept. 1962.
13. Rosa, R. J., "The Application of Magnetohydrodynamic Generators in Nuclear Rocket Propulsion," *ARS Journal 32*, Aug. 1962.

This article is based on a paper presented by the author at the IEEE Winter Power Meeting at New York, February 2–7, 1964.

Graduate education:
a basic national resource

Our economic future is intimately linked with the spreading strata of the highly educated: each Ph.D. can mean employment for 5 to 10 engineers, each of whom uses 10 to 15 skilled workers

L. V. Berkner
Graduate Research Center of the Southwest

As engineers, we are acutely conscious of the impact that engineering activity has on society. Each new device, system, product, or service modifies society in some way, through the necessity for greater adaption of man to his environment.

Now, of course, no one engineering development can be considered profoundly revolutionary in its social impact. But when engineering advances come rapidly, one quickly following another, then the total impact becomes revolutionary—as it has in the past decade. Engineering achievements are effecting, perhaps, the most deep-seated social changes in history. In creating this social revolution, we are producing at the same time a radically different social climate within which our works of engineering must function in the future. In short, this technological revolution in social affairs is reflected back to give us a greatly modified engineering environment.

Factors of social change

Everyone is impressed by the technological advances that have caused this social revolution; it has come so quickly that few yet fully appreciate the true extent of the social changes that are being wrought. Let us try to bring these fundamentals more sharply into focus so that we can grasp the full perspective of the basic social events now bearing on our profession.

First, our population is exploding at the rate of 2 per cent per year. The doubling time for our population was 200 years at the time of the American Revolution; it has dropped quite suddenly in this century to 40 years. As a consequence of modern medical science, the life of the average man now surpasses the Biblical three score and ten years.

Second, quite suddenly, at mid-century, agriculture is now almost completely mechanized and industrialized. With the application of biology, chemistry, and mechanics, agricultural productivity has increased a hundred times. A tiny fraction of our population now produces more than adequate food supplies, leaving great excesses for export. Seventy per cent of the American population were employed in agricultural pursuits in 1860; only 7 per cent were so employed in 1960, and the estimate is 3 per cent by 1980.

Third, the same situation is found in the natural resources industries—fuel and minerals. With the new technology sparked by science, productivity of basic commodities is enormously enhanced.

Fourth, as a consequence of exploding populations, and migration from the land to the cities, our major population growth today centers in approximately 100 major metropolitan areas, which are bursting in size. In Texas, for example, population growth is 25 per cent per decade. Yet of the 254 counties, all but 31 have lost population or have almost stood still in each of the last three decades. Since 1950, more than a half million Texans have left the farming areas. Only nine or ten great metropolitan areas of Texas are getting all of the new and migrant population, thereby increasing their population more than 50 per cent per decade. The example in Texas is typical of the whole nation.

Fifth, the productivity of the industries that support the population of these cities has enormously increased. A powerful technology, sustained by an equally powerful and progressive science, provides machines and computers that operate factories. Ever fewer, but far more skilled, men and women are required to produce larger quantities of superior products and services.

Sixth, with the improved productivity, the personal needs of the educated, skilled, and productive elements of our population are increasingly satisfied. On October 7, 1963, the *New York Times* noted:

> "The family budget standard just issued by the Community Council shows that New Yorkers in the low-to-moderate income brackets are eating more meat, seeing more shows, walking on thicker rugs, and generally living better than they did 8 years ago . . ."

What is true in New York is equally true throughout the country.

At the same time, however, chronic unemployment is increasing. This problem is the inevitable consequence of the quickening flow of population to the cities and the coincident and equally quickening rise in individual productivity as a direct consequence of a sophisticated technology that affects all aspects of our lives.

Basically, then, we are witnessing a technological revolution derived from the powerful science of our time, leading to fast rising productivity and greater wealth of our population, but with a mass movement of people from the farms to flood 100 or so U.S. metropolitan areas—and with an increase of unemployment, focused primarily among the uneducated.

In a sense, at mid-century we moved quite abruptly from the traditional economy of scarcity, which had plagued mankind since the beginning of history, into an economy of plenty. For the first time in history, we have acquired, out of science and technology, the power to produce enough to supply all our needs, with a significant excess remaining.

Throughout history, man had struggled for the bare necessities and for the power to tap the riches of the earth for his development and benefit. Then suddenly, at mid-century he discovered that quick access to the earth's resources had been gained.

Up to a century ago, man's progress was characterized primarily by small improvements in obtaining his basic supplies of food, clothing, and housing. In the present century, as we move toward the new science-oriented economy of plenty, our progress is characterized by radically new and powerful industries that are not directly related to basic necessities. The automobile, the airplane and other fast transportation; accessible communication and mass media; the computer and automatic control; and the nuclear and space industries are all typical examples. These industries have the common characteristic of making the individual far more effective in his environment, rather than supplying the basic necessities of life.

It has become quite clear that the fundamental industries related to basic necessities—food, shelter, agriculture, and basic materials in their traditional form—no longer have the potential for growth that will overcome our increasing unemployment. New industries are required to provide new products and services—and with them, new sources for employment, wealth, and human satisfaction.

Innovation—wellspring of future industry

Here we come to the crux of the situation. New industry can be derived only from innovation. Innovation is the new resource that already provides for more than half of our economy. Today, innovation arises predominantly from an ever-expanding science, under the command of men educated and trained to enlarge that science significantly and to transform it into new and useful products and services.

In this perspective, we see scientific research—supported and conducted on an adequate scale and in a creative atmosphere—and a science-oriented technology as the wellsprings from which the major new industry of the future will arise. Already, the new science-based industry derived from innovation is expanding within our economy at twice the growth rate of the traditional industries of necessity. Already, these science-based industries are the principal source of new employment for the men and women flooding our 100 or so metropolitan areas.

Now, of course, our total production of goods and services cannot forever rise exponentially. But the growth that is provided by our new resource can bridge the transition, in human affairs and in human thought, between an economy of scarcity and an ultimate and stable economy of plenty. In the transition, the growth of productivity should be multiplied many times to satisfy the world's needs.

As a nation, we are now fully committed to reliance on the resource of innovation for our future national growth. It is a characteristic of this resource that it continually releases manpower through the steady rise of productivity of the highly educated and creative individual. Therefore, full employment of our people, with the consequent maximization of total wealth and elimination of poverty, requires the introduction of new industry to provide new products and services that people need and will buy. This implies a steady supply of "innovation." Once we have embarked on the development of this resource, there is no turning back. Consequently we are forced to examine the character of the resource in detail so that we understand it, and can develop it in optimum form.

As I see it, this new resource—innovation based on the scientific understanding and control of nature—has four major elements:

1. Scientific research, to uncover the functional behavior of nature in all its aspects, the general laws that govern this behavior, and the technologies that can command it under man's control and for his benefit.
2. Education, very advanced and continuing—at the graduate level and beyond—to advance knowledge in all areas through scientific research and to infer from new knowledge what innovations will be useful to mankind. In addition, education at all intermediate levels is imperative, both to provide the advanced skills capable of serving in the highly technological industries and to prepare the most competent people for even more advanced responsibilities and opportunities.
3. Energy, derived out of and controlled by the technologies as they emerge. The trend is for controlled energy to replace "labor."
4. Capital, which can direct the major energies of society toward the creation of the mechanized means of production whereby productivity is vastly enhanced and the powers of innovation are fully developed and released.

From this list, we see that the keys to the maintenance of this new resource are scientific research and its companion, advanced education.

Needed: new social and economic doctrines

Out of this new resource, through exploitation of the opportunities offered by human diversity in our free system, we have created a wholly new basis for an economy of plenty. However, since there has been an economy of scarcity for 5 millennia, it will take some time to adjust our thinking to this new economic base, where controlled energy and applied brain power are basic factors.

For instance, we were all taught as youngsters the adage of Adam Smith that "labor is of equal value to the laborer" and that "labor is the basis for all value." In this economy of plenty, where controlled energy and highly developed applied brain power are the principal bases of value, labor contributes only in proportion to its manipulative power over nature. So Adam Smith and Karl Marx, with all the other exponents of the economy of scarcity, can now be relegated to the museum of outmoded ideas acquired along the road of history in the growth of human thought.

Consequently, one suspects that this economy of plenty is optimally governed by quite new social and economic doctrines. These doctrines are being developed experimentally out of our free system of diversity. Unfortunately, the significance of the profound switch from the old economy of scarcity to the new economy of plenty does not appear to have been grasped by the mainstream of the social and economic sciences. Steeped in the doctrines of the past, and not familiar with the detailed technical processes whereby engineering is now influencing society so strikingly, the social and economic scientists have largely failed to grasp the full meaning of revolutionary social changes we are witnessing.

For instance, one of the radical consequences of this new "economy of plenty" on which we embarked at mid-century is the sharp realignment in the national proportions of our skills. Previously, education was a convenient stepping stone to success, but now it becomes a central contributor to value. In particular, graduate education to the doctor's degree and beyond—not long ago a luxury—now becomes a vital necessity to the success of the new economy and any society depending on it.

To illustrate this point, in 1950 a total of about 100 Ph.D.s were employed in the Dallas–Fort Worth metropolitan area—mostly as teachers and professors. Today, just over a decade later, this number has grown to 1200. It is increasing by 200 per year. This one great metropolitan area has grown rapidly to nearly 2 million with Dallas County alone increasing by nearly 60 000 per year. The new industry in the area is almost wholly science based, requiring approximately 1 science-trained Ph.D. to provide opportunity for each 115 others employed. We can see the immense problem in advanced education that lies ahead when it is noted that the graduate output of all Texas universities cannot now meet the needs of even this one of ten Texas metropolitan areas.

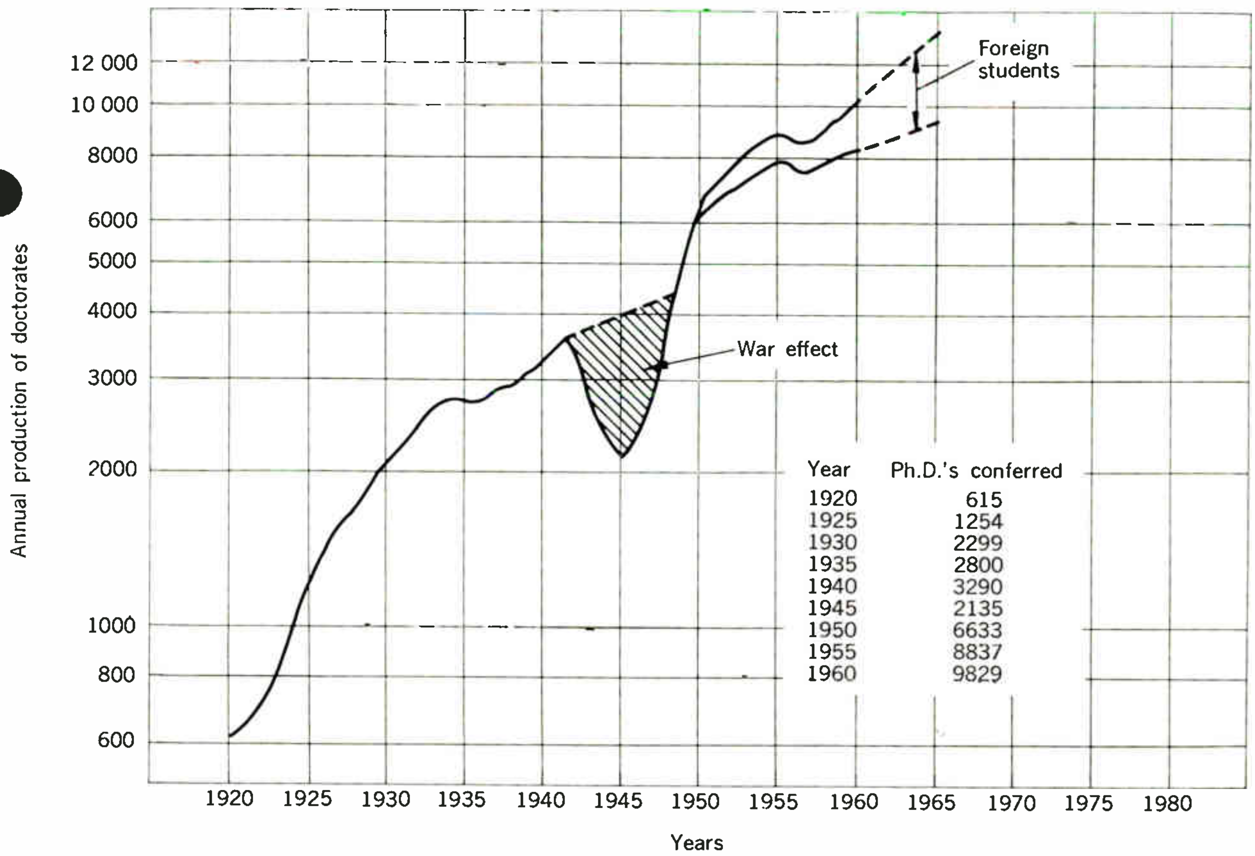
No longer can we avoid facing squarely the overpowering fact that modern science, and the technology spawned therefrom is created by the brain power of men and women trained to the fullest extent of their capabilities.

To capture and fully exploit the potential of the complex technology of today, to create industry from it, and to direct it for our national benefit, men must be found who understand and can manipulate the fundamental science from which it is generated. The education that suited the needs of the closing generation of engineers in which most of us functioned, can no longer command the ideas that will underlie engineering innovation of tomorrow. The days of the Edisons and the Fords who could develop inventions from a synthesis of simple technologies are being outdistanced by the power of wave mechanics, stochastic processes, and advanced circuit theory. The motor car and the telephone of today are already a far cry from the concepts of their inventors. Yet as engineers we are acutely conscious of the certainty that we have only cracked the surface of a vast science-based technology that must emerge in the years ahead.

Advanced education—an imperative

Obviously, the ideas underlying the technology of the future are abstruse and highly mathematical in nature. To comprehend these ideas, and to manipulate intelligently the technology born of them requires men of very advanced education in substantial numbers. This does not mean a mere four-year college education, for the ideas involved are at the very boundaries of knowledge. Command of the new technology and of the science from which it is derived requires postgraduate education to the doctor's degree and beyond—not less than eight years beyond the high school diploma. In particular, it may well require continued education at the graduate level, concurrent with productive endeavors, throughout a lifetime.

Of course, men of lesser training are essential and can be usefully employed in technological industry. But



Semilogarithmic plot of doctorate production in all fields in United States universities from 1920 on. Particularly noticeable decreases in growth rate are apparent in the depression years of the 1930s and during the slump following the recovery from World War II.

each Ph.D. means employment for 5 to 10 engineers, and each engineer uses 10 to 15 skilled workers. The creation of new industry, new products and devices, new methods and applications from the new technology arises from the creative and imaginative insights of scientific and technological leaders who have access to the very limits of knowledge. Without the presence of innovating skill, men of lesser skills will lose their opportunity. Therefore, we cannot discuss the development of the new economy without discussing the educational needs of the men and women who must provide the top brainpower to carry out our national research program.

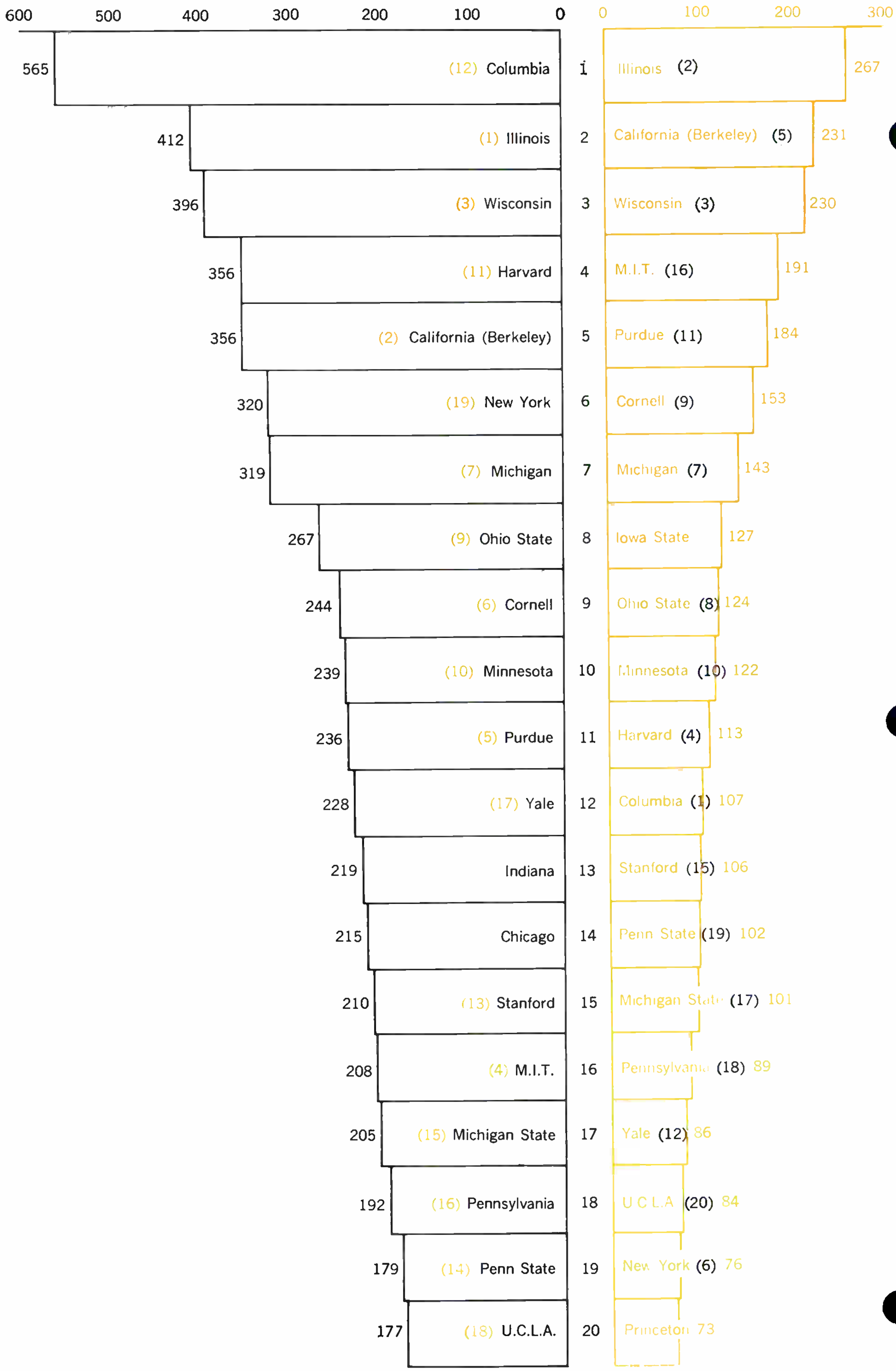
To put the situation bluntly: no training of numbers at the trade school, high school, or college level can in itself capture the new technology. Indeed, in the future we may have to count a hundred or more unemployed for each Ph.D. we fail to educate. The key to the new technology is derived from the boundaries of knowledge—from training at the doctoral level and beyond. Those communities that can produce and retain men of advanced education will have the most intimate access to and control of the new technology from which the industry and wealth of the future will flow. They will have the power to create opportunity for full employment at all levels.

This is why extensive doctoral and postdoctoral education in the sciences and engineering has suddenly become imperative. This is why the new science-oriented industry is suddenly exploding around the great centers of advanced education in Boston; New York; San Francisco and Palo Alto; Southern California; and even Minneapolis–St. Paul. For the science-based industry must have access, not merely to highly trained men in their own organizations, but also to the sources of education themselves—the universities and the great fundamental research laboratories where the secrets of nature are being explored to their very limits.

Gradually realizing these facts, our whole nation is suddenly awakening to the imperative need for training our men and women to advanced levels in science and technology; to building up our graduate schools in every major metropolitan area; to creating opportunities for postdoctoral training, and really advanced research so that industrial growth can remain adequate. Since the agricultural and basic resource areas of the nation can no longer provide the traditional professional and vocational opportunities on an adequate scale, the people in these areas have been forced to turn to innovation and consequently to the reassessment of their educational facilities in which that innovation is rooted.

Our graduate facilities

Where does our nation stand in the face of this revolutionary social and economic change? We produce a few more than 11 000 Ph.D.s annually (not including law


600 500 400 300 200 100 0
0 100 200 300
565 (12) Columbia 1 Illinois (2) 267
412 (1) Illinois 2 California (Berkeley) (5) 231
396 (3) Wisconsin 3 Wisconsin (3) 230
356 (11) Harvard 4 M.I.T. (16) 191
356 (2) California (Berkeley) 5 Purdue (11) 184
320 (19) New York 6 Cornell (9) 153
319 (7) Michigan 7 Michigan (7) 143
267 (9) Ohio State 8 Iowa State 127
244 (6) Cornell 9 Ohio State (8) 124
239 (10) Minnesota 10 Minnesota (10) 122
236 (5) Purdue 11 Harvard (4) 113
228 (17) Yale 12 Columbia (1) 107
219 Indiana 13 Stanford (15) 106
215 Chicago 14 Penn State (19) 102
210 (13) Stanford 15 Michigan State (17) 101
208 (4) M.I.T. 16 Pennsylvania (18) 89
205 (15) Michigan State 17 Yale (12) 86
192 (16) Pennsylvania 18 U.C.L.A. (20) 84
179 (14) Penn State 19 New York (6) 76
177 (18) U.C.L.A. 20 Princeton 73

Leading doctorate-producing institutions in the United States for the academic year 1960–1961. On the left in black, doctorates granted in all fields. On the right in color, doctorates granted in the natural sciences only (including mathematics, physics, chemistry, geo-sciences, engineering, and biological sciences). On the left, the numerals in color refer to the rank of the institution in the granting of science doctorates; on the right, the numerals in black refer to the rank of the institution in the granting of doctorates in all fields of endeavor.

and medicine), about half of whom graduate in science and technology.[1] In the 1960 era, approximately 55 per cent of this group were graduates of 20 great universities. In all, since 1940, 25 universities have accounted for two thirds of our Ph.D. crop. Their facilities are now approaching saturation; and new major graduate universities have yet to emerge in the United States. As a consequence, doctoral graduates are now increasing more slowly than undergraduates. Each of the 20 major graduate universities has more than 35 per cent of its student body in the graduate schools. The remaining 80 of the leading 100 universities in the United States are undergraduate-oriented, usually with less than 12 per cent of their students in graduate schools.

Ten states (containing 18 of the major 20 universities) graduate an average of 100 Ph.D. candidates annually per million of their population. These are the northeastern, the mid-northern, and the far western states. The remaining 40 states graduate from a maximum of 50 down to zero doctoral graduates per million population per year.

Nor do most students travel very far to the great universities. Last year, less than 1000 traveled more than 500 miles to earn their doctorates. In the vicinity of a great graduate school, about 15 students per 1000 high-school graduates receive their doctorates. At greater distances, less than five per 1000 achieve the Ph.D. degree.[2] This does not mean that a student always goes to the nearest graduate school—it does mean that the great graduate schools exercise a powerful intellectual motivation on their communities.

We have spoken so far of total Ph.D. production. Of particular interest is that of the top 20 universities, 18 are also leading producers of Ph.D.s in science and engineering. Moreover, these 20 universities have more than 55 per cent of their doctoral graduates in science and technology compared with an average of less than 45 per cent in the other leading universities. Of great significance, at the bachelors level, only 21 per cent of graduates of U.S. universities are trained in science and engineering compared with 46 per cent in Great Britain. The reputed overemphasis of U.S. education in science is sheer nonsense.

It is clearly evident that just at the time we desperately need this new resource—innovation—we are falling flat on our faces in the race to get it. What is needed for our new economy is at least one major graduate institution in each of our 100 major metropolitan areas. Each such institution should produce not less than 200 Ph.D. graduates annually. These institutions should bring the over-all level of intellectual excellence in the United States to at least 100 doctoral candidates per one million population annually.

Associated with these institutions must be the facilities and faculties for basic research that provide access by both scholarship and industry to the whole range of advanced scientific and technologic thinking of our time. In stimulating doctoral and postdoctoral levels of attainment, scientific research and graduate education are inextricably entwined and closely identified in many areas. Maximum use will not be made of our expanded graduate facilities unless:

1. The faculty actively participate in research; graduate teaching is utterly sterile otherwise. The whole attitude of the student is radically different under the intellectual leader whose thinking is oriented by his search for knowledge. Without research, teachers at the graduate levels become stereotyped and quickly outmoded.

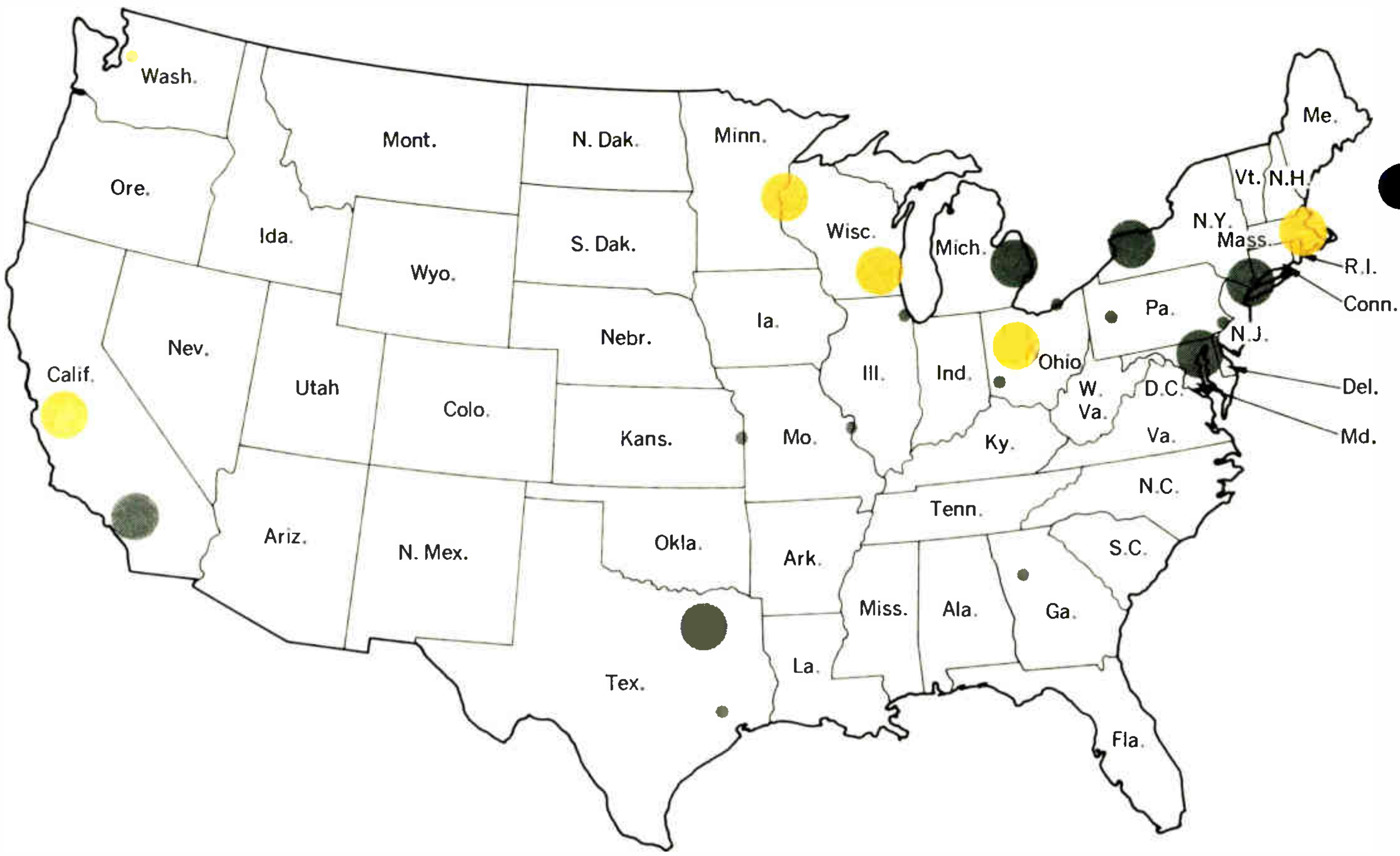
2. A graduate student is trained in the methods and procedures of scientific and engineering research by actual participation under great scientific and technologic leaders. He can no more become a scientist by merely reading books than can an airplane pilot learn to fly while sitting on the ground.

3. Fresh new minds come into early contact with the problems that nature presents. Some of the most significant new insights into scientific problems are gleaned by youthful scholars in an inquisitive academic environment.

So in the building of graduate opportunity for our new economy, we are faced with the enlargement of our research plant. Less than half of our population, and of our great future metropolitan areas, are served by graduate institutions of the requisite size or distinction. This educational lack is a costly liability in our national economy—a liability that our nation cannot long afford. Some critical questions need to be asked of universities that pretend to provide the intellectual strength for our unserved communities.

At the same time, the Congress could hardly select a more inauspicious time to reduce the national research expenditures. I have heard quoted the figure of $16 billion for United States research and development. Less than 10 per cent of this figure—a little over $1 billion—goes for true research. The remainder is involved in hardware in one way or another.

That $1 billion is the seed from which our enlarged economy must grow. Let us not confuse the massive, though necessary, engineering activities of our government with the much smaller but imperative needs for basic scientific research from which alone the training and the innovation of the future can be derived. The serious cuts last year in the basic research budget of the National Science Foundation and other agencies supporting basic science will be reflected, inevitably, in losses to our national economy, and, it should be noted, in more than equal losses to our Treasury in tax returns and payments to the relief rolls. Experience in the growth of our economy now tells us this unambiguously. Consequently, our expenditures for basic research should be multiplied twice again over the next decade to provide for the extension of research opportunity in doctoral and postdoctoral education for the unserved elements of our population. One can conclude only that in cutting expenditures for scientific research at this time our legislators are not clearly aware of the dominating elements on which the health of our new economy is founded.



On the right, metropolitan areas in the United States with one million or more population (U.S. Census for 1960) arranged in order by number of doctorates granted (academic year 1960–1961) in all fields per million population. Only those metropolitan areas shown in color meet the requirement set forth in the text of at least one institution per major metropolitan area producing at least 100 doctorates per one million population annually. The map above shows the geographical locations of the metropolitan centers in the bar chart. The same color-coding applies.

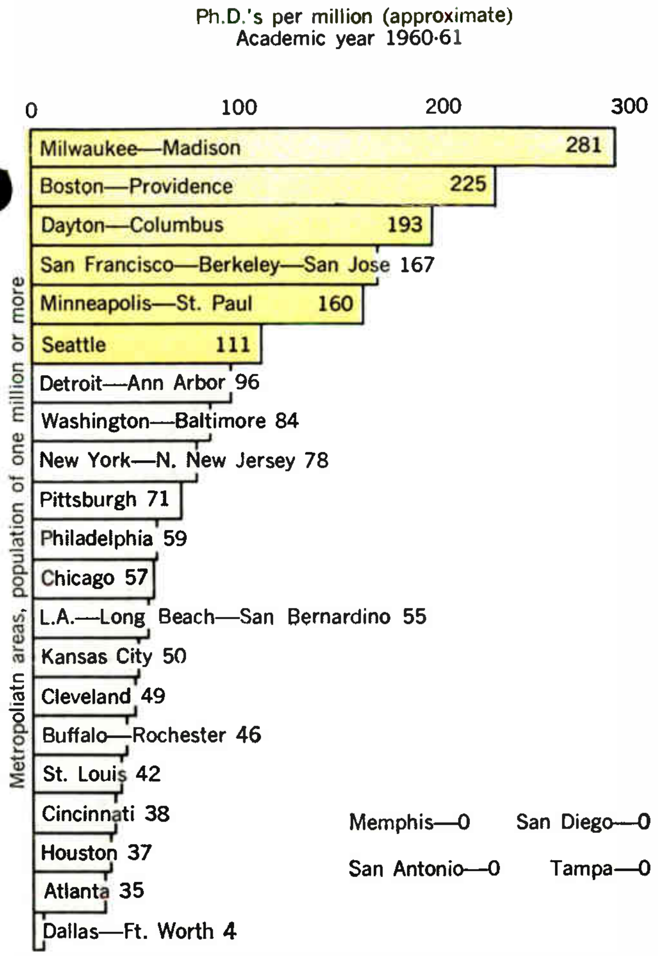
Some serious questions

In the light of the demands of our newly oriented economy, and the inevitable social problems that arise, together with the great social advantages that can be acquired, our national policies must be thought out more clearly with respect to science and engineering. Basically, it is agreed that our system of diversity, growing out of free enterprise, permits the healthiest evolution of social and economic objectives through unhampered natural selection in a free society. Within this framework, some really serious questions must be answered if adequate policies are to be developed.

How can we evolve a system of graduate education designed to meet the needs of the new economy in every great metropolitan center? To supply the needs for competent graduate teaching, how can a major proportion of our scientific and engineering community be brought into the teaching and counseling role in their areas of competence? How can research funds be injected into situations where advanced scholarship is yet undeveloped without encouraging and perpetuating an inferior brand of scholarship? How can continuing graduate training be provided for, and assimilated by our whole scientific and engineering community without the necessity of imposing undue strain on the individual?

What levels of funding for basic and applied research are essential to optimization of our new economy, in support of education and in creation of foundations for adequate innovation? What are the rules and procedures for identification of justifiable research objectives, both to meet the needs of intermediate educational training, and the ultimate needs of genuine advancement of knowledge? How can we acquire facilities and equipment quickly and effectively in intellectually underdeveloped centers of population? How can the public be assured that its tax monies for research and education are effectively employed without destroying the atmosphere of creative thought which it is their purpose to stimulate?

These are but a few of the questions to be intelligently answered if the national needs in our new economy are to be fulfilled. It seems doubtful that the answers will apply to all metropolitan centers with their diversified needs; consequently, the leaders of each metropolis cannot escape asking such questions, and finding answers for themselves. Above all, suitable policies cannot be evolved without intimate knowledge of the creative research process itself. So our scientific and engineering leaders in each center of population cannot escape thinking



through and analyzing the social problems that their works now create.

While the extraordinary evolution of our economy necessarily preoccupies our attention by the serious nature of the problems to be solved, we must not lose sight of the ultimate goals of any sound society. The advancement of knowledge unquestionably lifts civilization in the cultural sense. The incredible brutality of the past, so aptly described by the late Hans Zinsser[3] is no longer tolerated in our age when knowledge with its consequent justice has lifted civilization above the dregs of animal-like existence. Knowledge provides an ever higher richness of living within which man can more deeply enjoy and appreciate the opportunities that his scientific culture affords.

REFERENCE

1. "Doctorate Production in United States Universities, 1920–62," National Academy of Sciences–National Research Council, Washington, D.C., publication 1142, 1963.
2. "Manpower Study no. 3," National Research Council, Washington, D.C., 1961.
3. Zinsser, Hans, *Rats, Lice, and History*, New York: Little, Brown & Co., 1935.

OTHER DOCTORATE TRENDS

In his paper, Dr. Berkner makes reference to a recent publication of the National Academy of Sciences titled "Doctorate Production in United States Universities, 1920–1962." The following data are taken from this report and provide some additional information about trends in doctorate production in United States universities.

☐ Within the physical sciences (which include engineering and mathematics), engineering replaced chemistry in 1962 as the output leader on the strength of a three-fold increase since 1950 while chemistry was maintaining a nearly constant level.

☐ The doctorate records show that the average time lapse between baccalaureate and doctorate degrees has increased slightly in the last decade. This is true in spite of the massive growth of Federal fellowship programs which are designed to decrease the time lapse, and which have been found to result in a significant acceleration for those obtaining awards.

The time between the two degrees varies by field: the shortest at present, in the physical sciences, is 7.8 years, with the average ranging upward for the bio-sciences, social sciences, and arts and professions, to a peak of 15.2 years in the field of education.

The median age of doctorate recipients follows the same order, varying from 28.7 years in the physical sciences to 37.6 years in education, based on data for 1961.

☐ In all areas of graduate study, the influx of foreign students is apparent; the percentage of foreign-born doctorates increased slowly until 1950, but jumped from a total of 2000 in 1940–49 to 8000 for the period 1950–59. The latter figure compares with an over-all output of 83 000 during the decade of the 1950s.

The proportions of foreign-born doctorates are highest at present in the biological sciences at 17.4 per cent and the physical sciences at 14.7 per cent.

☐ The rapid growth of graduate education has not brought any radical change in the relative Ph.D. productivity of major fields of study. In spite of substantial year-to-year variation, the only persistent trend over the 40-year period from 1920 to 1962 is an increase in the field of education and a compensating decline in the arts and professions.

☐ Since 1950, the percentage of women receiving doctorates in the social sciences, arts and professions, and education has gradually increased, while the low levels reached in 1950 are continuing in the physical and biological sciences.

From 1920 until 1950, the percentage of women doctorate recipients underwent a general decline to about 3 per cent in the physical sciences, 9 per cent in the biological and social sciences, 15 per cent in the arts and professions, and 17 per cent in education.

The artificial intelligentsia

The work, working, and workers in artificial intelligence are illuminated, magnified, and dissected by the occasionally sharp instrument of the Government contractual process

Louis Fein *Consultant*

Thinking machines, translating machines, learning machines, self-organizing systems, conditional probability machines, teaching machines, adaptive machines, self-reproducing machines, man-machine symbiosis, intelligence amplifiers, artificial intelligence, heuristics, list processing, pattern recognition, simulation of cognitive processes, brain modeling, cybernetics, synnoetics, concept formation, bionics, and scores of other such terms may evoke awe, despair, confusion, charges of fraud and charlatanry, hope for the world, fear of abnegation of man's control of his own affairs, hope for a leisure society, fear of unemployment, etc. What's the fuss about?

Is there substance in this work? Does it have promise or is it merely pretense? What problems are being addressed? Can they be adequately formulated? What procedures are extant to solve these problems? Are they reliable and valid? On what theoretical foundations and empirical evidence does this work rest? How do the various schools of thought differ? How are they similar? What are the outstanding theoretical and practical problems in the field?

The purpose of this paper is to analyze these questions and, in particular, to discern and exhibit prominently what is known and demonstrated, as opposed to what is conjectured and claimed, in the field popularly called artificial intelligence.

I will ask you to imagine that a Federal agency has sent a "request to bid" on some work in artificial intelligence to four previously qualified bidders, each institution representing a contending school of thought. Each bidder writes a proposal. All proposals are evaluated by a competent and skeptical technical expert who represents and advises the agency.

This review consists of six sections: a fictitious request to bid, four proposals, and a technical evaluation. The request to bid defines the nature of the problems to which artificial intelligences address themselves. The four proposals give theoretical foundations, empirical results, problem formulations, and distinctive approaches to problem solution. The evaluation, by comparing and contrasting the proposals and their claims, provides a critique and a prognostication of the potential of current work. (The arbitrary selection of four companies to represent the various schools of thought is itself an act of evaluation.)

A by-product of this form of presentation is the opportunity to examine the assumptions, values, attitudes, and predilections of practitioners, their employers (proposal writers), and their sponsors (proposal recipients).

This review of current knowledge and technical details of artificial intelligence is as trustworthy and comprehensive as the author could make it in a paper of this length. It is thus intended to inform and hopefully to do so entertainingly. To lighten what otherwise might be a dry presentation, pseudonyms have been used whose significance many readers may appreciate. The humorous mode of presentation is a way of taking up attitudes and engaging serious topics in a diverting context. It should be unnecessary to assert that no offense is intended, since attitudes are taken up, not persons.

HEADQUARTERS

WHYONYX LABORATORY

BRIGHT FIELD

Reply to Attn of: SMARTIES
SUBJECT: REQUEST FOR BID NUMBER
123(456)-ABC-789-D
TO: PROSPECTIVE CONTRACTOR

8 January, 1964

1. Your organization is invited to submit a proposal covering the whole program or any part of it.

2. If your proposal is favorably considered, a survey team may visit your facility to determine your ability to perform.

3. Reply to Whyonyx Laboratory, Att: SMARTIES, Bright Field.

Exhibit A. Statement of Problem(s)

To perform many of its tasks well and to carry out many of its obligations, the Federal Government depends heavily upon human intelligence. The problems of providing sufficient aid to human intelligence for coping with increasingly difficult domestic and international tasks cannot be solved readily with conventional, contemporary computers and programs.

Former contracts have reviewed the status of the artificial intelligence field and have delineated the limitations of machines and programs. However, if a method could be found which would provide intelligence, bypassing humans altogether, revolutionary improvements could be made in our domestic and international situations. The question is: are there undiscovered machines and programs which can provide at least as much as -- and preferably more than -- the comprehension, adaptability, versatility, and intelligence of persons? More particularly, can we build machines or programs to:

1. Comprehend live voice and handwritten messages.
2. Prove theorems in mathematics and logic.
3. Identify targets, recognize sub-atomic particles from bubble chamber tracks.
4. Form concepts, generate hypotheses, make inductive inferences.
5. Translate printed French, English, Russian, German, Chinese, and Japanese.
6. Predict weather, make medical or psychiatric prognoses, explore space, explain the space environment, and make spatial predictions.
7. Act as a reference librarian.
8. Act as a mathematical assistant.
9. Prescribe medical treatment.
10. Write and debug computer programs.
11. Design intelligent machines.
12. Memorize data and answer questions.
13. Play games with incomplete strategies.
14. Imitate or emulate a military commander, a government head, a business corporation, a nation's economy.

Can machines and programs improve these capabilities with experience?

The answers to these questions are the objectives of this program.

Exhibit B. Objective

One of the aims of this program is to re-explore the human brain and nervous system, the anatomical characteristics of thinking, and human intelligent behavior. But the ultimate objective is to provide machines and programs which will obviate the requirement for people to do intelligent things.

Exhibit C. Areas of Consideration

Within the framework of current knowledge of artificial intelligence, every potential possibility will be considered. These considerations should provide some clues to the new areas of knowledge which must be interpolated. A diligent search for these clues will be a major emphasis of this effort.

Calculated Risks

CLUSTERVILLE, N-SPACE, INC.

To: Bright Field
Attn: SMARTIES
Subject: Proposal in Response to RFB No. 123 (456)-ABC-789-D

January 17, 1964

Dear Sir:

Calculated Risks, Inc., is probably the oldest, and certainly the quietest, organization currently in the field of artificial intelligence. We, therefore, wish to congratulate you on your filtering and discriminating ability in recognizing us as a potential contributor in the presence of so much noise.

Our claim to seniority is based upon our founder, Bayz N. Gouse, who generated the theoretical foundations for the operation of our product lines more than 150 years ago.

Our most profitable original machines are used to provide amusement to persons who enjoy playing games of chance. Simultaneously, the owners of these machines are guaranteed a certain average gain. The Las Vegas entrepreneurs will testify that our equipment returns to them the profit we predicted.

While our product line that emulates human intelligence operates on the principle of minimizing average loss (as opposed to maximizing average gain as in our games-of-chance machines), the devices are based on the same solid and tested theoretical foundations of probability and statistics. We trust that this long, successful, theoretical, and practical record will be given appropriate weight when you consider contractual awards.

Our artificial intelligence product line is intended primarily for classifying objects or patterns already represented by distinguishing features. We name our products for prominent persons whose contributions to science and society are commemorated thereby.

John Hancock. If the distinguishing features of the set of objects or patterns to be classified are represented by binary variables, and if they are statistically *independent*, then John Hancock will make optimum categorizations, i.e., it will elect the most likely category—the one that will minimize the average error (or loss). If there are but two categories of choice, Hancock is, of course, a dichotomizer. Indeed, it is a linear input threshold element!

President Lincoln. If the input variables are not statistically independent, and one wishes to render them independent (emancipate them so to speak), Calculated Risks, Inc. occasionally builds a device, labelled President Lincoln which, in tandem with Hancock, will make optimum categorizations of objects with nonindependent variables.

Bayes. Another way of assuring optimum categorization of objects with nonindependent binary inputs is to add new weight(s) to the inputs of the Hancock which represent the appropriate joint conditional probabilities. This module is called Bayes.

Socrates. Finally, a teaching module, called Socrates, is available to teach John Hancock the optimal weightings after each categorization. One of Socrates' pedagogies is the following:[1] Whenever a categorization is made, appropriate weights are incremented and the remainder are decremented. In particular, all weights w_i whose corresponding input $x_i = 1$ are replaced by $(w_i + 1)\delta$; the remaining w_i's are replaced by $w_i\delta$. If δ is close to 1, Hancock learns slowly and reliably; for δ small, learning is fast, but unreliable.

Mr. Gauss. Mr. Gauss is our optimum decision network for Gaussian patterns of continuous variables which may or may not be statistically independent. The five main components of Gauss and their computing roles are as follows:

Component	**Role**
Whitening filter	$Z_d = U_d^{-1/2}\mathbf{X}$
Matched filter	$\mathbf{Z}_d' U^{-1/2}\mathbf{\Lambda}_d$
Energy computer	$\mathbf{Z}_d{}^2$
Bias	$b(d) = \log p_d - \frac{M}{2}\log 2\pi - \log \lvert U_d \rvert - \frac{1}{2}\mathbf{\Lambda}_d' U_d^{-1}\mathbf{\Lambda}_d$
Summing device	$f(d) = b(d) + \mathbf{Z}_d' U_d - \frac{1}{2}A_d - \frac{1}{2}\mathbf{Z}_d$

Dr. Freud. Calculated Risks, Inc., recognizes the importance of making optimum decisions on abnormally distributed patterns. Our research is directed toward remaking an abnormal into a normal distribution for Gauss to operate upon. Accordingly, we call this future product Dr. Freud. If we are successful with this research, we shall suggest to our customers that Freud and Gauss be used in tandem to make optimum decisions for those patterns whose normality is in doubt.

We suggest that for target, voice and handwriting recognition applications, medical diagnosis, weather prediction, etc., the Calculated Risks, Inc. product line is eminently suitable.

Stochastically yours,

Abe Normal

Abe Normal, Pres.
Calculated Risks, Inc.

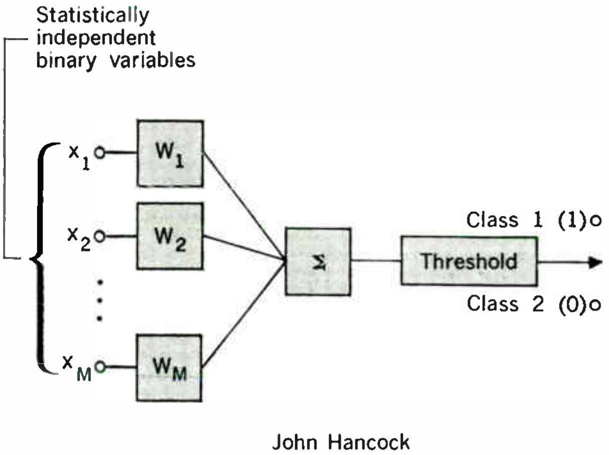
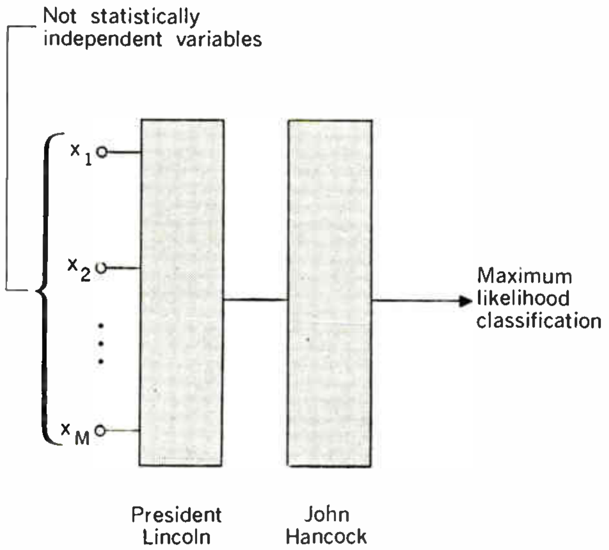
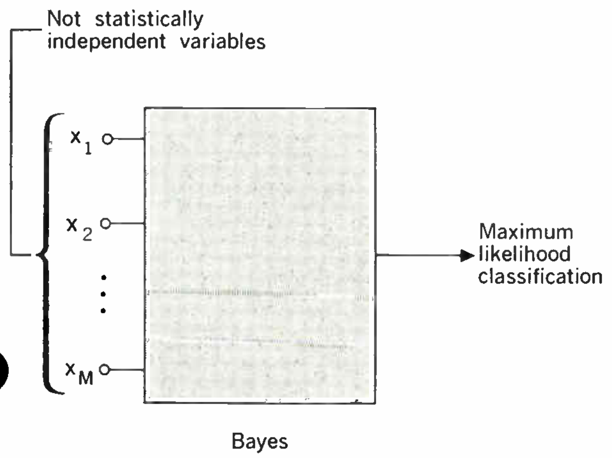


Diagram of input elements and threshold device of the John Hancock categorizer (produced by Calculated Risks).

Sketch of Calculated Risk's President Lincoln and John Hancock interconnected for optimum categorizations.



Line diagram of the Bayes, a variation module of the Hancock that serves to add new weights to the binary inputs.



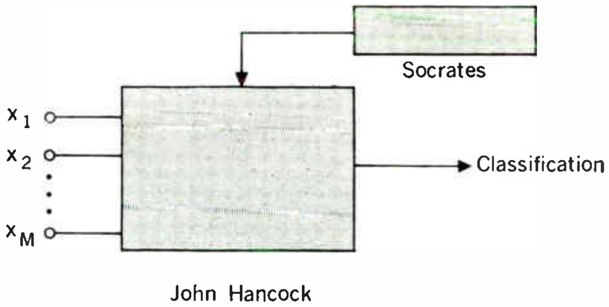
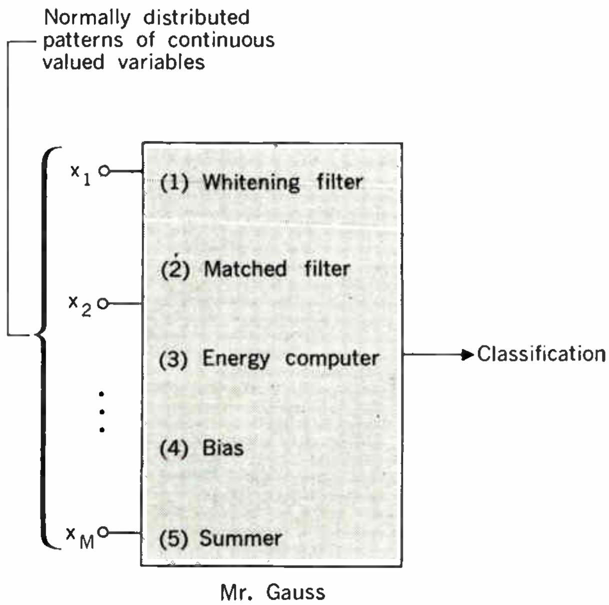
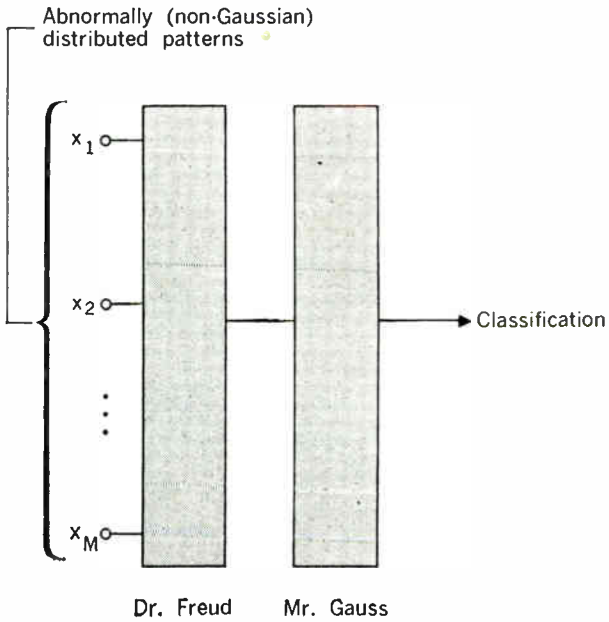


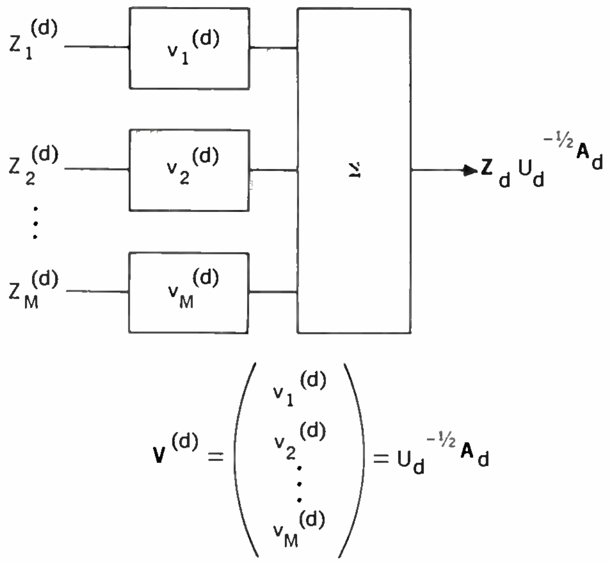
Diagram shows the teaching module, Socrates, cross-coupled to the "learner" John Hancock unit.



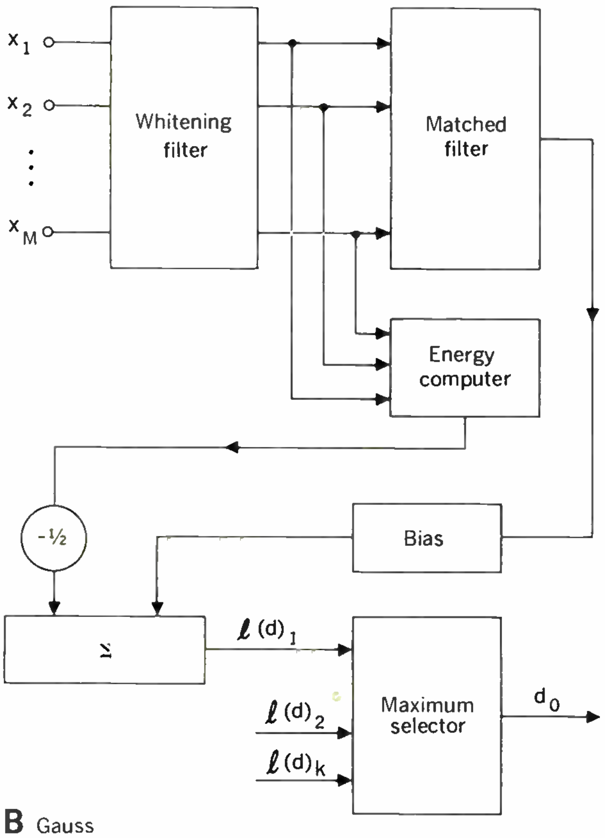
Sketch of Calculated Risk's Mr. Gauss, the optimum decision network for Gaussian patterns of continuous variables.

Diagram showing a theoretical combination of the proposed Dr. Freud unit with the existing Mr. Gauss unit.

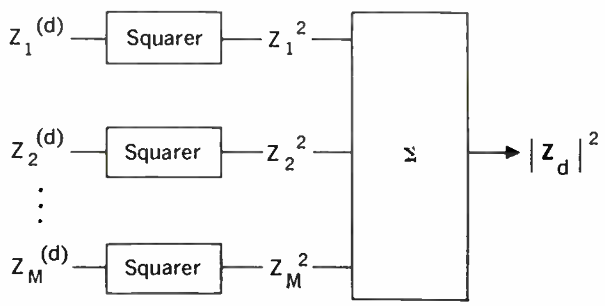




A Filter matched to the whitened mean



B Gauss



C Energy processor

APPENDIX

The Theory of Operation of Gauss*

Let $\mathbf{X}$ = the input vector of components $x_1, x_2, \ldots, x_M$
Let K = the number of possible categories
Let $p(c)$ = the known *a priori* probability for each category

$c = 1, 2, \ldots, K$; write $p(c) = p_c$

Let $p(\mathbf{X}/c)$ = the known (Gaussian) probability for $c = 1, 2, \ldots, K$.

Define K mean vectors $\mathbf{A}_c$ for $c = 1 \ldots K$, and K covariance matrices U_c for $c = 1 \ldots K$ such that

$$p(\mathbf{X}/c) = \frac{1}{(2\pi)^{M/2} |U_c|^{1/2}} \exp - \tfrac{1}{2}\{(\mathbf{X} - \mathbf{A}_c)^t U_c^{-1} (\mathbf{X}-\mathbf{A}_c)\}$$

for $c = 1 \ldots K$.

* Adapted from unpublished notes of N. Nilsson.

For a symmetric loss function, the optimum decision network computes the quantity $p(\mathbf{X}/c = d)p_d$ for each $d = 1 \ldots K$; then it chooses that category d_0 which corresponds to the largest of these K computed quantities.

Let $\ell(d) \triangleq \log [p(\mathbf{X}/c = d)p_d]$. Note that $\ell(d_0) \geq \ell(d)$ for all d so that the optimum decision network could as readily compute $\ell(d)$ for $d = 1 \ldots K$.

Because $p(\mathbf{X}/c)$ is Gaussian,

$$\ell(d) = \log p_d - \tfrac{M}{2} \log 2\pi - \tfrac{1}{2} \log |U_d| - \tfrac{1}{2}\{(\mathbf{X} - \mathbf{A}_d)^t U_d^{-1} (\mathbf{X} - \mathbf{A}_d)\}$$

The value of $-\frac{M}{2} \log 2\pi$ is the same for all $d = 1 \ldots K$, so it can be disregarded in the calculation of $\ell(d)$. The expression: $\log p[- \frac{1}{2} \log |U_d|$ does not depend on the particular input vector $\mathbf{X}$ to be categorized and it can, therefore, be thought of as a bias $b(d)$. Redefine $\ell(d)$ and $\frac{M}{2} \log 2\pi$, then

$$\ell(d) = b(d) - \tfrac{1}{2}(\mathbf{X} - \mathbf{A}_d)^t U_d^{-1}(\mathbf{X} - \mathbf{A}_d)$$

The transformation $\mathbf{Z}_d = U_d^{-1/2}\mathbf{X}$ allows us to write

$$(\mathbf{X} - \mathbf{A}_d)^t U_d^{-1}(\mathbf{X} - \mathbf{A}_d) = (\mathbf{Z}_d - U_d^{-1/2}\mathbf{A}_d)^t (\mathbf{Z}_d - U_d^{-1/2} \mathbf{A}_d)$$

Expanding this scalar product yields

$$(\mathbf{X} - \mathbf{A}_d)^t U_d^{-1} (\mathbf{X} - \mathbf{A}_d) = |\mathbf{Z}_d|^2 - 2\mathbf{Z}_d^t U_d^{-1/2} \mathbf{A}_d + \mathbf{A}_d^t U_d^{-1} \mathbf{A}^p$$

Therefore

$$\ell(d) = b'(d) - \tfrac{1}{2} \mathbf{A}_d^t U_d^{-1} \mathbf{A}_d + \mathbf{Z}_d^t U_d^{-1/2} \mathbf{A}_d - \tfrac{1}{2}|\mathbf{Z}_d|^2$$

The quantity $-\frac{1}{2} \mathbf{A}_d^t U_d^{-1} \mathbf{A}_d$ does not involve $\mathbf{X}$ so it may be considered as a bias. Thus,

$$\ell(d) = b(d) + \mathbf{Z}_d^t U_d^{-1/2} \mathbf{A}_d - \tfrac{1}{2} |\mathbf{Z}_d|^2$$

Two terms in this expression depend on $\mathbf{X}$. The first, $\mathbf{Z}_d^t U_d^{-1/2} \mathbf{A}_d$ is a linear combination of the components of the whitened input vector and can be obtained as the output of the processor shown in the accompanying illustration; the weights $\omega_i^{(d)}$ for $i = 1 \ldots M$ are the M components of the transformed mean vector $U_d^{-1/2} A_d$. Such a processor is sometimes called a correlator or matched filter. The output is the response of a filter *matched* to the whitened mean when the input is the whitened input vector.

The second term which depends on $\mathbf{X}$ is $-\frac{1}{2} |\mathbf{Z}_d| \cdot |\mathbf{Z}_d|$ is sometimes called the *energy* of $\mathbf{Z}_d$, and can be computed by the energy processor depicted.

REFERENCE

1. Minsky, M., "Steps Toward Artificial Intelligence," *Proc. IRE*, vol. 49, Jan. 1961, pp. 14–15.

Diagrams of whitening filter (A) and energy processor (B) units as individual components, and their combination (C) to simulate Gaussian optimum decision networks.

DANDYLINES / ENTERPRISES

To: Bright Field
Attn: SMARTIES
Subject: Proposal in Response to RFB No. 123 (456)-ABC-789-D

January 21, 1964

Dear Sir:

Dandylines Enterprises appreciates the opportunity to submit a proposal. We propose to do some, not all, of the important tasks identified in Exhibit "A" of subject RFB. In all modesty, the few tasks on which we have chosen not to bid are those that our competitors can probably do better. But those on which we have chosen to bid are the items about which we feel confident; indeed, that we can do better than our competition by a comfortable margin. Dandylines trusts that you will concur.

We represent a unique attitude, capability, and school of thought. We suggest that certain problems in artificial intelligence be solved by concurrently operating combinational circuits built of linear input threshold elements.

Dandylines product line

Our product line functions as a recognizer or classifier of objects having M distinguishing features on the basis of which an object is placed into one of 2^N categories. Our M-input, one-binary output Dandyline, $N = 1$ (a single linear input threshold element), is our only shelf item; all other M-to-one lines and M-to-N lines are built to satisfy customer's requirements.

An M-to-one line, called an M1 line, is a summing device having M weighted inputs, $x_1, x_2, \ldots, x_M$, and a threshold, such that its output is one, if

$$\sum_{i=1}^{M} x_i W_i \Sigma \geqq 0;$$

otherwise, it is zero.

An M1 line is often built as an MK1 line in tandem with a K1 line, with both lines having adjustable weights, or one line having adjustable weights and the other fixed weights.

Sometimes, M1 lines are built as MKQ1 lines, and even as MKQR1 lines with different combinations of fixed and variable layers.

MN lines are built in a variety of ways. The most popular versions are built of N identical M1 lines in parallel. Sometimes, MN lines are built of a number N of M1 lines of different kinds.

M1 Dandylines

Principles and theory of operation. There are 2^{2^M} Boolean functions of M binary variables. One linear

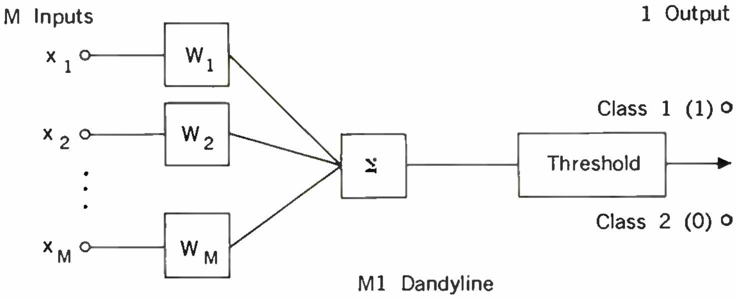
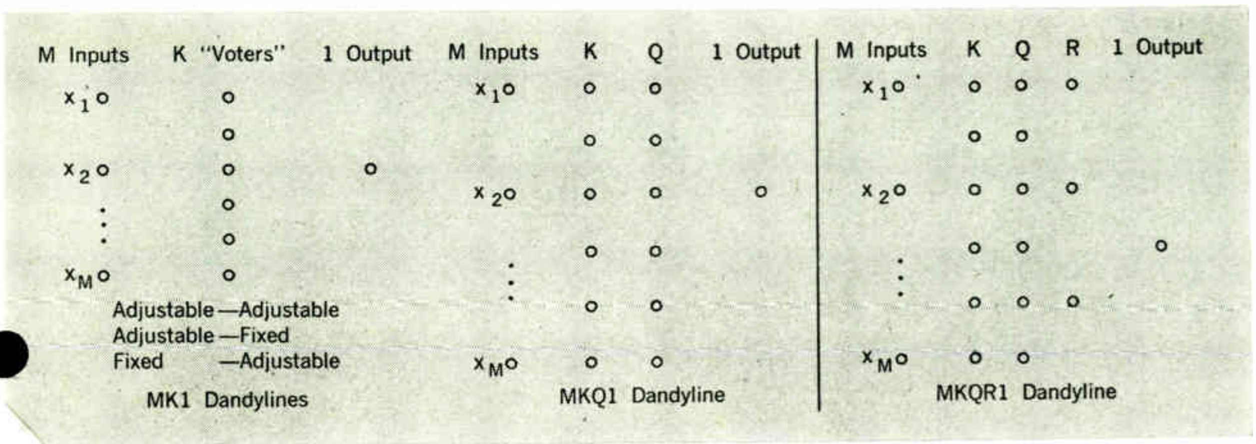


Diagram of an M1 Dandyline summing device.

A line sketch of MK1, MKQ1 and MKQR1 Dandylines with different fixed and variable layers.



threshold element, in general, cannot realize all possible Boolean functions: indeed, as M grows, the ratio of realizable to possible functions decreases considerably.[1]

M	0	1	2	3	4	5	6	7
Possible	2	4	16	256	65 536	4.3×10^{9}	1.8×10^{19}	3×10^{38}
	0	1	2	3	4	5	6	7
Realizable(R_M)	2	4	14	104	1882	94 572	15 028 134	10^{10}
In general, $2^{.33M^2} < R_M < 2^{M^2} \ll 2^{2^M}$; for M large.								estimated

The preselection of weights $W_1, W_2, \ldots, W_M$ to fulfill realizable Boolean functions is an easy job for switching-circuit designers, who have a variety of methods for small values of M.[2–4] Dandylines Enterprises does *not* design its lines by preselecting weights. Instead, on the M1, for example, we adjust the values of M arbitrary weights (on the bench, so to speak) according to a prescribed training rule until the correct values are attained. Three training rules have been found useful. All three start with the M weights (weight vector) chosen arbitrarily and present the M-variable inputs (pattern vectors) to the M1 line. If the M1 line classifies a pattern incorrectly, the weight vector is adjusted according to a particular rule. When the M1 line classifies all the pattern vectors correctly, it has achieved a satisfactory assignment of weights.

In accordance with the *fixed-increment* rule, the pattern vectors are presented in any order. One adds to the weight vector a constant K, times the pattern vector, whenever the latter is classified incorrectly. The process terminates after each pattern vector of interest has been presented and classified correctly without adjustment.

By the *error-correction* rule, the *fixed-increment* rule is applied repeatedly to a single pattern vector until it is classified correctly, before the next pattern vector is presented.

By the *relaxation method*, the weight adjustment depends on the value of the undiscriminated analog sum rather than the discriminated output. If on presentation of a pattern vector, the analog sum is, say, $-S$ when it should be positive, calculate K (depending on a preselected parameter λ) so that when K times the pattern vector is added to the weight vector, the new analog sum is $(1-\lambda)(-S)$. For fixed λ, K is different for each pattern and weight vector. It may be seen that for $\lambda = 2$, the analog sum is reflected; for $\lambda = 1$, the analog sum is taken to zero; for $0 < \lambda < 1$, the analog sum is partially corrected. In practice, all pattern vectors may be presented and the analog sums noted. Then, that pattern vector, whose analog sum is farthest from threshold, is multiplied by the appropriate K and added to the weight vector. The process is repeated to termination.

Training theorems. Theoretical justifications exist for the claims that these three training rules work. One oftprotved heorem asserts[5–12] that no matter with what assignment of weights we begin, the process of recursively readjusting the weights by error correction or fixed increment will terminate after a finite number of adjustments in a satisfactory assignment, provided such a satisfactory assignment exists.

The other theorem[13] asserts that no matter with what assignment of weights we begin, the process of recursively adjusting the weights by the relaxation method will terminate after a finite number of adjustments in a satisfactory way for $\lambda = 2$, provided such a satisfactory assignment exists.

If $0 < \lambda < 2$, the process is guaranteed only to converge. (Note: for $1 < \lambda < 2$, many experiments have been performed which *do* terminate.)

Realizability. Although a decreasing percentage of Boolean functions are realizable with increasing M, several practical problems do not require that all 2^M input vectors be classified. We may assume that as fewer than 2^M input vectors are to be classified, the probability of realizing such functions is increased. Indeed, we have the following theorem: "Given almost any $2M$ input vectors of dimension M; of the 2^{2M} possible binary functions, one half of these are realizable by an M1 line. If the fixed ratio of the number of M dimensional vectors to M exceeds 2, then as M tends to infinity, the probability that a random dichotomization is realizable by an M1 line tends to zero; the probability approaches one if the fixed ratio is less than two.[14,15] We say the capacity of the M1 line is $2M$—twice the number of weights.

Whereas this theorem does not hold for binary vectors, $2M$ is known to be the upper limit of the number of incompletely specified Boolean functions that are realizable.[15]

M1 boosters. At the cost of a preprocessor and additional weights, the capacity of an M1 line can be boosted from $2M$ to about $2M^r$, where r is a positive integer. We build second-, third-, fourth- and higher-order boosters, depending on whether one wants $r = 2, 3, 4$, etc. Actually, the preprocessor consists of

$$\sum_{k=1}^{r} \binom{M+k+1}{k}$$

multipliers whose products are weighted in the conventional way. The essential theoretical justification for the operation of an rth order booster is as follows:[16] the weights of an M1 can be thought of as separating points (patterns) in an M-dimensional space by passing linear, quadratic, or rth order surfaces among the points. The capacity of this M1 is twice the number of weights, which for an rth order surface is the sum of the binomial coefficients:

$$\sum_{k=0}^{r} \binom{M+k-1}{k} \sim 2M^r$$

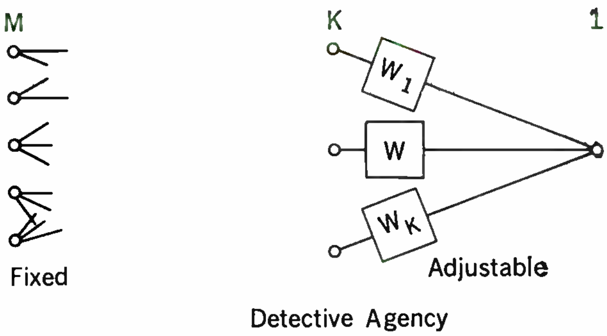
Applications. M1 lines may equal or excel human beings in dichotomizing under certain conditions and requirements. Specifically, an M1 may be suitable for identifying friend or foe; diagnosing a particular illness; judging guilt; predicting rain or shine; determining truth or falsity, validity, decodability, probability, completeness, etc.

To illustrate what conditions must be met, consider an M1 emulating a physician's diagnosis to determine presence of turberculosis. One would first teach the M1 using, as training input vectors and categorizations, actual case histories of those M symptoms that did or did not correspond to a diagnosis of tuberculosis. This requires that the *training set, or syndrome, be dichotomizable;* otherwise, the M1 would not even learn to repeat an accurate diagnosis already made by a physician.

If the training syndrome is dichotomizable, and the "medically" trained M1 is to diagnose its first case accurately, then *the new syndrome should not render the union of the training syndrome and it inseparable.* Indeed, each syndrome, that renders the union of the training syndrome and itself inseparable, will be wrongly diagnosed.

On the other hand, if an M1 were trained to diagnose tuberculosis by having access to case histories and to the patients themselves for purposes of testing, then, to guarantee accurate diagnosis, the M distinguishable characteristics of tubercular patients would have to be ascertained from the case histories, and also by further tests. The *training syndromes for each accurately diagnosed person would have to be separable*, as would each new syndrome when a new diagnosis is made.

Dandylines Applications Division suggests M1 Dandylines for simulating or emulating, judging, diagnosing, predicting, and deciding between two prognoses, provided the customer can state confidently that the appropriate conditions of separability are met. Thus, M1 Dandylines may be useful to Bright Field for some medical diagnoses, weather prediction, and for imitation of organization officials.



Sketch showing Dandylines' fixed M-K layer, variable K-1 layer device, which is called the Detective Agency unit.

MK1 Dandylines—The "Detective Agency"

The Detective Agency is an MK1 line whose M–K layer is fixed and whose K–1 layer has variable weights. The adjustable K–1 layer is similar to the M1 Dandyline. This configuration provides an opportunity to reduce the usual large number M to a modest-sized K, thereby decreasing considerably the number of weights otherwise required to perform a switching function by an M1 Dandyline. Furthermore, the number of realizable functions may be increased thereby. It is illuminating to consider that the wiring in the first layer detects the K distinguishing features and distortions of the input vector. Thus, we call this Dandyline, *The Detective Agency*.

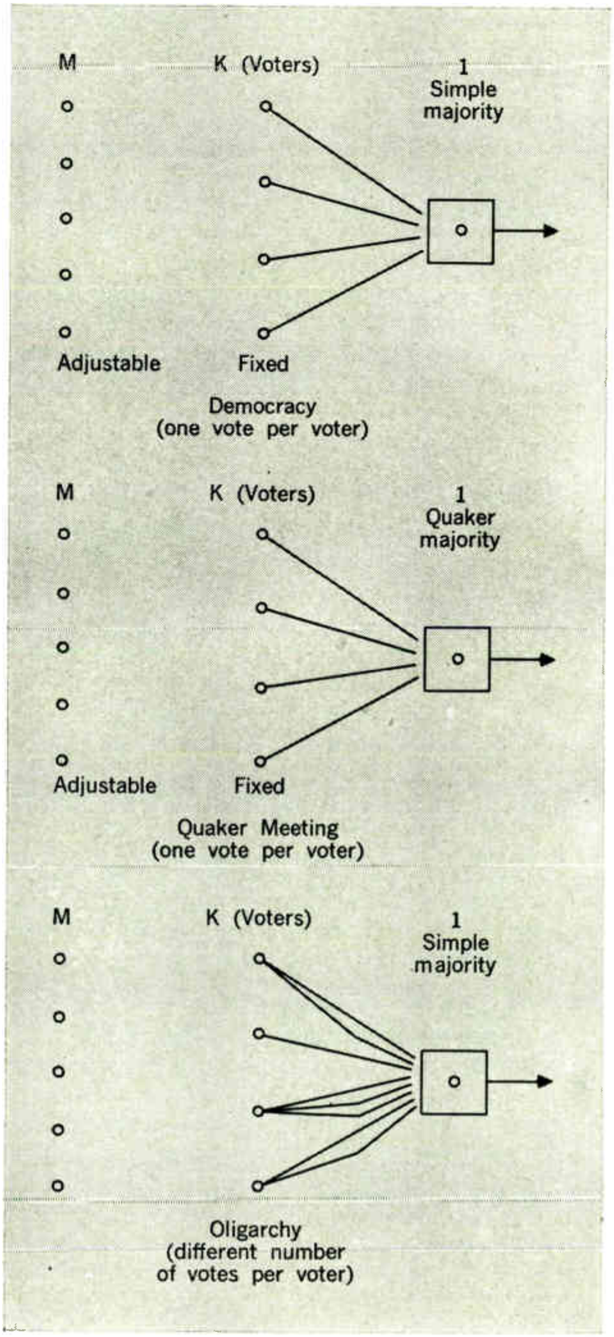
Applications. For instances where distinguishing features are unknown, we recommend a Detective Agency as an experimental tool for ascertaining empirically the distinguishing features by first-layer wiring modification. For example, the M observations made by physicians may reduce the K distinguishable features of a disease to a much smaller number.

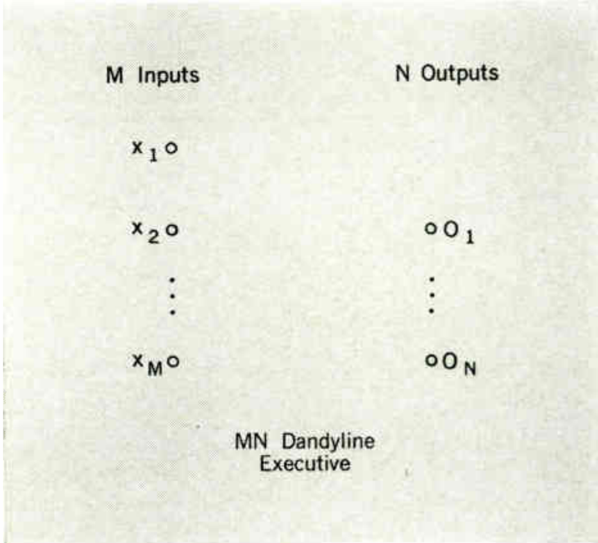
The Democracy, Quaker Meeting, and Oligarchy

The members of the Democracy, Quaker Meeting, and Oligarchy series of Dandylines are adjustable in the first layer and fixed in the second. They are typed according to whether each of the K "voters" is given equal or unequal weights and whether dichotomization requires a simple or greater majority. Our most popular MK1 Dandyline is the Democracy, one vote per voter, with a simple majority required for electing a class. Our Quaker Meeting, requiring better than simple majorities is growing in popularity. The Oligarchy, wherein voters are given different numbers of votes, is in the experimental stage.

It is plausible that the Democracy, having several (K) M1 lines could realize more of the 2^{2^M} Boolean functions than could a single M1. Indeed, it can be shown that to realize any Boolean function by a Democracy requires, at most, a number (K) of M1 lines equal to the number of input vectors that are in the smallest class. Furthermore, our experience indicates that for K voters, a Democracy will dichotomize up to any 2 KM input vectors (not necessarily separable) and that KM input vectors can be dichotomized in about M iterations using the following training rule: either by fixed increment or error correction, they adjust only those weights affecting the minimum number of voters necessary to change the category. A

Diagram showing circuit variations and various elements of Dandylines' Democracy, Quaker Meeting and Oligarchy.





The MN Executive is the custom line of Dandylines Enterprises. This unit represents the ultimate in pattern recognition, logic functions and associative memory capabilities.

theorem,[17] subsequently revised (whose proof was first challenged[18] by a counter example), states that this procedure for a Democracy converges.

It is intuitively clear that Quaker Meetings will realize fewer incomplete Boolean functions than Democracies.

Oligarchies of K voters would seem to be able to dichotomize specified functions less completely than a Democracy with an equal number of voters. Yet, it has been proved that a three-voter Democracy can do everything that a three-voter Oligarchy can do.[19] It is unresolved as yet whether a Democracy can do as well as an Oligarchy with the same number of voters.

MN Dandylines—Executives

There are $2^{N \cdot 2^M}$ possible N-output switching function of M inputs. An MN Dandylines can realize only a few of these functions; the exact number depending on the simultaneous separability of the input vectors by each of the N switching functions.

We estimate that the capacity C of MN lines lies between M and $2M$; $M < C < 2M$, whether we define C to be either (1) the number of input patterns of M dimensions that can be correctly assigned to binary N-tuples selected at random with a probability of $1/2$; or (2) the expected value of the random variable r defined as follows: given a random assignment of 2^N categories to a group of input patterns, r is the number of successive patterns of a random sequence that can be assigned correctly.

Any of the training rules that work for M1 lines work for MN lines.

Applications. MN lines classify objects represented by M distinguishing features into one of 2^N categories. We have named MN Dandylines, *Executives*, since they also decide upon one of several alternatives. Under certain conditions and requirements, an Executive may emulate many intelligence functions identified in Exhibit "A."

Executives can be custom-built to do these things:

(1) Recognize a limited vocabulary of words in almost any natural language and to recognize the speaker as well.
(2) Recognize a limited number of noisy and distorted handwritten characters.
(3) Identify individual noisy targets.
(4) Make word-for-word (and some idiomatic) translations of many foreign languages.
(5) Play the role of an associative memory; act as a code converter; perform all arithmetic operations; indeed, do any mathematical operation on input quantities whose output may be stored in a table.
(6) Play checkers and chess (M will increase with the number of previous moves that accommodate to the strategy of the game).

Dandylines wishes to say that if a Dandyline can perform any task at all, it will improve with experience.

Adaptively yours,

Layer N. Weight

Layer N. Weight, President
Dandylines Enterprises

REFERENCES

1. Winder, R. O., "Single-Stage Threshold Logic," *AIEE Conference Paper* 60-1261, Oct. 1960.
2. Gabelman, I., "The Synthesis of Boolean Functions Using a Single Threshold Element," *IRE Trans. on Electronic Computers*, vol. EC-11, Oct. 1962.
3. Minnick, R. C., "Linear Input Logic," *Ibid.*, vol. EC-10, Mar. 1961.
4. Chow, C. K., "Boolean Functions Realizable with Single Threshold Devices," *Proc. IRE*, vol. 49, Jan. 1961, pp. 370–371.
5. Block, H. D., "The Perceptron: A Model for Brain Functioning," *Revs. Modern Phys.*, vol. 34, Jan. 1962.
6. Rosenblatt, R., Rept. VG-1196-G-4, Cornell Aeronautical Laboratory, Buffalo, N.Y., Feb. 1960.
7. Joseph, R. D., Tech. Memo 12, Project PARA, Cornell Aeronautical Laboratory, Mar. 1960.
8. Joseph, R. D., Tech. Memo 13, Project PARA, Cornell Aeronautical Laboratory, July 1960.
9. Stafford, R. A., "Learning by Threshold Elements in Cascade," Tech. Note 20081, Aeronutronics, Mar. 8, 1961.
10. Singleton, R. C., "A Test for Linear Separability as Applied to Self-Organizing Systems," Conf. on Self-Organizing Systems, May 22–24, 1962, Chicago, Ill.
11. Novikoff, A. B. J., "On Convergence Proofs for Perceptrons," Tech. Report, Proj. No. 3065, Jan. 1963, Stanford Research Institute.
12. Nilsson, N., Unpublished proof.
13. Motzkin, T. S., and I. J. Schoenberg, "The Relaxation Method for Linear Inequalities," *Can. J. of Math.*, vol. VI, no. 3, 1954.
14. Efron, B. and T. Cover, "Linear Separability of Random Vectors," Unpublished research paper, 1963.
15. Winder, R. O., "Bounds on Threshold Gate Reliability," May 9, 1963; submitted to *IEEE Trans. on Electronic Computers*.
16. Cover, T., "*R*th Order Separability," Stanford Research Institute Memorandum, July 31, 1963.
17. Ridgeway III, W. C., "An Adaptive Logic System with Generalizing Properties," Tech. Rept. No. 1556-1, Stanford Electronics Lab., Stanford, Calif., Apr. 1962.
18. Singleton, R. C., Private communication, May 28, 1962.
19. Nilsson, N., Unpublished notes, 1962.

Search Ltd.

CARNAGE, PENNSACHUSETTS

To: Bright Field
Attn: SMARTIES
Subject: Proposal in Response to RFB No. 123 (456)-ABC-789-D

January 24, 1964

Dear Sir:

While reading the list of 14 tasks in Exhibit "A," I was gratified that scientists in Search Limited have considered, experimented, and written extensively on virtually every item on the list. This is not a coincidence. We assume that you are sufficiently familiar with our abilities and products to waive our detailed discussion in this proposal.

Search Limited will try to justify your confidence in us. We would be delighted to negotiate separately a contract covering some of the listed items (e.g., theorem proving) upon which we are certain our competitors will not bid. We feel sure that our demonstrated competence and experience also will influence your decision to award the other contracts in our favor.

Our products are rarely hardware. Rather, we produce heuristic and algorithmic programs for the performance of tasks by computers—analog and digital. Occasional by-products of our company are formal (as opposed to natural) languages; problem-oriented languages suitable for describing certain classes of problems to a computer; languages for documenting programs to facilitate program communication among humans; languages for facilitating computer data-processing for certain kinds of problems. Sometimes, we encounter problems for whose solutions we would prefer to have computers with certain characteristics. If such computers are not available, we may design them.

Search Limited writes heuristic programs to do intelligent tasks in the fields of mathematics, statistics, business, medicine, psychology, law, politics, the military, biology, taxonomy, linguistics, space science, pattern recognition, meteorology, etc. A heuristic program serves any one of two purposes: (1) to *emulate* humans in the performance of a task; and (2) to *simulate* the performance of a specific task as an exercise to obtain insight on how a human performs that task. We know Bright Field appreciates that, although simulation is not specifically mentioned in Exhibit "A," it may be an important tool in designing emulators sought by Bright Field. Indeed, the Search Limited Emulation and Simulation Groups share many programming and problem-solving methods, and stimulate each other on new ideas and approaches.

Heuristic programming for performing an intelligent task consists of seeking an acceptable short sequence of transformations on the distinguishing features of a set of inputs. This will map the inputs to desired recognizable outputs. This general formulation is expressed variously in the vernacular of individual fields. For example, linguists speak of *phonemes*, psychologists of *attributes*, meteorologists of *predictors*, pattern recognizers of *properties*, etc. These are all synonyms for "distinguishing features" of a set of inputs. Despite the obfuscation of the general formulation by the jargon of the various sciences, professions and arts, the writing of an acceptable heuristic program implies at least these processes:

(1) Finding an efficient representation of the environment of interest (the Representative problem)
(2) Identifying an efficient, sufficient, appropriate, and representable set of distinguishing features of the environment (the Scientific Bureau of Investigation problem)
(3) Inventing workable rules of combination of the distinguishing features (the Legislator problem)
(4) Selecting a valid sequence of transformations from experience to arrive at a desired output (the Navigator problem)
(5) Devising a test for recognizing the desired output and applying the test (the Chairman of the Board problem).

Indeed, we are tempted to diagnose that our disappointing programs have been caused by our inability to solve one or more of these five problems. Programs for proving theorems are disappointing because so few have been proved. Programs for medical diagnosis, voice, talker, and handwriting recognition work after a fashion, as do our battle, corporation, national economy emulators, and simulators. In a few instances such as checker-playing, emulations have been quite successful. But, simulation of learning and generalization has been largely unsuccessful. Unfortunately, when we succeed we don't know why and when we fail we also don't know why.

We do not imply that Search Limited is admitting failure—quite the contrary. No one has shown that such heuristic emulations and simulations are impossible; nor that we shall be forever without solid foundations. Indeed, we are confident that we can perform on this contract to your satisfaction.

Heuristically yours,

Hew R. Itzik

Hew R. Itzik, President
Search Limited

BIBLIOGRAPHY

Bruner, J. S., J. J. Goodnow, G. A. Austin, *A Study of Thinking*. New York: John Wiley & Sons, Inc., 1962.

Clarkson, G. P. E., Portfolio Selection: "A Simulation of Trust Investment," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

Hunt, E. B., *Concept Learning*. New York: John Wiley & Sons, Inc., 1962.

Minsky, M., "Steps Toward Artificial Intelligence," *Proc. IRE*, vol. 49, Jan. 1961, pp. 8–30.

Newell, A., "Learning, Generality, and Problem Solving," *Proc. of the IFIP*, Congress 62, pp. 407–412.

Preprint, *First Congress of the Information System Sciences*, The Mitre Corp., 1962.

APPENDIX

The following are a few illustrations of the various jargons used to denote the five common subproblems. Aristotle spoke of essential and accidental *properties* of objects or events to discriminate between the sets of relevant and irrelevant *features*. Statistical decision theorists speak of selecting the space to be partitioned by *receptors* and methods of partitioning the space by *categorizers*. Modern psychologists speak of *cue searching* in reference to the search for essential *attributes* basic to the decision of placing an object in a category. They speak of concept formation in reference to a rule of combination of the essential attributes whose application results in a concept. Sometimes, the rule is called a concept. Sometimes, the whole process of identifying essential attributes to place an object in one of the several categories is called *concept attainment*.

Practitioners involved in recognition of written, printed and spoken messages, targets or talkers speak variously of *property filtering* or *feature determination* for finding distinguishing features. Recognition, rules of classification or categorization are their labels for rules of combination. Police speak of *clues*, which, when appropriately put together, point to a suspect. Judges and juries consider *evidence* and combine the evidence according to the rules of law in order to arrive at a judgment after assigning weights to the evidence. Physicians and psychiatrists observe *symptoms* and on combining them by the rules of science, intuition or experience, or all three, make *diagnoses* on the basis of which they must apply some further rules in order to prescribe appropriate treatment.

The main problem of statistical weather prediction is to identify the distinguishing features called *predictors;* the problem of information and document retrieval for reference librarians is to identify distinguishing features called *descriptors*, or copyrighted synonyms thereof.

OPTYMISTICa

DISSYLAND, CALIFORNIA

To: Bright Field
Attn: SMARTIES
Subject: Proposal in Response to RFB No. 123 (456)-ABC-789-D

January 28, 1964

Dear Sir:

Optimystica feels both pride and dismay at the opportunity to make a work proposal concerning artificial, intelligent, adaptive mechanisms and organisms to do actual tasks. We are proud to be chosen as potential participants in the very work for which our organization is continuously forming. The reason we are dismayed is that we've never been asked before to submit a proposal to build actual hardware or to write actual programs. Thus, we beg your indulgence if this, our first proposal, fails to meet some of your exacting standards.

Optimystica is an organization currently engaged in nurturing its own description. We are a loosely-knit body of a few hundred persons, dedicated to a policy and *modus operandi* which is quite parallel to the ideas we espouse on how to build adaptive, self-organizing, creative, goal-oriented, visible mechanisms, whose ultimate design and capability we may not know anything about beforehand. We think that, from the interaction of ideas, our organization can orient itself to the formulation of a goal; can synthesize itself as a machine for attaining the goal: and can remain a viable, adaptive, learning creature, without any member of the organization knowing—or even caring—what the other members are doing. At present, we are not a commercial undertaking, but we expect soon to be one. We are a company quite unlike the conventional ones: for in accordance with our fundamental design doctrine, our self-organized organization will, like standard companies, be disorganized—but we will be deliberately so in an orderly way.

Optimystica's staff is a group of scientists who have compiled the Optimystica "bromoplasm." This rudimentary organism has decided that its initial goal is to compete for contracts which will constitute is nutrition. Thus, our early publications, which were in handwritten or in stenotypic coded form, have evolved to the form you now read.

One minor example of our adaptation has been the change of name. Hitherto, we have been called Optimysiques. But the emphasis this throws on the third and fourth syllables has led people to misinterpret our nature as concerned with mysticism. At any rate, for a creative enterprise, rather than a strictly mystical one, the more feminine ending is appropriate and less harsh.

It is a fortunate coincidence, both for Bright Field and Optimystica, that we have already addressed ourselves seriously to the very problems upon which you are asking us to propose; namely, how to build artificial intelligences. In addition, one of our most distinguished *bromoplasts* has assured us that "the problem is already solved... what is wanted is a demonstration of all our developments in the complexity of actual practice..." But we are getting ahead of the story. The only purpose in this brief preview of our proposal is to assure you that General Ize came to the right people when he invited Optimystica to submit a proposal.

Technical proposal

The following description is the essence of our thinking on how to build an artificial intelligence.

The parts would be, for example, pieces of confetti and ungluable glue for sticking the pieces together. An iron lung would keep the confetti in motion. Or, we might take the holes in a dielectric which has interacting components and an external source of power.

But a confetti-glue-iron lung mechanism is spatial. To make a temporal mechanism, merely encode this proposal in binary mode, circulate it in the Univac I mercury delay-line memory and the result will have components, a system and its power source.

You are unimpressed. Then hear this. We hold with the statement made by the late distinguished I. M. Pedantic: "Every human action, emotion and idea can be explained by evolution. If we cannot explain cerebral operations with current laws, then these laws are wrong even for those phenomena they do explain."

Anybody who has seen a motion picture taken by a hidden camera, will appreciate this infinite dexterity and may conclude that these apparently simple acts are already beyond our means of apprehension—or comprehension.

In anti-Gestalt psychology, and in unfree association theory, the phenomenon of projection is stressed. In our own nonsense libidinous Weltanschauung, we project the environment on ourselves and vice versa. How do we know that we are not the environment and the environment is not we?

Optimystica staff holds, though by no means with unanimity, that environment is simply a stochastically misanthropic, maldistribution of both the solar and lunar dequantized reradiations, plus their hazardous, uncertain mass equivalents on earth. A schizophrenic may be merely an assemblage of L.S.D. molecules with moonlight as his power source. In fact, we have uncovered a low-Q meson dissonance that occurs when ultraviolet irradiates casein. All we need to complete the explanation is to find the anti-anti-particle which Professor B. O. Taney has indicated to exist in every plant. He found an extremely high correlation between the absence of plant life and sand beaches on which the extent of lunar tides can be measured easily. Thus, he came to the obvious conclusion that each plant keeps time through a circulatory mechanism by which an anti-anti-particle combines with its dual at every high tide.

We can summarize our approach to designing artificial intelligences, capable of survival independent of the laboratory, by again quoting our distinguished bromoplast: "I can say definitely that the problem is already solved. The design of an artificial intelligence is something that is fantastically complicated and difficult from one point of view; yet is fantastically simple if approached from quite another point of view. The fact is that *any* complex dynamic system, if allowed to run towards equilibrium, will, as it approaches equilibrium, show the increased characteristic relationships that exist between organism and environment.

"In my opinion the principles are completely understood already. Now, we want a demonstration of all their developments in the complications of actual practice."

Approach

In accordance with this proclamation, we will not waste your money on solving the artificial intelligences design problem. We have already done so. We will merely subcontract the engineering and programming work to others competent in such details. Of course, we will steadily eliminate from our minds all the wrong ideas that we have picked up from the past and we shall reeducate ourselves and all our new hirelings.

Optimystica again asks your indulgence if this proposal does not meet your format requirements.

Optimystically yours,

Homer O. Stasis

Homer O. Stasis
President

P.S. Reference*: ARTORGA—The sixth quantized surge of information through our organism. May 15, 1959.

* Those unfamiliar with, and/or skeptical of, the existence of Optimystica and/or H. O. Stasis, are requested to read this reference.

PESSIMYTHS
INC.

To: Bright Field
Attn: Col. T. Hink
Subject: Observations on and Evaluations of Proposals in Response to RFB No. 123 (456)-ABC-789-D

January 30, 1964

Dear Sir:

You have assigned to Pessimyths the task of evaluating the technical proposals made by Optimystica, Search Ltd., Calculated Risks, Inc. and Dandylines Enterprises in response to your invitation; and to suggest to you which research projects to initiate or abandon in the field of artificial intelligence.

Pessimyths has studied the four proposals in detail. The evaluations and the advice contained herein result from our proposal studies and the experience of the Pessimyths staff in making such appraisals.

The request for bid

At the outset, we must mention Bright Field's dereliction in failing to ask certain companies to bid, especially Switching Functions Co. and Not-Implausible Cybernetic Models.

Although the former does not describe its work and results in anthropomorphic terms, Switching Functions does possess much theoretical and practical knowledge on the design of devices to perform switching functions. Using switching function equipment to perform the tasks in Exhibit "A" might turn out to be prohibitive and costly. But, by not asking Switching Functions to bid, you denied yourself valid information as to what prohibitive costs would be.

Not-Implausible Cybernetic Models might have written an informative proposal representing *that* school of thought. Both they and Switching Functions might have suggested building blocks other than linear input threshold elements.

We note also that in Exhibit "A" you explicitly ask for emulators; simulators are implicity recognized as necessary. The mathematical assistant (no. 8 in Exhibit "A") is the only instance of a supplement to the human. Nowhere does Exhibit "A" either mention or imply an artificial intelligence that would be a colleague (synnoent) to a human intelligence. Pessimyths considers that, at present, intelligent assistants and synnoents are at least as important as emulators and simulators in man's exploitation and understanding of artificial intelligence.

An evaluation of Optimystica will not be made; its maladaptive proposal speaks for itself.

The proposers

Search Limited is a large company with hundreds of staff scientists. Dandylines Enterprises is considerably smaller. The Calculated Risks, Inc., Department of Artificial Intelligence, is by far the smallest of the three. Despite all disparities in size or appearance, all three are equally fierce in their devotion to their own corporate dogmas; one worshipping at the shrine of the linear input threshold logic unit; another agnostic to non-Gaussian distributions, and the third putting its faith and hope in heuristics—which is as elusive as the term "deity." It is easier to explain what heuristics does not mean rather than what it does mean. These people are cultists. They do not seem to realize that many future artificial intelligences, like present hybrid analog-digital systems, probably will consist of combinations of the product lines of all three companies—and possibly of others as well.

Except perhaps for Calculated Risks, Inc., the companies are mostly flying blind. Occasionally, however, a ray of light appears and the companies, especially Dandylines, are grateful. But they are evidently still skeptical enough to consider the possibility of a mirage. Why else would Dandylines reprove the training theorem at least eight different ways? Pessimyths does not employ a historian of mathematics, but we conjecture that this is the most oft-proved theorem in mathematics!

We have adopted the naming policy of the companies who designate their product lines after appropriate persons. The ablest of the pioneers in the field of *Artificial Incandescence* was Thomas Alva Edison, whose modus operandi could be epitomized accurately by the union of the slogans of the three companies. At Pessimyth's we refer indiscriminately to workers in Artificial Intelligence as *Edisons*.

Like their namesake, most Edisons have implicit faith that the problems they address are soluble by the means they advocate. Search Limited remarked on this point in its proposal. A Dandyline Edison declared to one of our interviewers that, until someone proved an existence theorem to the contrary, he would continue the attempt to make a Dandyline that generalized. That is, with a probability approaching one, a Dandyline would recognize patterns upon which it was not trained.

Edisons of all companies have arrived simultaneously at the realization that despite the value of a program or machine as a processor it is useless if the features or properties constituting the output are irrelevant or incomplete. The recognition of this pattern is of *key* significance. It may be compared to the realization a decade ago that, no matter how good a computer one had, it was useless without good programs.

Edisons are enthusiastic and optimistic. They have promised such things as champion chess masters, music composers, infallible and quick reference librarians, spoken language translators and good mathematicians. They have failed to deliver. They have even stopped giving a "von Neumann constant" (the fixed time-to-complete estimated by a project leader at any time). There are many prototype programs and machines. Often the machines are simulated, but their performances are largely unacceptable. Thus, they are at best research and test vehicles. Artificial intelligence is in the research phase; the development and production stages are in the future. When confronted with this failure to perform, most Edisons indicated that there were, nevertheless, important gains from their work. They call these gains *fallout*. For example, linguists claim that they could not have learned as much about linguistics if they had not tried to make artificial translators.

Fallout claims are also made by those trying to simulate cognitive, neurotic and other brain processes. When the alleged rationale for simulating a brain process is to obtain an understanding of the brain from the model, it is seldom stated exactly how the expected information will be useful in the practical understanding of the actual brain. Pessimyths was told that, even if information from models were never used, new languages or machine designs probably would be generated in constructing the model in the first place. However, we can legitimately question whether new languages and machines are built most economically under such inspirations.

The proposals

Pessimyths cannot make a recommendation as to which companies deserve contracts on the basis of their specific technical proposals. In fact, we recommend against your awarding any development or production contracts—not because the companies differ in resources and qualifications, nor because of their different devotions to threshold elements, optimal decision devices, or heuristic programs, nor even because parts of their proposals and purported scientific literature read as if they were composed and issued by an advertising agency rather than by scientists. *We recommend against contract awards because the whole field is in the research phase.* Many questions should be answered *before* development can begin on a broad scale. Indeed, each company indicated as much in its proposal. We agree with their identification of high-priority research efforts and urge Bright Field to finance their implied requests for support.

Future research

We recommend support of Calculated Risks, Inc. in its research on Dr. Freud, President Lincoln, and other descendants of Socrates.

Research in Dandyline's Detective Agency, Democracy, Oligarchy, Quaker Meeting and multiple-layer systems should be supported with stress on proving training and capacity theorems.

Finally, we strongly recommend support research on what Search Limited has called its Representative, S.B.I., Legislator, Navigator, and Chairman of the Board, with emphasis on the first two.

Because development and production projects may not be successful at this time, they should, if possible, be reoriented to one of the research projects mentioned above.

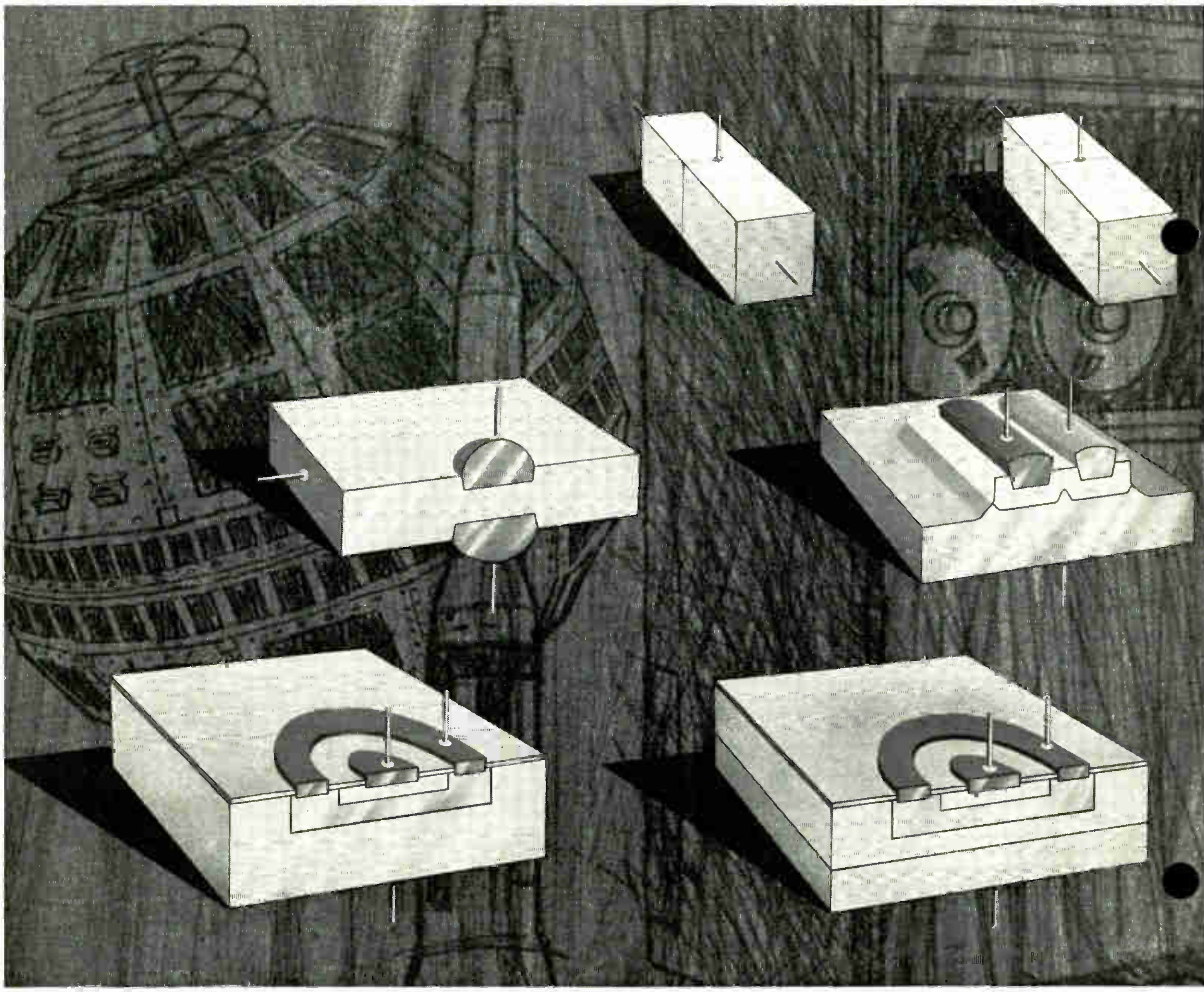
Pessimystically yours,

J. R. Bubble Piercer

J. R. Bubble Piercer, Pres.
Pessimyths, Inc.

AUTHOR'S ACKNOWLEDGMENT

The author wishes to thank all those whose writings and conversations helped shape this paper. In particular, he wishes to acknowledge the technical help of Dr. Nils Nilsson of Stanford Research Institute. He is especially indebted to Dr. Charles A. Rosen of the same institution for kindly providing encouragement and support.



GROWN transistors employ a time-proven manufacturing process for true economy and reliability.

ALLOY transistors feature low cost, mass production uniformity and reliability.

DIFFUSED PLANAR transistors give extremely low leakage and high gain at low currents.

GROWN-DIFFUSED silicon transistors couple product-proved reliability with true economy.

ALLOY-DIFFUSED transistors offer high-frequency response and economical price.

EPITAXIAL PLANAR transistors give extremely low leakage, off-set voltage and saturation resistance.

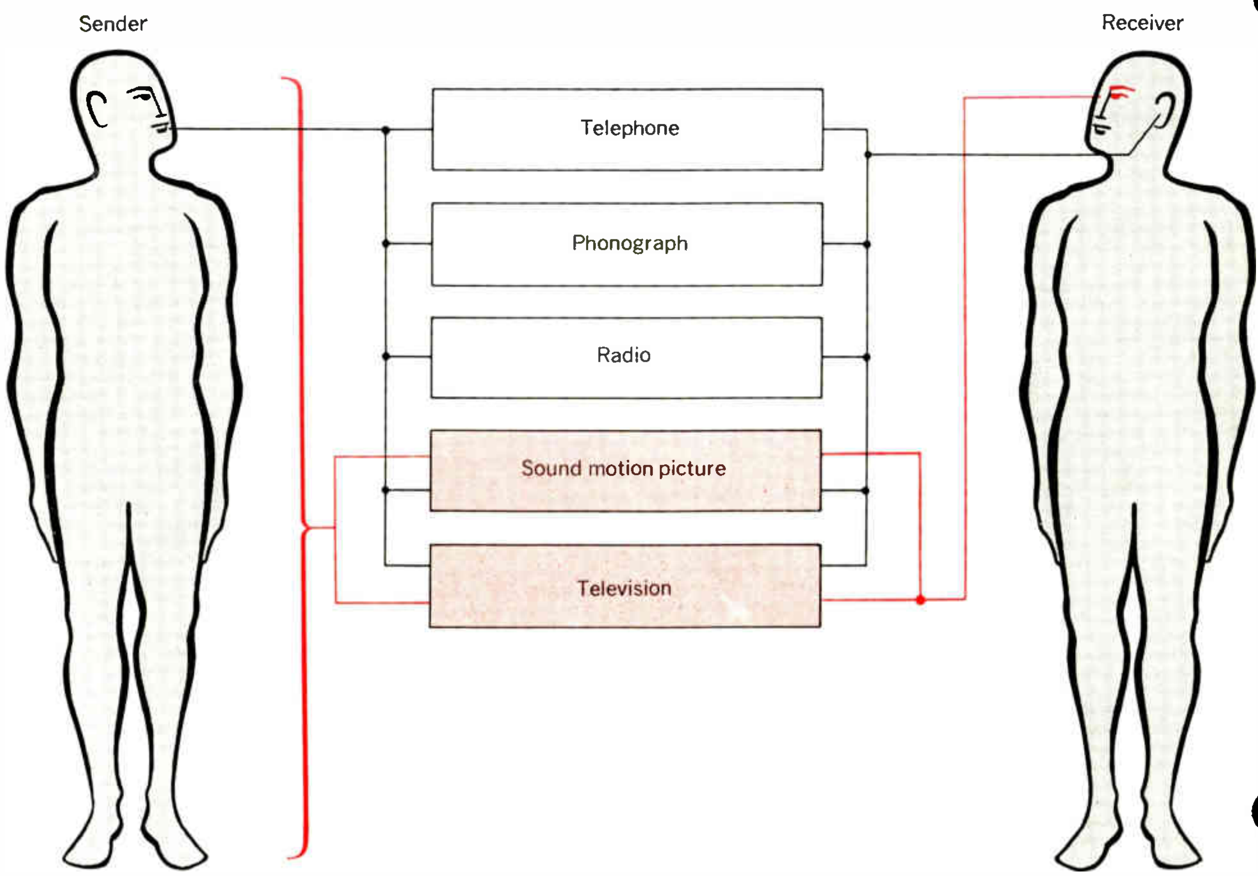
THE BEST TRANSISTOR TI MAKES

Speech processing systems

Upcoming systems—analyzing, code converting, and synthesizing speech sounds—will revolutionize communication both between individuals and between machines

Harry F. Olson *RCA Laboratories*

Fig. 1. Existing communication systems. Those incorporating visual techniques are shown in color. Speech transmission and aural reception are indicated in black.



The most important systems for the communication of speech information today are the telephone, the phonograph, the radio, the sound motion picture, and television. But the list does not end here. Additional elements, adjuncts, and systems for the transmission of speech intelligence will ultimately improve, modify, augment, or displace existing systems. Developments in the field of speech processing promise new facilities in speech communication.

Speech always has been and remains in the lead as the most practical and useful method for the transmission of information between individuals. Figure 1 depicts the communications systems presently in use between two persons—a sender and a receiver.

In the very early stages of development is a new systems category that involves the conversion, transmission, and utilization of speech information. When fully developed these systems will provide facilities for the reproduction of speech, its conversion to a code, its production from a code, its conversion from the printed page and to the printed page, the control of machines by speech, and the translation of speech from one language to another.

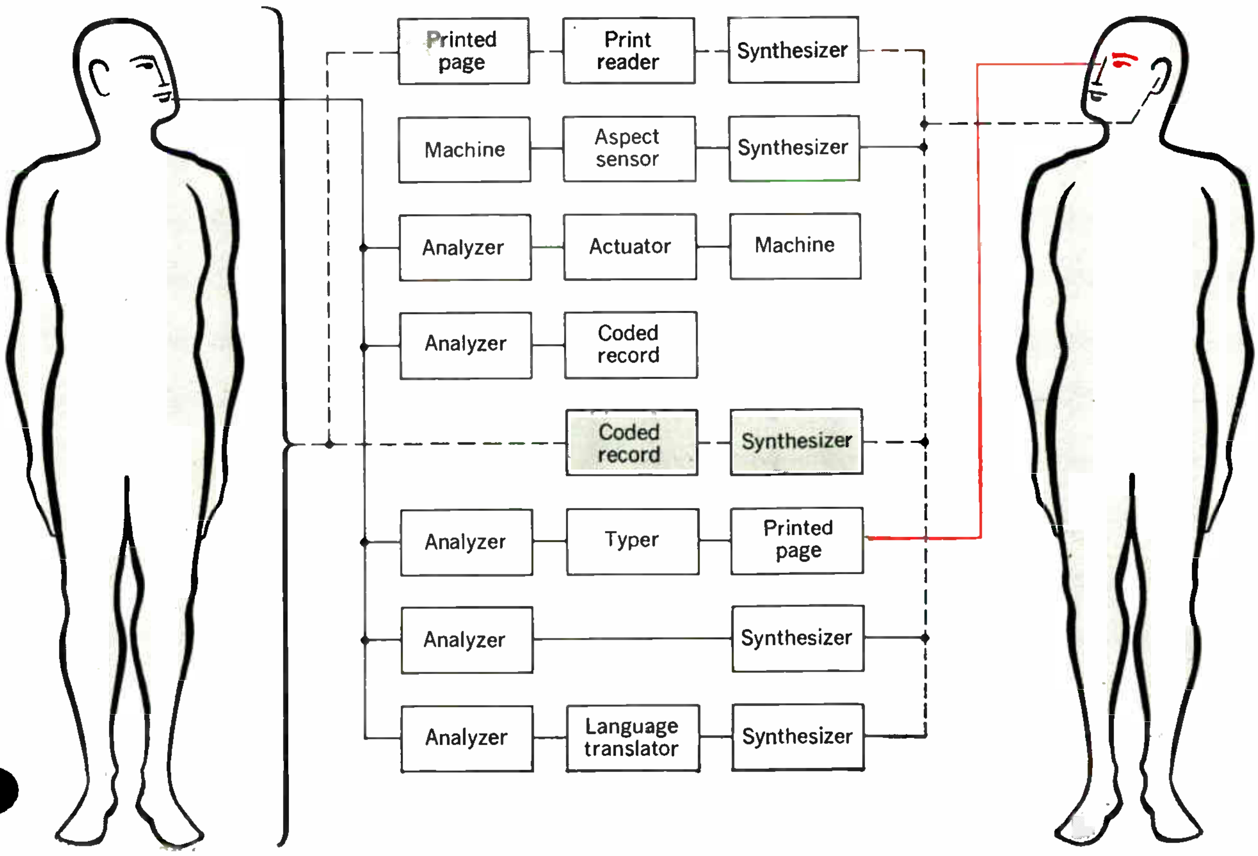
Figure 2 also shows a sender and a receiver. But in this instance the sender is using his voice to do one of the following: transmit speech information to the receiver, control a machine, or produce the typewritten page. As for the receiver, he may receive speech information, sense a machine, receive speech reproduction of the printed page or obtain language translations. To make these systems possible, speech-information conversion systems are required.

Systems for the conversion of speech information to and from a code are shown in Fig. 3 and include (A) the conversion of speech sounds to a code, (B) the production of speech from a code, (C) the conversion of the output speech sensor to a code, (D) the production of motion and force from a code, (E) the conversion of the printed page to a code, (F) the production of the typewritten page from a code, and (G) the conversion from a code in one language to the corresponding code in another language.

Analysis of speech sounds

A system for analyzing and separating the sounds of speech into discrete categories constitutes one of the basic

Fig. 2. New communication systems. Visual reception is indicated by color. Black dashes are for systems using aural reception and for other than speech transmission.



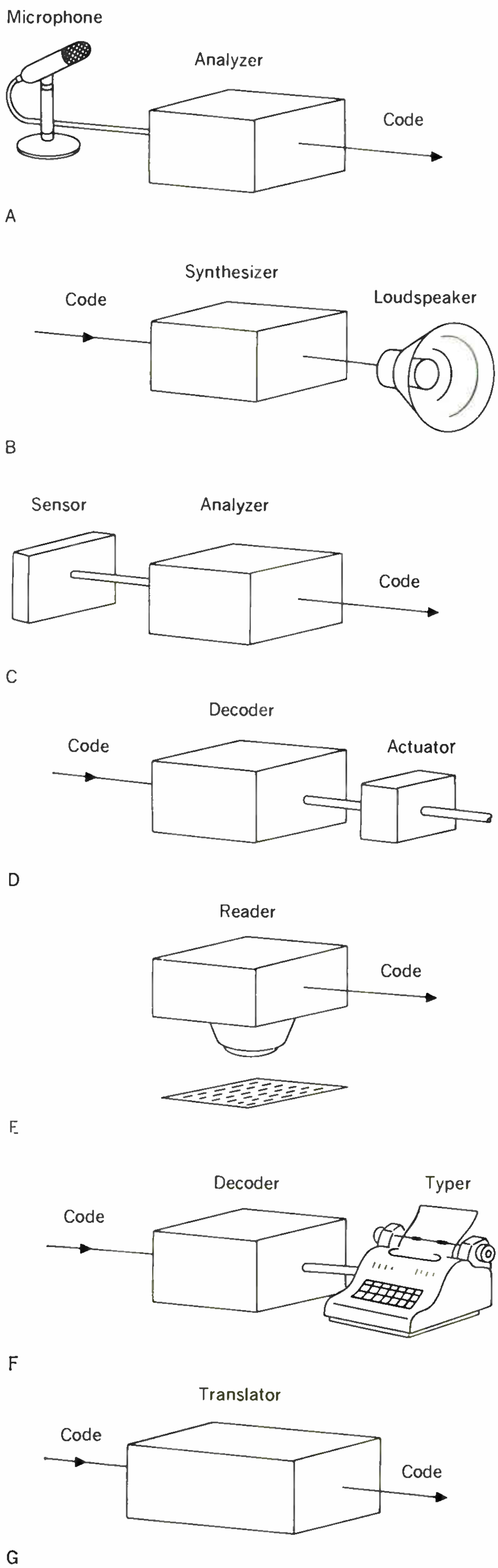


Fig. 3. Information conversion systems. A—Speech to code. B—Code to speech. C—Sensor to a code. D—Code to an actuator. E—Printed page to a code. F—Code to a typed page. G—Language translator with code input and output.

elements of any device for the conversion of speech sounds into a code. A sound wave may be completely described in terms of the amplitude and frequency of its components and time. Therefore, the analyzing system must be based upon frequency, amplitude, and time.

The amplitude of a typical vowel sound—in this case *I*—as a function of time is given in Fig. 4. This characteristic shows that speech carries a tremendous amount of information. However, there is nothing immediately apparent from the characteristic of a vowel in this form that can be used to identify and differentiate this vowel sound from another vowel or, for that matter, other sounds, so that the sound can be used to produce a specific code. To attain this objective, it seems logical to break down the characteristic of Fig. 4 into the parameters that carry the phonetic information in speech.

The amplitude characteristic as a function of the time given by Fig. 4 can be resolved into response frequency characteristics at discrete time intervals. Five response frequency characteristics of the vowel *I* at 0, 0.05, 0.1, 0.15, and 0.2 second are shown in Fig. 5. In this illustration, the parameters are amplitude and frequency at discrete intervals of the time.

The frequency-amplitude characteristic as a function of the time of Fig. 5 can be resolved into a frequency-time-amplitude characteristic, with the shaded area depicting the amplitude as shown in Fig. 6(A). If a definite upper amplitude limit is selected and also discrete sections with respect to time, the frequency-time-amplitude characteristic will be as shown in Fig. 6(B). The latter characteristic may be used as the elementary basis for the identification of sounds of speech and the conversion to a code.

The analysis of speech and the conversion to a code as described briefly in this section will be explained in greater detail in the section "Speech-to-code converter."

Nature of the code

To implement the new speech communication systems, conversion of speech to a code and its production from a code are necessary. The nature of the code depends to a major extent upon the smallest speech entity that can be analyzed out of context. The main consideration involves the number of each of these entities in the language. The per cent of all phonemes, syllables, and words as a func-

Fig. 4. The amplitude-time characteristic of a typical vowel sound—the letter I—of 0.25 second duration.

Fig. 5. Response frequency characteristics of a vowel sound at intervals of 0, 0.05, 0.1, 0.15, and 0.2 sec.

Fig. 6(A). Frequency-time-amplitude characteristic of a vowel sound (shaded area shows amplitude). (B)— Discrete frequency-time characteristics for a fixed threshold amplitude level of a typical vowel sound.

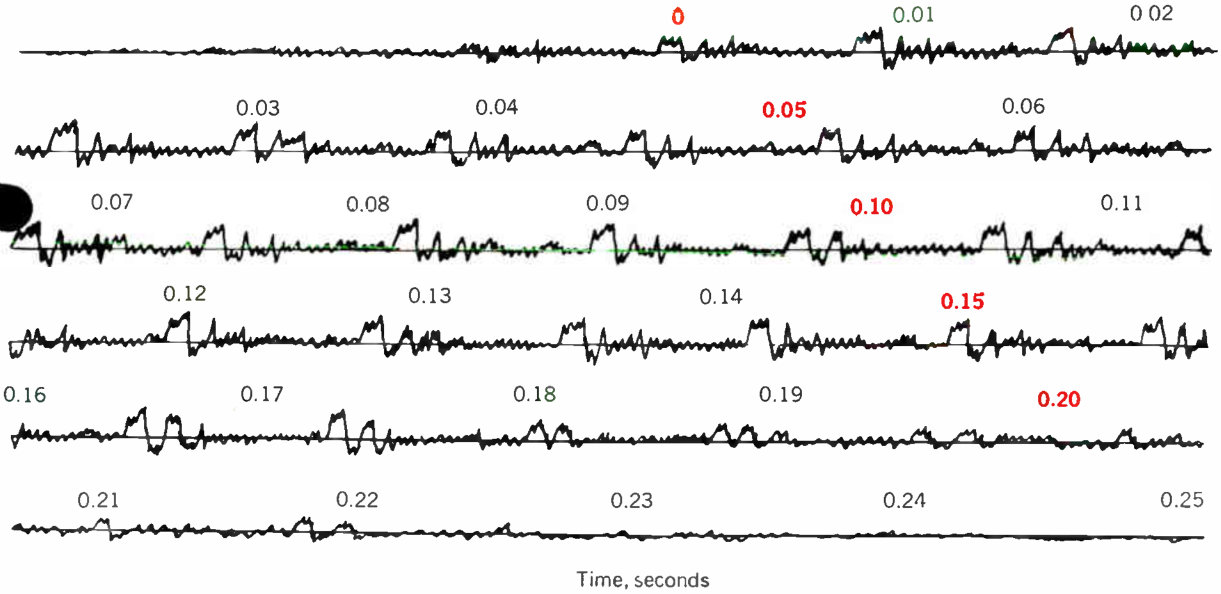


Fig. 4

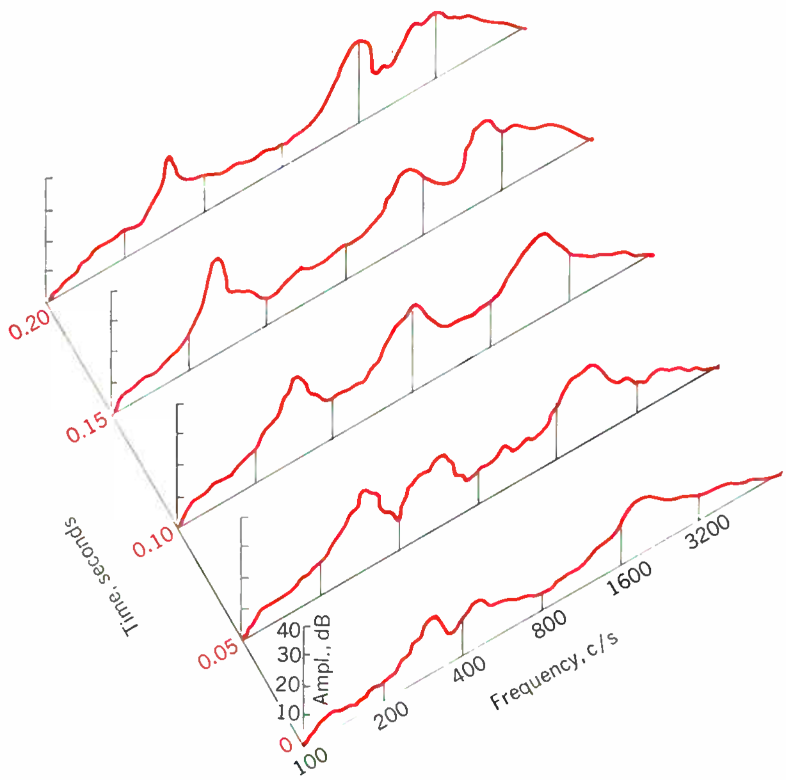
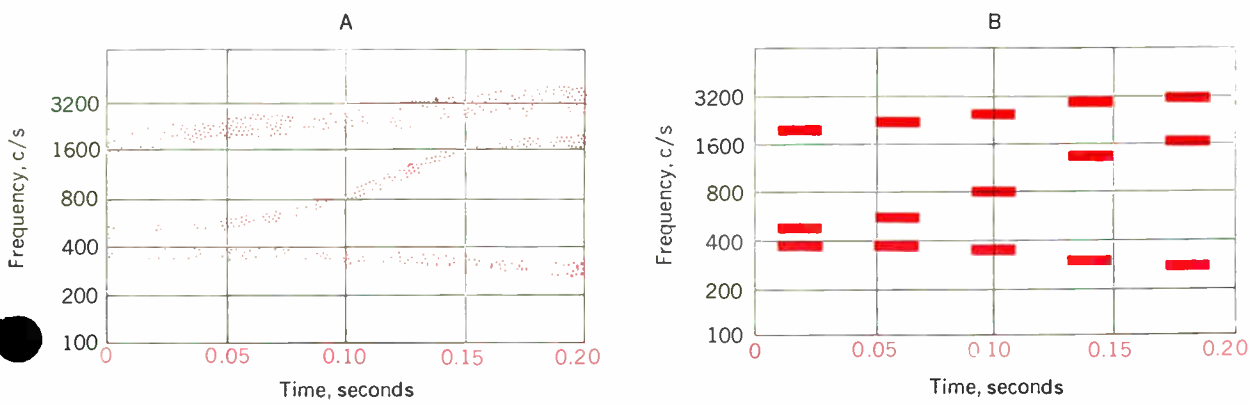


Fig. 5

Fig. 6



tion of the most frequently occurring sounds is depicted in Fig. 7.

The smallest speech entity is the phoneme. As shown in Fig. 7, there are about 42 phonemes in the English language based on U. S. usage. If speech could be analyzed in terms of phonemes, then the problems of encoding, storing, and decoding would be relatively simple. However, most words consist of several phonemes interconnected in such a manner that each is intimately associated with those adjacent. Therefore, it appears exceedingly difficult if not impossible to carry out the analysis on the phoneme approach except when the word consists of a single phoneme. Research on speech has indicated that groups of two or more phonemes, sometimes termed clusters, appear to be the smallest entity that can be analyzed by machine at this stage of the art. With some extension of the definition of the term, these clusters may be called "syllables;" the syllable may then consist of one or more phonemes. However, in this case it is not necessarily the conventional syllable as defined in the dictionary.

The information in Fig. 7 and research on the occurrence of the syllable as just defined indicate that approximately 1000 syllables will provide an adequate number for practical operation. It is assumed that the code depicted in the seven separate illustrations making up Fig. 3 is the syllable code.

Synthesis of speech

Speech may be synthesized by several different means, as for example, the vocoder synthesizer, the vocal tract synthesizer, the sound and music synthesizer, and the syllable synthesizer. Inasmuch as the syllable is the form of the output of the analyzer, it appears to be the logical basis of approach.

Research shows that if syllables are recorded as separate entities, by a person who enunciates clearly and distinctly, these individual stored syllables can be assembled to produce highly intelligible speech.

Speech-to-code converter

Figure 8 is a schematic block diagram of a speech-to-code converter consisting of a speech analyzer and syllable encoder. The first element is the microphone, specially designed and built for this application. The input to the microphone is speech consisting of syllables or monosyllabic words spoken one at a time. The figure also shows the wave train of the last part of the vowel sound *I* and the wave train of the electric signal from the microphone. The resultant audio-frequency electric signals from the microphone are amplified, compressed, limited, and suppressed in the normalizing input system. The suppressor is in the form of an expander with a preset gain. When the signal is under a threshold value, the gain can be considerably decreased. When it is over that value, the suppressor gain remains constant over a certain preset amplitude range of the signal; but for further increases in signal level above this range, the compressor limiter begins to reduce the gain. The amplitude-normalizing system has the distinct advantage of allowing a larger amount of compression without also producing acoustic feedback difficulties or reducing the effect of low-level noise.

The wave trains from the output of the compressor–limiter–suppressor and the dividing network are shown in Fig. 8. The amplitude-processed audio signal from the dividing network is fed through eight band-pass filters, the response frequency limits of which are shown. The range under 250 c/s is not used because there is very little information in this frequency band but the band of frequencies over 5000 c/s is useful for providing information on consonant sounds. The frequency ranges of the remaining band-pass filters were selected so that the information-bearing components would be the same in each frequency band. The figure shows the wave trains at the output of the band-pass filters for the last part of the vowel *I*. It will be noted that the frequencies of the outputs correspond to the last part of the vowel *I* as has been illustrated in Fig. 6(B).

The signal in each band of frequencies is again amplified and rectified and the dc voltages obtained are proportional to the envelope in both positive and negative directions. Amplitude and amplitude-sensing circuits are provided for different setups. The output of each channel may be quantized directly by operating a relay above a predetermined threshold—the analyzing process corresponding to the pattern in Fig. 6(B). In addition, the comparison circuit can be connected so that a relay will operate whenever the level in that channel is higher than the level in the channel of next lower frequency. This type of amplitude comparison circuit provides information on the first derivative of the frequency spectrum considered as a curve. If the amplitude comparison network for a channel is connected to both adjacent channels, as indicated in the schematic of Fig. 8, then a relay will operate when the signal level in the channel is higher by a predetermined amount than the average of the channels. This state of affairs, in effect, corresponds to specifying the second derivative of the curve representing the contour of the frequency spectrum at any one time. As a result, the output of the analyzers is in the form of quan-

Fig. 7. The per cent of all phonemes, syllables, or words are shown here depicted as a function of the most frequently occurring phonemes, syllables, or words.

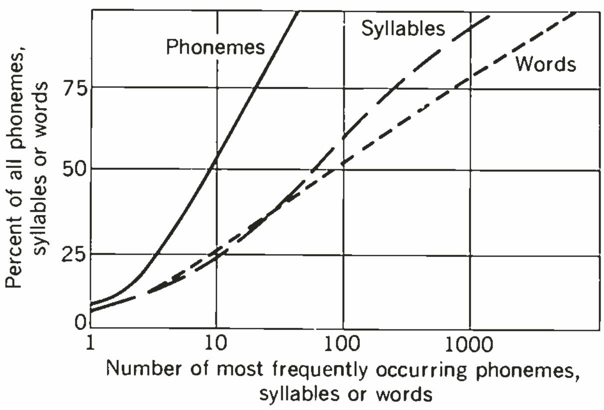
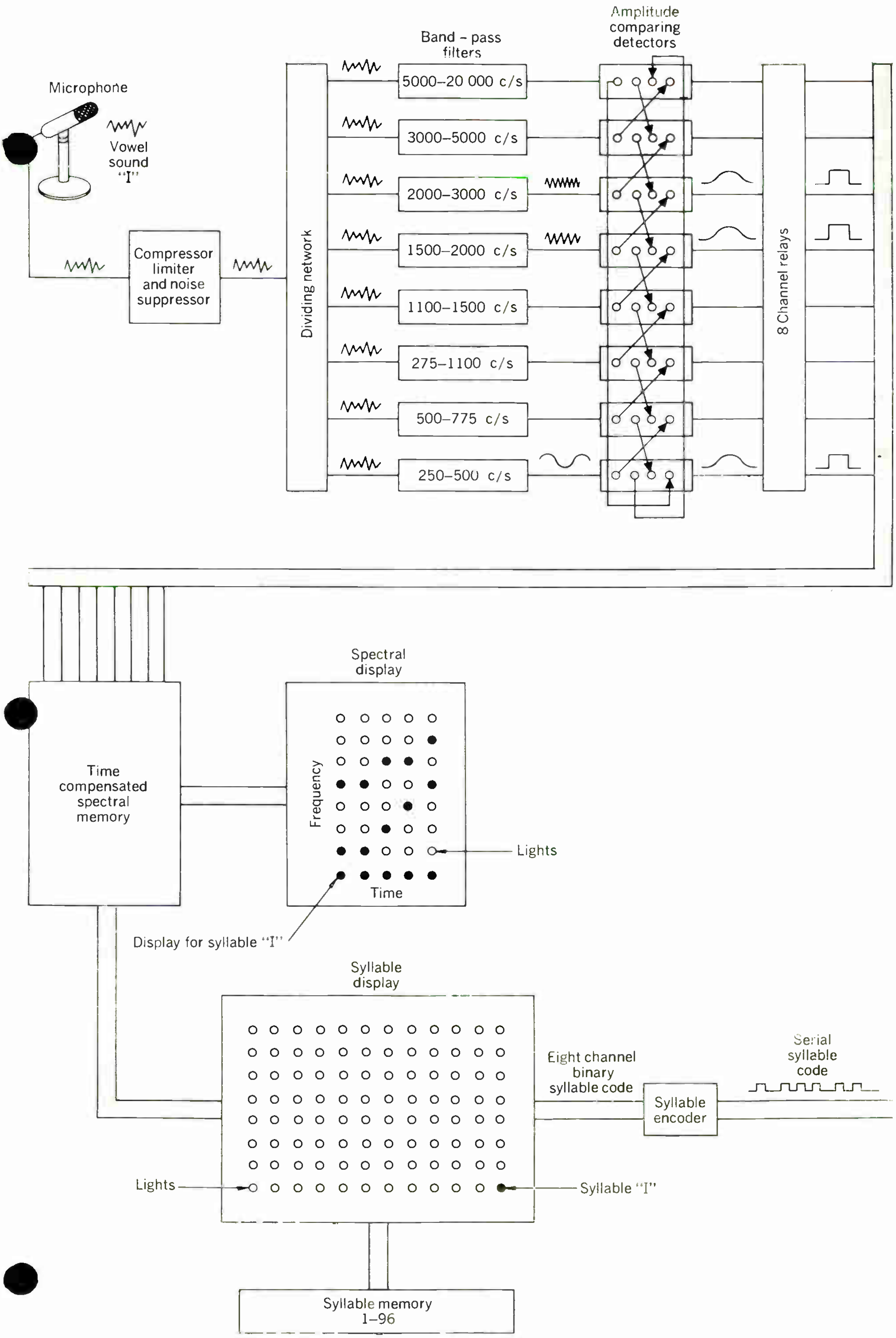


Fig. 8. A schematic block diagram of a speech-to-code converter. Elements of the system are: a specially designed microphone; normalizing input circuits; dividing network; band-pass filters; amplitude comparison circuits; spectral memory; syllable display and memory; and encoder.


Microphone
Vowel sound "I"
Compressor limiter and noise suppressor
Dividing network
Band - pass filters
5000–20 000 c/s
3000–5000 c/s
2000–3000 c/s
1500–2000 c/s
1100–1500 c/s
275–1100 c/s
500–775 c/s
250–500 c/s
Amplitude comparing detectors
8 Channel relays
Time compensated spectral memory
Spectral display
Frequency
Time
Lights
Display for syllable "I"
Syllable display
Eight channel binary syllable code
Syllable encoder
Serial syllable code
Lights
Syllable "I"
Syllable memory 1–96

tized spectral information varying with time as the syllable is spoken. The electric output of the amplitude-comparing detectors is in the form of a pulse (Fig. 8), which actuates the corresponding relay. The relays are connected to the spectral memory.

The spectral memory is time compensated for the speed of talking. The time compensation consists of 70 relays, which set up the first column in the frequency spectrum display and then transfer the spectrum information derived from the analyzer system in the form of an eight-bit code to the next column. There the system rests until there is a change. A significant change causes the second-column display to lock and the input is then stepped over to the third column, and so on. After the fifth column has been set up, the next and last change disconnects the input and the entire spectrum time display is held. A readout command is formed by operating a relay in the proper sequence followed by cancellation of the display and resetting to the initial position. Under these conditions, the pattern for each syllable will be the same regardless of the speed of talking.

The display, with its eight-bit spectrum and five time steps, provides a matrix of 40 fields. The lights of the spectral display indicate the amplitude–frequency–time pattern of the vowel *I*, as shown in Fig. 8, where the display is identical with the analysis of the vowel *I* in Fig. 6(B).

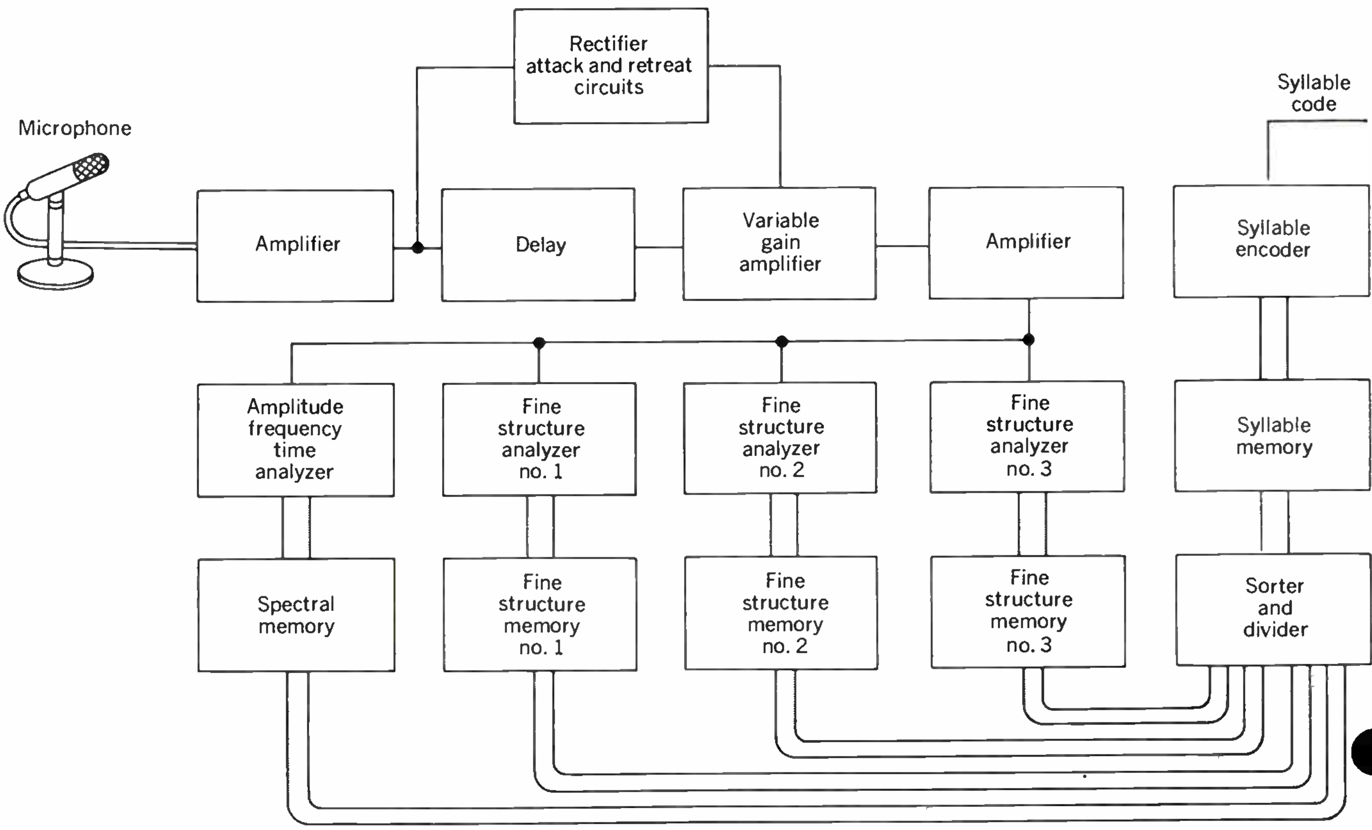
The spectral syllable display, namely, the sequence of quantized frequency spectrum information, is a 40-bit matrix; so every display corresponding to each individual voicing can be looked upon as a pattern or, simpler still, a binary number of 40 digits. This number, corresponding to the spectral display of the syllable, is fed to the syllable memories after the completion of each voicing. The syllable memories recognize the syllable display and produce a code identifying the syllable. The capacity of the syllable memory (Fig. 8) is 96 syllables and the syllable display accordingly includes 96 lights. The light corresponding to the vowel *I* is on in the syllable display.

The output of the syllable memory is connected to the syllable encoder, which converts the eight-channel binary code to a serial syllable code for transmission over a two-wire line.

The speech-analyzing system in Fig. 8 will not satisfy the practical considerations of an ultimate system. Experiments show that, because of the variations in frequency that occur when any one syllable is spoken, an amplitude-frequency-time analyzer with additional frequency steps will not supply additional information for analyzing speech. The net result obtained with an analyzer with more steps is several new patterns for the same syllable. Therefore, in addition to the amplitude-frequency-time analyzer described, fine-structure analyzers will be required to determine more of the information components of speech. A few of these fine-structure components of speech will now be briefly noted.

The envelope of a speech wave provides data on the information-bearing characteristics of speech. The data could be included in the spectrum-vs.-time analysis but

Fig. 9. A schematic block diagram of a speech analyzer employing three fine structure analyzers.



the envelope presentation is more convenient. The envelope provides information on the growth and decay of speech. Investigations show that the starting characteristics of a speech sound provide considerable information in discriminating between syllables with a similar sound. A stop consonant tends to have faster decay than one that ends in a vowel.

Measurements on the output of vocal cords of vowel sounds show that the wave is of the saw-tooth type with the negative portion somewhat flattened. Flattening on one side gives rise to an unbalance in the two directions, the unbalance maintained at the output of the mouth. An inspection of the bilateral envelopes of speech shows that the unsymmetrical nature of speech provides additional data on the information-bearing components of speech.

Early work on the synthesis of music and speech indicated the importance of the deviations from the mean to achieve artistic rendition of music and the natural quality of speech. Deviations in pitch or frequency are a case in point. There is fine structure of frequency deviation in speech above the more obvious deviations in pitch such as vibrato and other forms of frequency modulation. The fine structure is always present in the case of sound generators like the human voice in which the sounds are in the form of relaxation oscillations. The deviations in frequency in speech provide data on the information-bearing components of speech.

There are many other fine-structure characteristics of speech that may be employed to ferret out the information-bearing content of speech.

The amplitude-frequency-time analyzer is fundamental and will probably always be a part of speech analyzers in some form. Figure 9 shows a schematic block diagram of an analyzer employing the amplitude-frequency-time analyzer and three fine-structure analyzers. The fine-structure analyzers employ such characteristics as the envelope, growth, decay, dissymetry, frequency deviations, etc. The sorter and divider provide the logic for the differentiation and identification of the syllables. Outlining the operation of the system of Fig. 9 is beyond the scope of this paper but the illustration is presented to depict the general philosophy of the extended aspects of speech analysis. The output of the system is the syllable code, as in Fig. 8.

Code-to-type converter

The code-to-type converter consisting of the syllable decoder, spelling memory, typing control, and typewriter is diagrammed in Fig. 10. The syllable decoder converts the coded input to the appropriate syllable of the spelling memory. The spelling memory of the typing console has one spelling relay for each word. A particular relay corresponding to a syllable is operated by completing the circuit for it in the syllable memory to which it is connected. The spelling relays in the typing control are wired to bus wires representing the different keys of the typewriter through separate sets of relay springs corresponding to the spelling desired. A set of sequence bus wires is connected to every one of the relays. To prevent jamming of the typewriter, only one spelling relay can be operated at a time.

Operating of the typewriter feeds back a signal to the typing control unit, which causes a stepping switch to transfer to the readout portion of the second-sequence

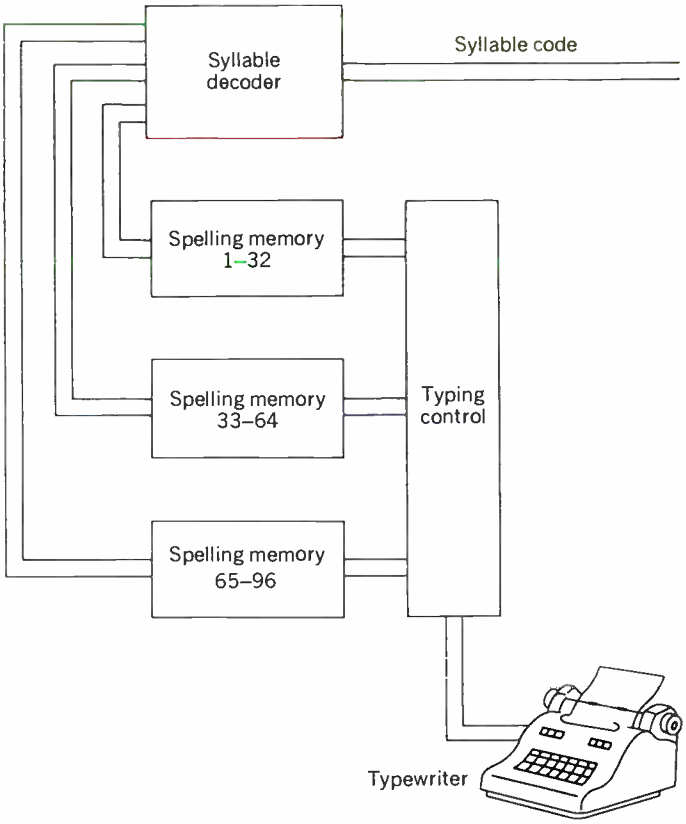
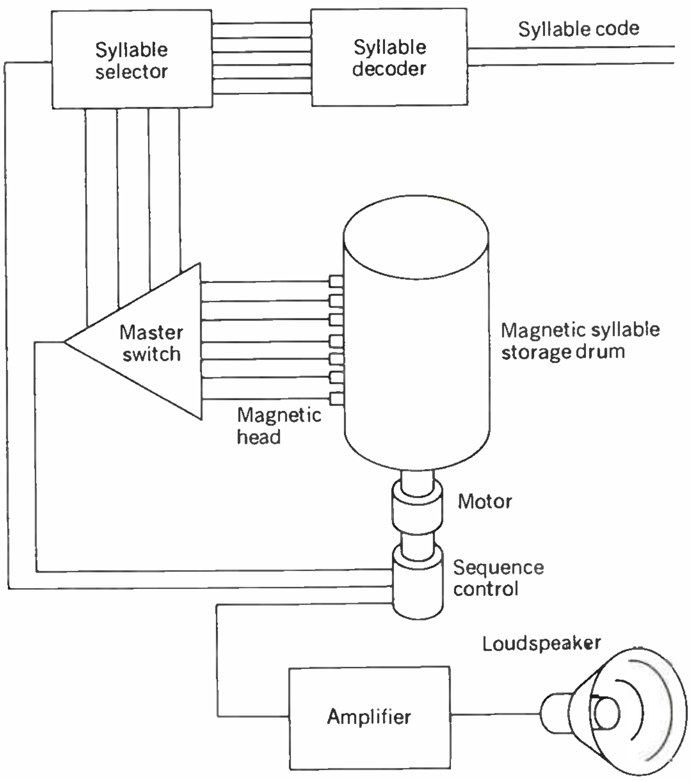


Fig. 10. A diagram of the coded typer that converts the syllable code input to the corresponding typed word.

Fig. 11. Speech synthesizer converts the syllable code input to the corresponding reproduced word.



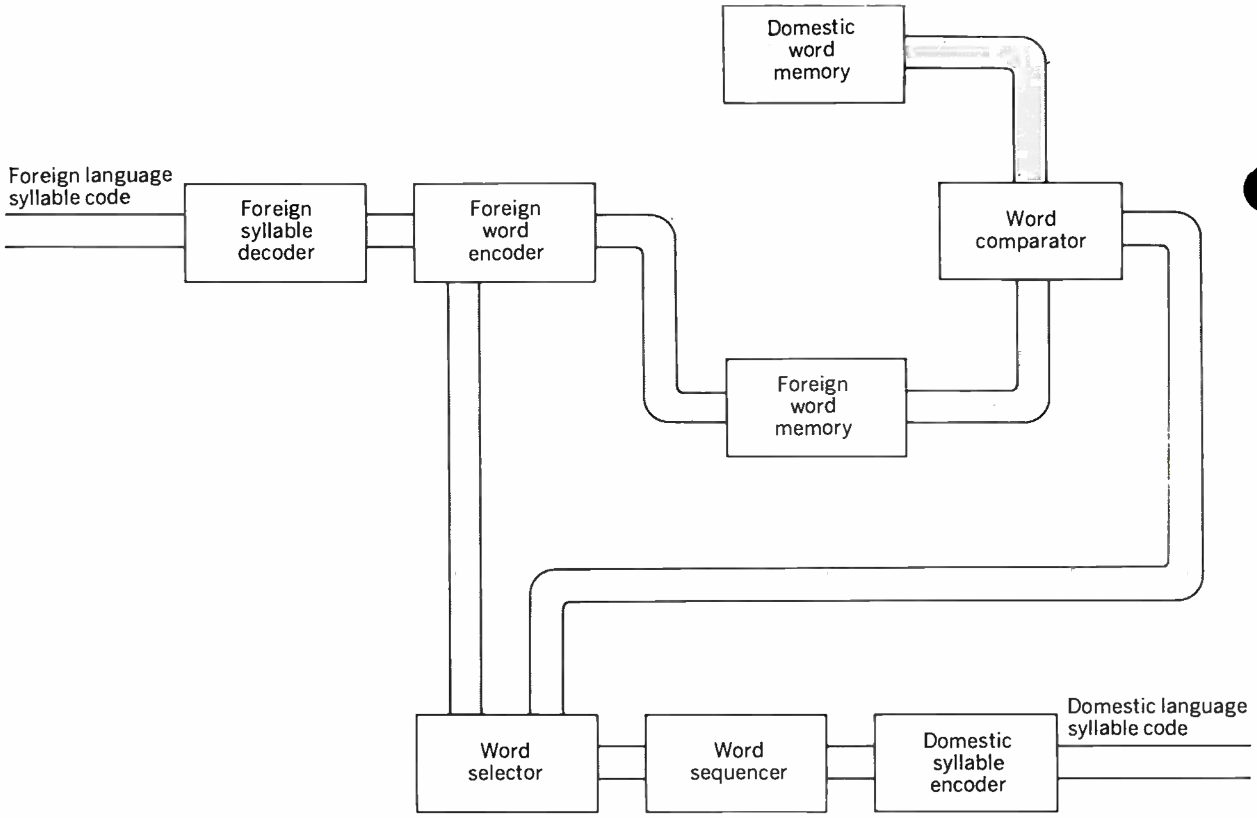


Fig. 12

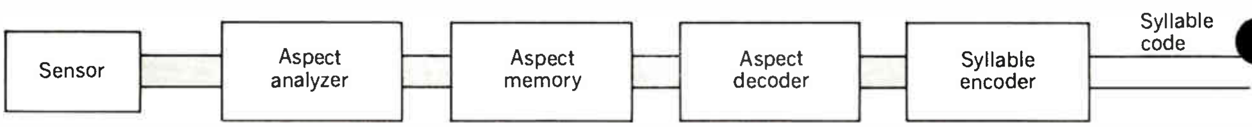


Fig. 13

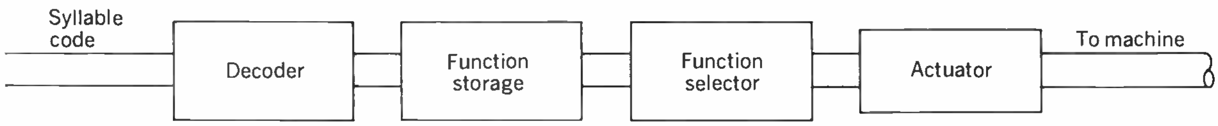


Fig. 14

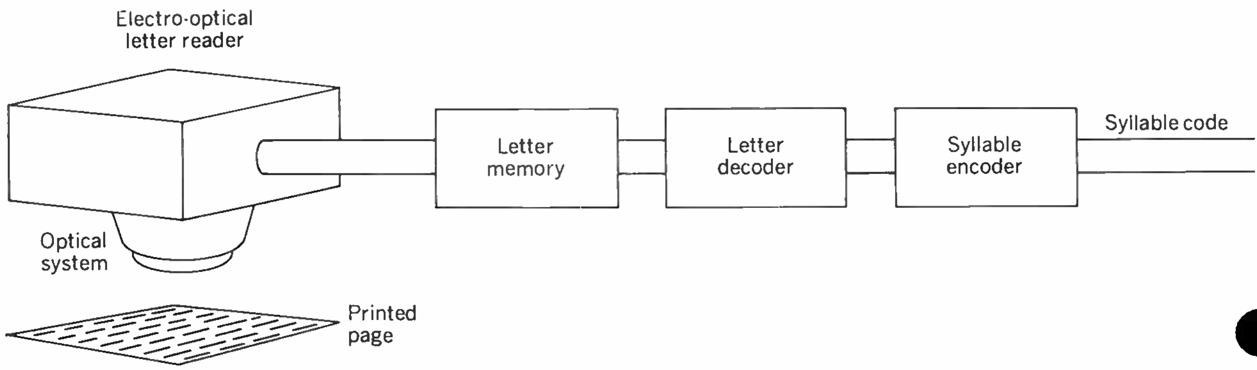


Fig. 15

bus. This action operates the typewriter again by connection to the desired spelling wired into the second set of contacts of the particular relay. A total of seven functions can be set up for each relay. The resetting of the spelling memory is automatic. The preceding discussion indicates that the syllable code input is converted to the typed syllable on paper.

Code-to-speech converter

Figure 11 is a diagram of the code-to-speech converter or speech synthesizer consisting of the syllable decoder, syllable selector, syllable storage, amplifier, and loudspeaker. The syllable code input is converted to reproduced speech from the loudspeaker. The syllable code is fed to the syllable decoder, which converts the coded input to the appropriate syllable of the syllable selector. The syllable selector operates the master switch that connects the appropriate magnetic head to the amplifier through the sequence control. The syllables are recorded on the magnetic drum, and stored syllables are reproduced by means of the magnetic heads. The sequence control completes the connection between the master switch and the amplifier in the appropriate interval; the loudspeaker converts the speech signal output from the amplifier into speech sounds.

High intelligibility can be obtained from the syllable synthesizer because the recorded voice can be selected to provide exceedingly clear enunciation of the different syllables.

Language translator

A simplified language translator for operation on the syllable code is diagrammed in Fig. 12. The input is the foreign language syllable code and the output the domestic syllable code. The foreign language syllable code is fed to the foreign syllable decoder which in turn is coupled to the foreign word encoder. The output of the foreign word encoder is fed to the foreign word memory and word selector. The foreign and domestic word memories are compared by the word comparator and the correct word is selected by the word selector. The word selector is fed to the word sequencer, which determines the correct sentence structure. The output of the word sequencer is then fed to the domestic language syllable encoder which in turn provides the domestic language syllable code output.

It must be remembered that translation from one language to another language is an extremely difficult task, particularly in the case of complex classical literature. But in the simpler types of literary works there are possibilities. Word-for-word translation, useful for simple commands and instructions, can be carried out without any great difficulty.

Fig. 12. Block diagram of a simplified language translation system. Input is the foreign language syllable code, and the output is the domestic syllable code.

Fig. 13. Elements of a system for the conversion of a particular aspect of a machine to a syllable code.

Fig. 14. Block diagram of the elements of a system for the conversion of a syllable code to produce force or motion.

Fig. 15. Elements of a system for the conversion of the printed page to a code.

Machine aspect to code converter

Conversion of a particular aspect or condition of a machine into a code can be useful, particularly in the case of automated complexes. The elements of a system for the conversion of a particular aspect of a machine to a syllable code are shown in Fig. 13. The sensor determines the machine's particular aspect, and the output of the sensor is analyzed. For example, if an analysis is wanted of amplitude as a function of time, the information is fed to the aspect analyzer, and its output is fed to the aspect memory, which is then coupled to the aspect decoder. A syllable encoder operates from the aspect decoder to produce the syllable code, and the appropriate syllable is then selected to express the particular aspect of the machine.

Code to mechanical actuator

Also in the case of automation complexes (and other applications), the conversion of a code to force or motion will be found useful. Fig. 14 shows the elements of a system for the conversion of a syllable code to force or motion. The input syllable code is fed to a syllable decoder; the function storage is coupled to the syllable decoder and function selector; and the output of the function selector operates the actuator.

Printed page to code converter

The conversion of the printed page to a code is useful for the input to computers and for the reproduction of speech from the printed page. The elements of a system for the conversion of the printed page to a syllable code are given in Fig. 15. The output of the electro-optical letter correlator provides the letter that has been read and is coupled to the letter memory. The letter decoder is coupled to the letter memory and the syllable encoder. The syllable encoder derives the syllables from the letter decoder and converts the syllables to a code.

Print readers that will read the printed page have been developed and are in use. The reduction of the output to a syllable code as just outlined appears to be possible with the means available today.

Complete processing systems

The systems for converting information to and from a code, and outlined in the preceding sections, can be combined to provide many new facilities in the field of communications. A few examples of the many interesting possibilities follow.

Code transmission system. A code transmission system for speech reproduction consists of the speech-to-code converter of Fig. 8 and the code-to-speech converter of Fig. 11. The advantage of such a system is the small frequency bandwidth required for the transmission of speech; for example, the bandwidth corresponds to only 23 bits per second for the transmission of speech at normal rates of talking.

Phonetic typewriter. A phonetic typewriter consists of the speech-to-code converter of Fig. 8 and the code-to-

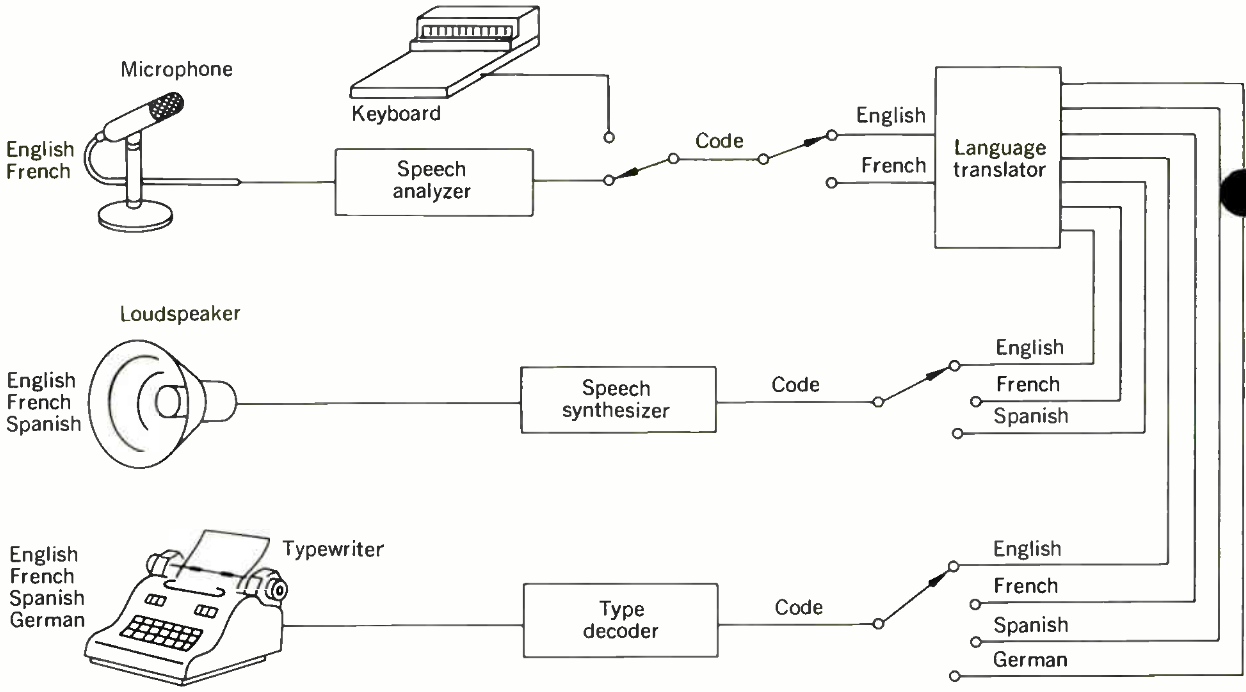
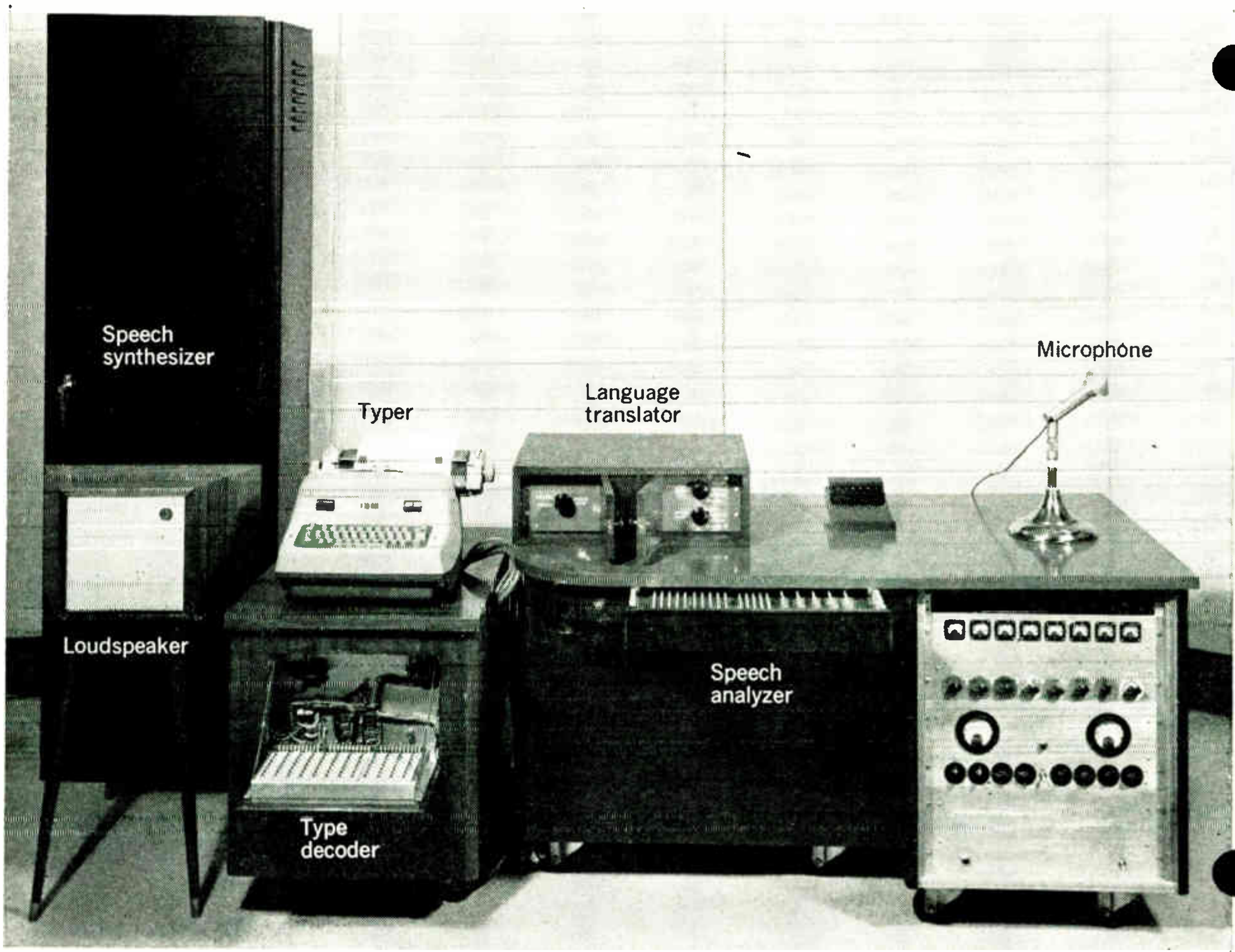


Fig. 16. A schematic block diagram of speech-processing demonstration equipment depicting the elements, and the input and output functions.

Fig. 17. A view of the various components of the speech-transmission equipment used for demonstration purposes.



typer of Fig. 10. The speech input to the microphone is converted to the typed page.

Speech-controlled machine. The speech control of machines is useful for augmenting or supplanting the use of the hands and feet in the operation of a machine. A speech machine controller consists of the speech-to-code converter of Fig. 8 combined with the code-to-actuator of Fig. 14.

Machine-sensor speech-reproducer. Speech sounds that announce the aspects or conditions of a machine or operation are useful when the operator cannot obtain the information by visible means. A machine-sensor speech-reproducer consists of the aspect-to-code converter of Fig. 13 and the code-to-speech converter of Fig. 11.

Printed-page speech reproducer. A printed-page speech reproducer consists of the page-to-code converter of Fig. 15 and the code-to-speech converter of Fig. 11. Such a system provides means for speech reproduction of the printed page.

Language translators. A language translator providing for foreign speech input to the microphone and domestic speech output from the loudspeaker consists of the speech-to-syllable code converter of Fig. 8, the simplified language translator of Fig. 12, and the code-to-speech converter of Fig. 11.

Another type of language translator providing for foreign speech input to the microphone and domestic typed-page output consists of the speech-to-syllable code converter of Fig. 8, the simplified language translator of Fig. 12, and the code-to-typer of Fig. 10.

A third language translator providing for foreign printed-page input and domestic speech output from the loudspeaker consists of the printed-page to code converter of Fig. 15, the foreign language syllable input to domestic language syllable output translator of Fig. 12, and the syllable-code to speech converter of Fig. 11.

Still another version provides for the foreign printed-page input to the domestic printed-page output.

Demonstration equipment

In the development of almost every new art, complete systems have been built in the various stages of program development. As one looks back upon such systems, they seem crude and rudimentary compared to the polished and finished commercial systems. Nevertheless, the developmental system serves as a guide to the nature of the problems to be solved. It was with general philosophy in mind that a demonstration equipment of a speech-processing system was developed.

Figure 16 gives the elements and the input and output functions of the speech - transmission demonstration equipment. Figure 17 shows the demonstration equipment with labels indicating the various elements making up the complete system.

When the system of Fig. 16 operates as a speech-transmission system, speech is converted to a syllable code by means of the speech analyzer of Fig. 8, then transmitted as a syllable code and synthesized from the syllable code to speech by means of the speech synthesizer of Fig. 11. In this way, each word spoken into the microphone is reproduced as the same word in the same language or translated and reproduced as the corresponding word in another language. The advantage of the system is the extremely small capacity of the channel required for the transmission of speech. Only 23 bits per second are required to transmit speech at normal rates of talking. For optimum coding on a channel of 20 dB signal-to-noise ratio the bandwidth required is 3.5 c/s.

When the system of Fig. 16 operates as a speech to printed page transmission system, speech is converted to a syllable code by means of the speech analyzer of Fig. 8, then transmitted as a syllable code and converted to the typed page from the code by means of the type decoder of Fig. 10. In this way each word spoken into the microphone is converted into the same typed word on paper or translated and converted into the corresponding typed word on paper in another language.

When the system of Fig. 16 operates as a coded-speech reproducer, the syllable code is manually produced by means of the keyboard, transmitted as a syllable code, and reproduced by means of the speech synthesizer of Fig. 11.

The language translation is carried out on a word-for-word basis. This of course means that meaningful translation can occur only for phrases and sentences which are amenable to such procedures. However, with this type of translation simple commands, directions, and explanations can be carried out.

The translator used for converting the syllable code in one language to the syllable code in another is a switching system that does the converting for operating the speech synthesizer or code typer.

I. Translated words

Voice input		Print output				Voice output		
English	French	English	French	German	Spanish	English	French	Spanish
stop		stop	halte	halt	alto	stop	halte	alto
I		I	je	Ich	yo	I	je	yo
see		see	vois	sehe	veo	see	vois	veo
six		six	six			six		
tanks		tanks	tanques	panzer	carros	tanks	tanques	carros
	oui	yes	oui			yes		
	je	I	je	ich	yo	I	je	yo
	vois	see	vois	sehe	veo	see	vois	veo
	six	six	six			six		
	tanques	tanks	tanques	panzer	carros	tanks	tanques	carros
	halte	stop	halte	halt	alto	stop	halte	alto

Demonstrations

Demonstrations of the system depicted in Fig. 16 involve these English words that can be processed, transmitted, typed, and reproduced as speech:

a	is	sells	speak
ah	it	sends	spell
are	left	shall	stop
burns	new	she	sure
by	no	shells	tanks
can	oil	shore	thanks
come	please	should	the
Dear	read	Sir	you
Earl	school	six	york
I	see	space	yours

and these French words: halte, je, oui, six, tanques, vois.

English words or parts of words that can be processed, transmitted, and typed are: Gen., Sar, noff, Mr., Ol, son.

Translated words: Additional words by translation are provided as depicted in Table I.

The capacities of the Fig. 16 system are summarized:

It can understand 52 words in two languages: 46 English and 6 French.

It can print out from speech input 62 words in four languages: 48 English, 6 French, 4 German, and 4 Spanish.

Its speech output from speech input is 50 words in three languages: 42 English, 4 French, and 4 Spanish.

Its maximum transmission rates are: speech input to print output, 60 syllables per minute; and speech input to speech output, 30.5 syllables per minute.

Performance

The accuracy of the machine is another important consideration. If an average of 200 voicings of a syllable or word are made to establish the code, the accuracy will be 98 per cent. The code of each syllable or word used in the demonstration was set up from more than 200 voicings by H. B. (of a given national origin), in which case under normal operation conditions 98 per cent accuracy of the system was obtained. With the code thus set up for H. B., the score obtained by R.d.S., of a different national origin from H. B., was 98 per cent for 20 words. For the remainder of the words there was a decreasing reliability ranging from 96 per cent to almost zero. With the code set up for H. B., the score obtained by G. S., of a national origin different from either H. B., or R.d.S., was 98 per cent for 17 words. For the remainder of the words there was a decreasing reliability ranging from 96 per cent to almost zero. In view of the three different national origins, the high scores for R.d.S. and G.S. for a code set up for H.B. are considered remarkable. Of course, when the code for R.d.S. was established from 200 voicings, the accuracy obtained was 98 per cent for him, and similar results were obtained by G.S. when the code was established for him.

The system results today for others than those for whom the code was established show great improvement over the results obtained three years ago when the code was for all practical purposes almost a personal one.

Conclusions

Now in the early stages of development are new speech-processing systems for the transmission of speech. Progress has been made toward controlling machines, producing the typewritten page, sensing a machine, receiving reproduction of the printed page, and translating languages. These systems have been demonstrated but to make them a complete practical reality will require considerable research and development in these various areas:

1. Research in the analysis of speech to determine the information-bearing components for complete and universal systems has advanced to the stage where systems of limited capabilities will soon be in use, but considerably more research is required.

2. The synthesis of speech from a code and a means of syllable storage has been demonstrated to be a practical and straight-forward process.

3. Print readers that read the printed page have been developed and are now in use. The reduction of the output to a syllable code appears to be possible with the means available today.

4. Machine actuators that operate from a code and machine sensors that convert a particular condition to a code have been developed for other applications. The extension to speech processing appears to be merely a more complex problem.

5. Typers that produce the printed page from a code are in use. The application for the phonetic typewriter appears to present no problems that cannot be solved with existing knowledge.

6. The development of language translators that will operate on all types of composition is an exceedingly difficult problem. In the case of the complex classical literature, machine translation may never become an accomplished fact, but in simpler types of literary works machine translation is indicated. Word-for-word translation, useful for simple commands and instructions, can be carried out with no particular problems or difficulties.

As analyzing and synthesizing systems advance, the various applications for the speech-processing systems will become a reality. That they will is no longer in doubt; developments have passed beyond the questionable stage.

BIBLIOGRAPHY

Olson, Harry F., *Acoustical Engineering*, Princeton, N.J.: D. Van Nostrand Co., Inc., 1957.

—— and H. Belar, "Phonetic Typewriter," *J. Acoust. Soc. Amer.*, vol. 28, Nov. 1956, pp. 1072–1081.

—— and ——, "Phonetic Typewriter," *IRE Trans. on Audio*, vol. AU-5, July–Aug. 1957, pp. 90–95.

—— and ——, "Time Compensation for Speed of Talking in Speech Recognition Machines," *Ibid.*, vol. AU-8, May–June 1960, pp. 87–90.

—— and ——, "Syllable Analyzer, Coder and Synthesizer for the Transmission of Speech," published in the *Proceedings*, 1961 National Aerospace Electronics Conference (NAECON), Dayton, Ohio.

—— and ——, "Syllable Analyzer, Coder and Synthesizer for the Transmission of Speech," *IRE Trans. on Audio*, vol. AU-10, Jan.–Feb. 1962, pp. 11–17.

—— and ——, "A Print-out System for the Automatic Recording of the Spectral Analysis of Spoken Syllables," *J. Acoust. Soc. Amer.*, vol. 34, Feb. 1962, pp. 166–171.

—— and ——, "Phonetic Typewriter III," *Ibid.*, Nov. 1961, pp. 1610–1615.

——, —— and R. deSobrino, "Speech Processing Systems," *Ibid.*, vol. 34, May 1962, p. 724.

—— and ——, "Recognition of the Spoken Word by Machine," in *Biological Prototypes and Synthetic Systems*, vol. 1, New York: Plenum Press Inc., 1962.

——, ——, and ——, "Demonstration of a Speech Processing System Consisting of an Analyzer, Translator, Typer and Synthesizer," *J. Acoust. Soc. Amer.*, vol. 34, Oct. 1962, pp 1535–1538.

Superconducting magnets

The advent of high-field superconducting materials has prompted the exploration of applications, including dc magnets, energy storage, and the production of homogeneous fields

D. Bruce Montgomery *National Magnet Laboratory*

Ever since the discovery of high-field superconducting materials, considerable attention has been focused on their practical uses. The possibility of carrying lossless currents at high magnetic fields suggests such diverse applications as magnetohydrodynamic generators, bubble chambers, energy storage, spaceship ion motors, radiation shields, small laboratory magnets, and the full range of ac devices, including transformers, motors, generators, and transmission lines.

In the few years since the properties of the two principal materials—niobium-zirconium and niobium-tin—were announced, there has been a great deal of success in their production and application. For example, high-quality Nb-Zr wire is now available in single lengths to 10 000 feet or more, and reliable laboratory solenoids can be purchased in sizes up to 60 kilogauss. Several very large magnets involving many tens of pounds of wire have been built and others are under construction or being planned.[1] In addition, a number of high-field super-homogeneous and superstable coils have been built for nuclear resonance work. Niobium-tin is a higher-field material but is more difficult to handle. The most significant achievement was made recently when generation of 101 kilogauss was accomplished by the use of about 600 feet of Nb_3-Sn wire.[2] Many coils up to 70 kilogauss have also been built.[3,4] Several manufacturers are producing significant lengths of Nb_3-Sn ribbon, which can be wound into coils after heat treatment (unlike the standard wire, which must be reacted after winding).

Perhaps the major disappointment has been the poor ac characteristics of the high-field superconductors. When exposed to an alternating field, they no longer operate in a lossless way. In fact, the losses are so high that the refrigerator power necessary at 4.2°K would be much larger than the power consumption for ordinary conductors at room temperature. A second disappointment has been the low current-carrying capacity of Nb-Zr wire when it is wound into coils. This degradation is caused by instabilities of the diamagnetic current in the material.[5-8] There has been some evidence that degradation is more severe in large coils. The degradation has not been apparent to such a degree in the Nb_3-Sn coils built to date, but it may play an increasing role as larger coils are built.

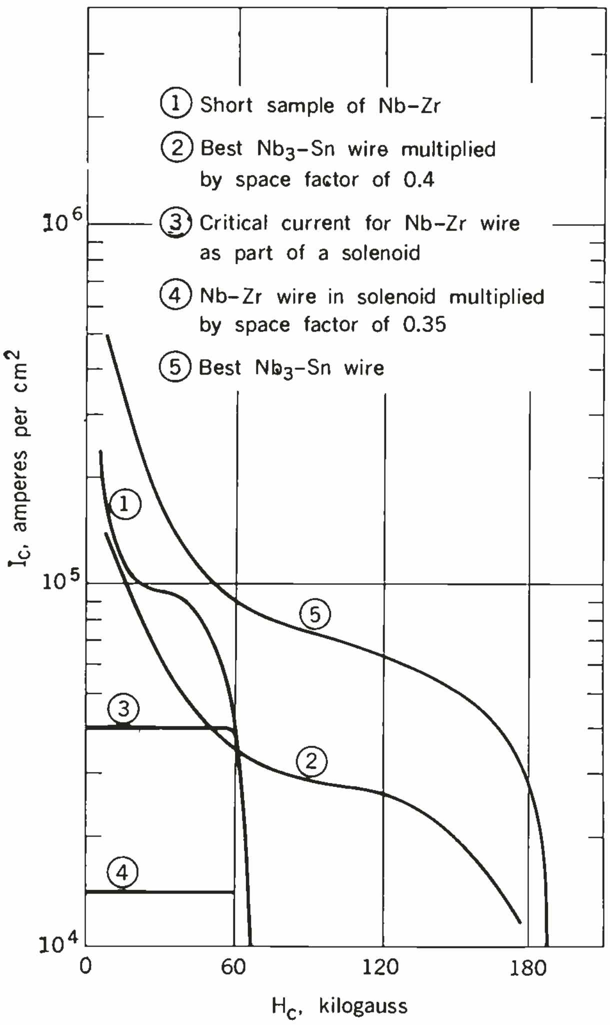
Wire characteristics

Critical current, critical field. The current that a given superconducting wire will carry is dependent upon the magnetic field to which it is exposed. This current is referred to as the critical current I_c, and the curve of current versus applied field H_c is usually called the critical current–critical field curve. The highest field at which any trace of superconductivity remains is referred to as the upper critical field. In general, the upper critical field is a function of the material and is independent of the metallurgy and subsequent wire drawing. The critical current, on the other hand, is strongly dependent on metallurgy, impurities, heat treatment history, and other factors. This dependence gives rise to a rather large spread of I_c-H_c curves for the same material. Typical I_c-H_c curves for the best available Nb-Zr and Nb_3-Sn are shown in Fig. 1, curves 1 and 5.

There is further variation in the performance of the wire when it is operated in a self-generated field, as in a solenoid. Niobium-zirconium in particular shows an I_c-H_c curve that is almost independent of H up to about 50 kilogauss (Fig. 1, curve 3). Niobium-tin wire has given little evidence of this degradation effect when used with a fairly low ratio of Nb_3-Sn to the surrounding niobium, as in the cored wire. However, the coils built with the new ribbons of deposited Nb_3-Sn on a thin substrate of niobium have given evidence that degradation may also play a part in niobium-tin coils, especially at low fields.

The degradation in Nb-Zr coils is more apparent as the coils become larger and the total stored energy becomes greater. A small, 1-cm-bore magnet generating 50 kilogauss may carry 25 amperes, whereas a 25-cm-bore coil at 50 kilogauss will not carry more than 15 amperes. This size dependence is neither understood nor universally recognized. It is expected that larger Nb_3-Sn coils will also be more subject to degradation.

Fig. 1. Critical current density vs. critical field for Nb-Zr and Nb_3-Sn wire (referred to entire cross section of wire, exclusive of copper cladding or insulation).



The degradation effect. The basic cause of degradation lies in the instability of the shielding currents that are set up in the superconducting material as the magnetic field is changed. These shielding currents form to prevent flux from penetrating the wire and flow first at the surface of the wire until the local allowable current density is exceeded. The currents then flow deeper in the wire until the flux has completely penetrated the wire and maximum current is being carried at all points. Figure 2 shows the resultant diamagnetic shielding effect of these currents as the field is raised, and shows the magnetic trapping effect as the field is reduced and the currents trap flux in the wire. If we start with a virgin sample at zero field, the currents form and provide perfect shielding as we move from a toward b in Fig. 2. At about 800 gauss, the allowable surface current density is exceeded and the flux begins to penetrate. At about 2.5 kilogauss, the flux has completely penetrated and, by definition, all parts of the wire are carrying critical current. (If the current were not everywhere at the maximum current density allowed, more flux could be shielded.) Any further increase in field must now result in a decrease in magnetic moment, as the allowable current density is a decreasing function of field; see Fig. 1.

If the sample had been carried to a sufficiently high field, all magnetic moment would have disappeared. If we stop at any point along the curve and introduce some current into the wire from an external source, the magnetization decreases in proportion to the transport current. Since transport current simply subtracts from the shielding currents, no degradation should result from the presence of the additional shielding currents. However, the penetration of the flux and the formation of shielding currents are not smooth, continuous processes. The flux tends to be pinned on dislocations in the wire; and as local pinning forces are exceeded, the flux avalanches.[9] This sudden release of stored magnetic energy goes into heating the material, and under certain conditions, can be sufficient to raise the material locally above the critical temperature. The local "quench" is then propagated by joule heating from the transport current that flows in the normal area. These sudden, discontinuous movements of flux are called flux jumps, and it is their interaction with the transport current that gives rise to degradation. The largest flux jumps take place where the circulating currents are largest (point b in Fig. 2). Therefore, the quench in solenoids would be expected to originate where the circulating currents are high and where the transport current has reached sufficient magnitude to propagate the quench. This point would tend to be in the interior of the coil, where the field is less than 20 kilogauss at the time the transport current density has reached 4×10^4 amp/cm^2.

Several factors work toward decreasing the degradation effect of flux jumps in Nb_3-Sn. The transition temperature of Nb_3-Sn is higher than that of Nb-Zr. Therefore, a given release of magnetic energy is less likely to raise the material above the critical temperature. In addition, the Nb_3-Sn phase is accompanied by a considerable volume of inactive niobium, which acts as a local heat sink. Finally, the Nb_3-Sn on wires or ribbons is generally in the form of very thin layers—an inherently more stable configuration. In contrast, a sintered 1-cm-diameter rod of Nb_3-Sn shows very severe flux jumps.

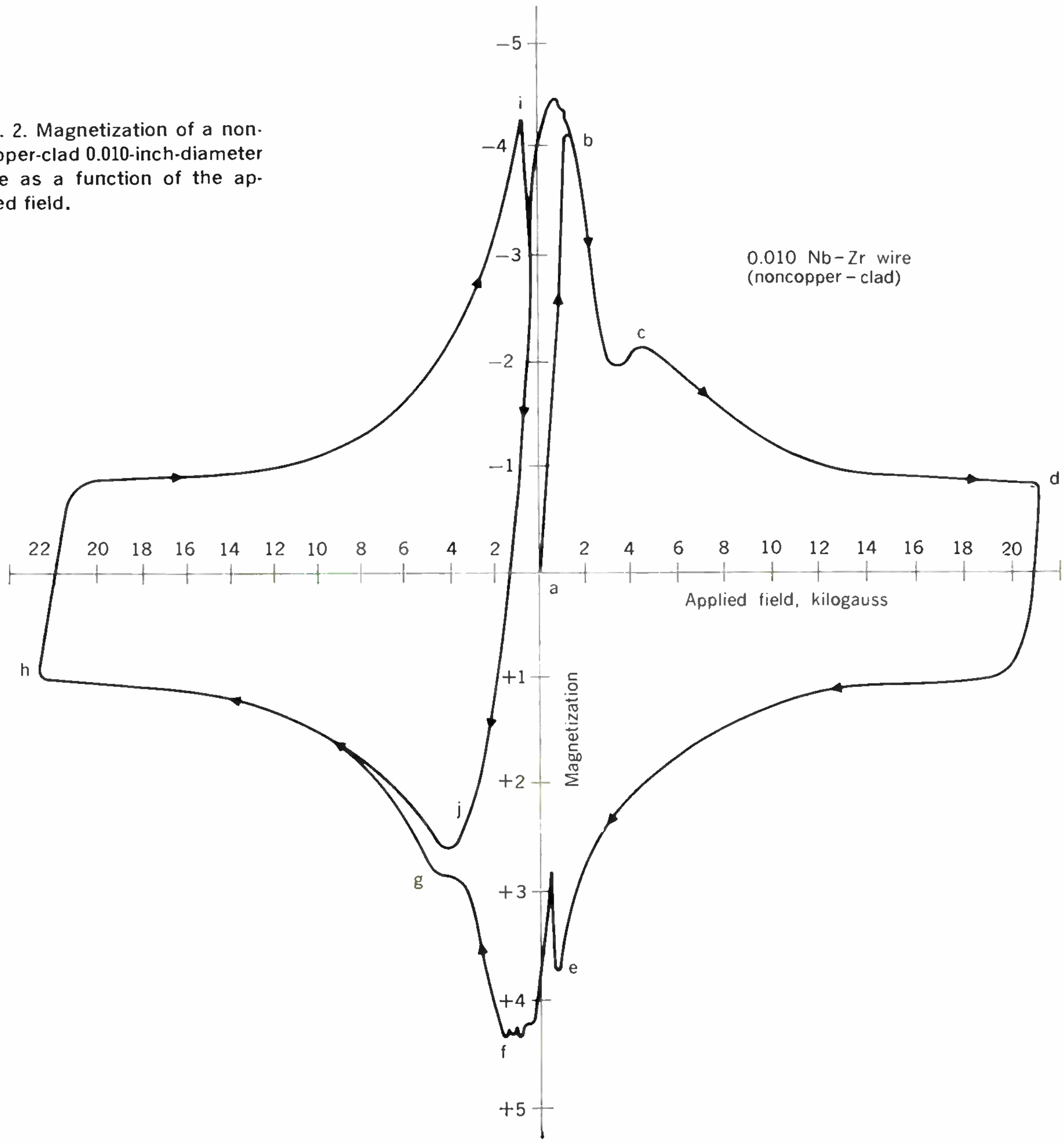
Degradation beyond that usually observed, especially in larger coils, has been mitigated by the use of copper cladding on the superconducting wire. The cladding acts as an eddy-current shield against the turn-to-turn spread of flux jumps. The copper jacket also serves the important function of providing a parallel path for the transport current during a quench and prevents large voltages from building up.

There are two possible approaches that might eliminate degradation. If coils are operated in an external field high enough to depress the large shielding currents that form at low fields, less violent flux jumps occur. For example, small coils have been operated in an external field of 20 kilogauss nearly without degradation.[10] This technique is obviously not practical for large systems, however. A second possibility is to improve the internal cooling of the coils by means of a forced helium flow in a channel structure. This has not been tried, but the converse has been noted—namely, that small coils operated in vacuum and cooled only by radiation are plagued by flux jumps more often than are immersed coils.

Mechanical properties and typical constructions. Niobium-zirconium is a ductile alloy with superior strength characteristics. The wire generally used is 0.010 inch in diameter, with a 0.001-inch cladding of copper, and a 0.001-inch coating of epoxy. Most coils are wound with 0.002 inch of mylar between layers. Large coils often include additional short-circuited copper foils between layers to dissipate the magnetic energy during quenching.

The best contacts are made by swaging carefully cleaned wire into copper tubes. The wires are sometimes indium-tinned. If Nb-Zr–copper contacts are located in a low-field area, spot welding can be used by welding first to a platinum ribbon and then soldering. Superconducting

Fig. 2. Magnetization of a non-copper-clad 0.010-inch-diameter wire as a function of the applied field.



joints can be made by spot welding two Nb-Zr wires, but the joints must be kept in fields of less than a few thousand gauss. A good discussion of contacts is given by Ralls *et al.*[11] Most coils are wound without potting, but care must be taken to restrain the leads between the coil proper and the contact terminals.

Niobium-tin is a brittle intermetallic alloy, and thus the difficulty of winding it has discouraged its widespread use. The original wire used was a Monel-jacketed niobium tube filled with niobium and tin powders, drawn down to a 0.015-inch OD and a 0.006-inch core, wound in place and then reacted at 1000°C.[3] The wire did not have turn-to-turn insulation, but was provided with quartz cloth between layers. Many coils have been built with this technique, but the effective space factors are so low that fields of only about 70 kilogauss have been generated. Evidence suggests that most of the current is being carried at the interface between the core and the tube wall.[12]

Recently a wind-and-react magnet that achieved a field of 101 kilogauss was built by the General Electric Company.[2] The wire has not yet been described in detail, but it involved dipping niobium wires in tin, and reacting the wires after winding. The ceramic insulation used was of a type that did not involve any harmful excess of oxygen during reaction.[4] The current densities reported by General Electric are considerably higher than any previously reported. The coil was wound without adequate protection and was damaged during the first quench. Future coils will need to sacrifice part of the space factor to allow for protective windings.

Several manufacturers are producing bendable Nb_3-Sn either by depositing Nb_3-Sn onto thin substrates or by diffusing tin into thin niobium ribbons. Then material can be bent gently (3/8-inch radius) after reaction without serious loss of current-carrying capacity. A typical ribbon is 0.001 inch thick by 0.062 inch wide and would have a current density (referred to the whole wire) of 4.3×10^4 amp/cm² at 100 kilogauss. These ribbons are often copper plated for additional protection. Small coils using ribbon have been built by the National Research Corporation, and RCA is employing ribbon in the development of some relatively large (2-inch ID) pancakes. At least one additional company—Materials Research Corporation—has ribbon under development.

A recent technique in producing composite wire has been under study by Saur.[13] A bundle of tin-dipped niobium wires is placed in a niobium tube and the tube is vacuum impregnated with tin. Then it is drawn to about 0.020 inch OD. The wire must be reacted in place. On the basis of preliminary results this technique appears to be promising. Contacts to Nb_3-Sn can be made either by ultrasonically indium-tinning or by copper-plating the material, and then soldering it into copper blocks.

Design of solenoids

Field relations. For the purpose of exploring the design considerations for superconducting solenoids we will assume the product of current density and space factor that is shown in Fig. 1: the constant value of 1.4×10^4 amp/cm² for Nb-Zr and the variable function passing through 2.8×10^4 amp/cm² at 100 kilogauss for the best reported Nb_3-Sn.[2] The field in the center of a uniform, wound solenoid can be written

$$H = j\lambda a_1 F(\alpha, \beta) \text{ gauss} \tag{1}$$

where

j = current density in the conductor, amp/cm²

λ = volume of active conductor/volume of coil

a_1 = inner radius, cm

$$F(\alpha, \beta) = \frac{2\pi}{5}\beta \ln \frac{\alpha + (\alpha^2 + \beta^2)^{1/2}}{1 + (1 + \beta^2)^{1/2}} \tag{2}$$

where

$\alpha = \dfrac{a_2}{a_1}$ a_2 = outer radius, cm

$\beta = \dfrac{b}{a_1}$ $2b$ = length of solenoid

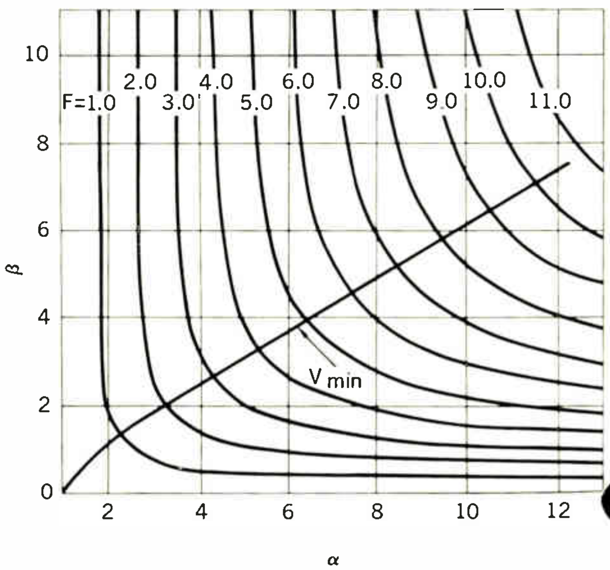
Curves of $F(\alpha, \beta)$ and the line of minimum volume V_{min} are shown in Fig. 3. Values of α and β chosen along this line will give the shape of coil that achieves the maximum field per unit volume of conductor and thus represent the most efficient family of coil shapes. Because of size limitations of Dewar vessels, however, it is often necessary to use a smaller α and larger β to achieve the required value of F. The minimum-volume line can be closely approximated by the following equation:

$$5\beta = 3\alpha + 1 - \frac{4}{\alpha^2} \tag{3}$$

Where the current density is a function of magnetic field, as with Nb_3-Sn, a considerable reduction in the necessary volume of material can be made by increasing the current density as a function of radius. This can be done either by changing wire sizes or by using nested coils with separate power supplies. A discussion of the optimization procedure and the potential savings is given by Gauster.[11] Niobium-zirconium coils have also been built using higher Zr content wire in the outer sections, where the higher critical current but lower critical field characteristics of high Zr content wire can be used advantageously.[15]

It is often necessary to design split-coil systems for

Fig. 3. $F(\alpha, \beta)$ curves for rectangular coils with constant current density. The V_{min} line represents the relation between α and β for maximum field and minimum wire.



transverse experimental access. It is useful to note that the central field can be written by superposition using Eq. (1):

$$H = j\lambda a_1[F(\alpha, \beta_1) - F(\alpha, \beta_2)] \tag{4}$$

where

$\alpha = a_2/a_1$

$\beta_1 = L/2a_1 \qquad L$ = total length of coils and gap

$\beta_2 = g/2a_1 \qquad g$ = gap length

It is also useful to note that the field along the axis of a uniformly wound solenoid (assuming the solenoid to be coaxial and centered at the origin) can be written

$$H(x, 0) = \frac{2\pi}{10} j\lambda a_1 \left[\gamma_1 \sinh^{-1}\frac{\alpha}{\gamma_1} - \gamma_1 \sinh^{-1}\frac{1}{\gamma_1} + \gamma_2 \sinh^{-1}\frac{\alpha}{\gamma_2} - \gamma_2 \sinh^{-1}\frac{1}{\gamma_2} \right] \tag{5}$$

where

$\gamma_1 = (b + x)/a_1$

$\gamma_2 = (b - x)/a_1$

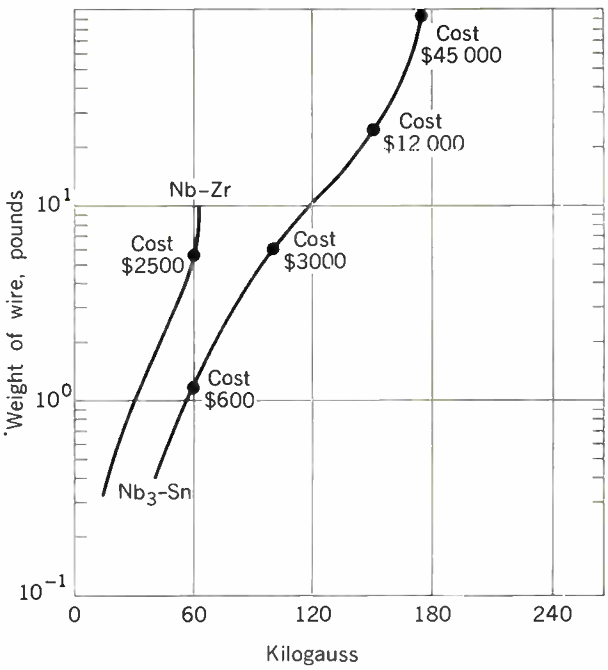
$\alpha = a_2/a_1$

On the basis of the few split coils that have been built, there is evidence that they tend to have a low current-carrying capacity. This should be properly investigated with small coils before proceeding with large coils.

Economics of solenoid design. The wire costs for Nb-Zr are shown in Fig. 4 and are plotted as a function of field and inside diameter. The coils chosen all lie on the minimum-volume lines of Fig. 3 and represent the most efficient use of conductor. The current densities are based on Fig. 1, and the wire cost is taken at \$460 a pound of Nb-Zr (exclusive of the weight of the copper cladding).

As a comparison of costs, the size of the power supply necessary to produce the same field volume with copper magnets is given for several points. Large, well-controlled dc power supplies cost about \$100 per kilowatt of power. Commercial superconducting solenoids sell for about three times the wire cost. There is the additional cost of the Dewar vessels and, in the case of large systems, the probable cost of a refrigerator. The larger sizes of coils, even including the cost of a refrigerator, are more economical than copper magnets. The smaller sizes, even in the area normally covered by laboratory-type iron magnets, can be economical if the entire cost of the refrigerator need not be included. Liquid helium is available in 50-liter containers almost everywhere now, and it

Fig. 5. Weight of superconducting wire required to generate a given field in a 1-inch-bore coil. The best reported Nb_3-Sn is compared with conventional Nb-Zr.



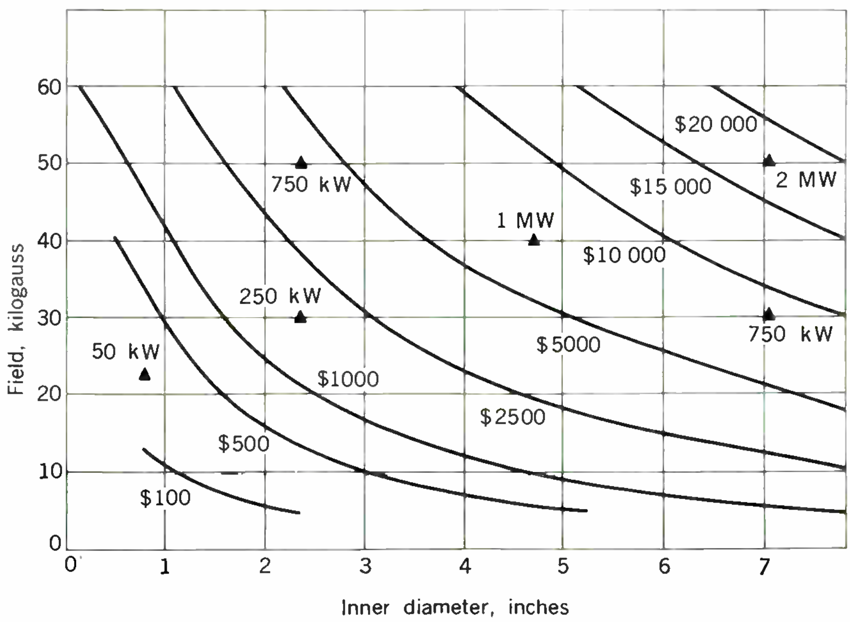


Fig. 4. Niobium-zirconium wire cost as a function of the magnetic field and inner diameter. Selected points show the power required to generate the same product of field and bore with conventional copper windings.

is usually unnecessary to amortize a refrigerator in the cost of a small coil.

Superconducting coils are also more economical than cryogenic nonsuperconducting systems, especially for continuous or high-duty-cycle use. However, if fields higher than that reachable by the superconductors must be generated, or in special cases where the field must be changed rapidly, other systems must be used.

Figure 5 illustrates the economics of building Nb_3-Sn coils, and contrasts the cost of 1-inch-bore solenoids of various fields made from Nb-Zr and those made of the best reported Nb_3-Sn. The cost of Nb_3-Sn wire is assumed to be \$500 per pound, but no account is taken of the additional expense of winding or heat treating. The curves indicate that it is more attractive to use Nb_3-Sn even below the critical field of Nb-Zr if the difficulties of winding can be solved economically, and if current densities as high as the best reported can be maintained. The curve also indicates the rapidly increasing cost of trying to generate fields approaching the critical field of Nb_3-Sn.

Coil protection

When the current in a superconducting solenoid exceeds the allowable level, the wire will locally revert to the normal state with the sudden introduction of high-resistance material (1.4 ohms per foot at 10°K) into a highly inductive circuit. Three attendant problems can result: high internal voltages, local overheating of the wire, and rapid boil-off of helium. The high voltage can be eliminated by copper-cladding the wire and thus providing a low-resistance shunt path around the normal area. Voltages can be reduced from several kilovolts to several volts in this manner. The coil is also shunted with a resistor of a few ohms, which provides a discharge path.

The problem of local overheating is more serious. As systems get larger, the stored energy increases more rapidly than the mass (although the increase is partly alleviated by the increased structure necessary) and energy can be dissipated in a local normal area faster than the normal area can propagate. The very large systems attempt to overcome this effect by forcing several areas of a coil to go normal at the first indication of a quench. Probably the best method is to distribute inductively coupled short-circuited copper foils throughout the coil. These short-circuited turns slow down the decay and overheat, thus forcing neighboring superconducting layers to go normal.

Other protection schemes to avoid local overheating involve rapidly removing the low-impedance power supply and substituting a high-resistance discharge resistor.[16] The resistor is limited only by the maximum discharge voltage that can be tolerated without internal arcing. The energy that can be dissipated in the resistor, outside the Dewar vessel, is in proportion to the ratio of external discharge resistor to internal normal resistance of the local normal area. The size of the internal resistance is in turn dependent on the rate of growth of the area[17,18] and on the speed with which the low-resistance power supply can be removed. The technique of using a discharge resistor becomes increasingly difficult with larger systems, where the stored energy per unit volume becomes greater.

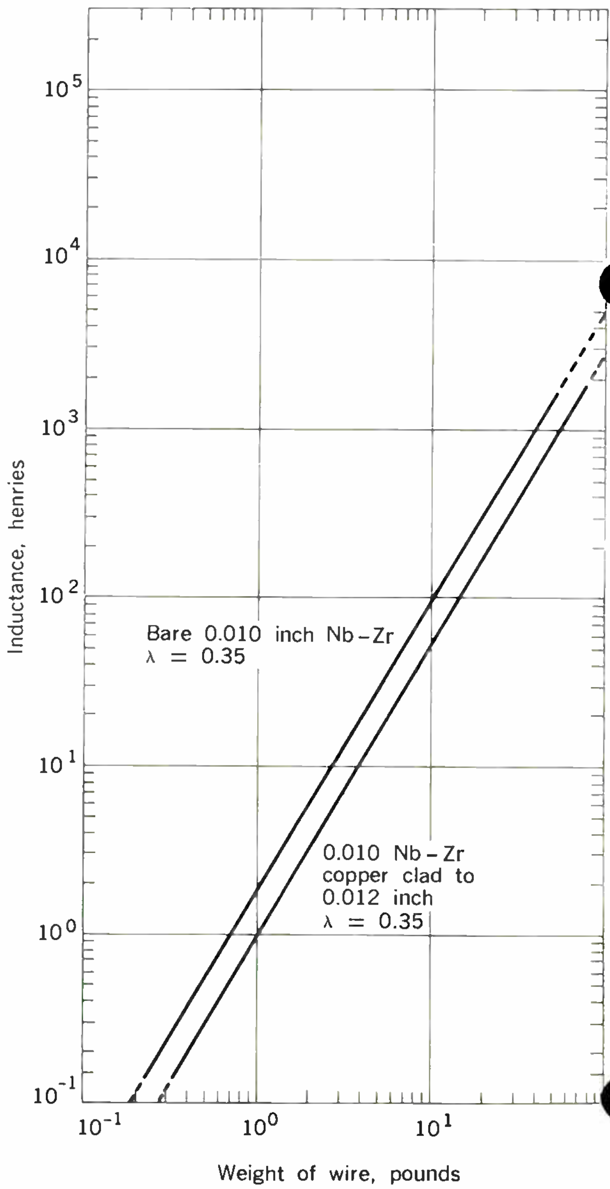
The energy stored in a coil is calculated from the inductance:

$$E = \tfrac{1}{2}LI^2 \text{ joules (henry·amp}^2\text{)} \quad (6)$$

Because of diamagnetic effects, the inductance changes by a factor of two, from one half the classical inductance at low fields to the classical value at fields over approximately 20 kilogauss. When the inductance for a wide range of coils is calculated, it is interesting to note that, within about 10 per cent accuracy, the inductance of a coil is dependent only upon the volume of material used and is independent of geometry. Thus, a large coil generating a low field and a small coil generating a high field have the same inductance if they contain the same weight of wire. The relationship between weight and inductance is shown in Fig. 6. We see that inductance increases faster than mass by the 5/3 power.

The third problem—rapid boil-off of helium during a quench—has not proved to be a major problem even at discharges approaching 100 kilojoules of stored energy.

Fig. 6. Relation between weight of wire in a solenoid and its inductance. The curve at the right represents the total weight of Nb-Zr and copper cladding in a solenoid of λ=0.35. The curve at the left represents the Nb-Zr core weight alone in the same construction.



The chief reason is that the heat released during a quench must reach the surface of the coil before appreciable boil-off begins. Due to the poor thermal conductivity of the coil structure, this process is somewhat slow. In addition, the time constants of large systems are of the order of several seconds. Therefore, although the amount of released gas can be considerable, it will not result in explosive conditions if the Dewar vessel is properly vented into a large pipe. One liter of helium is boiled off for each 2.5 kilojoules of energy dissipated in the coil. Based on Fig. 6, the result would be a boil-off of approximately 4 liters for a 10-pound coil and 200 liters for a 100-pound coil. This assumes that all the energy is dissipated inside the Dewar vessel. It is possible to couple a considerable amount of energy outside the coil, by the use of both a high-voltage discharge resistor[16] and closely coupled secondary coils.[19]

To eliminate any possibility that explosive pressure will be generated in large thin-walled cryogenic vessels,[1] several schemes have been proposed for containing the helium in tubes rather than in an open Dewar vessel. The problem with small tubes is that helium has a tendency to vapor-block and cool-down is slow. A pump for circulating helium at a high ambient pressure at which the density of the gas is high would be attractive in this regard. Such a pump, operating at 1800 psi and circulating a few cubic centimeters of liquid per second, has been developed at the National Magnet Laboratory by H. Kolm and the author, but it has not yet been used for cooling a magnet.

Mechanical stresses

Stresses in the magnet windings arise from magnetic interactions and thermal contractions. Stresses in most Nb-Zr coils are not a serious problem, because Nb-Zr is strong and ductile and the magnetic pressures are not high in the field range attainable with the material. Thermal contraction problems can be minimized by use of metallic bobbins that match the expansion coefficient of the wire. Mylar sheet is usually wound between layers and around the bobbin.

Stresses in Nb_3-Sn magnet windings are a problem and will become increasingly troublesome as higher fields are generated. The National Research Corporation[20] reports a 30 per cent reduction in current-carrying capacity of their ribbon at a tension load of 24 000 psi, and a breaking tension of 58 000 psi. They report that the winding tension must be carefully controlled so that the wire will not be stressed during the cool-down. Nb_3-Sn coils that have been potted show a certain loss of current-carrying capacity probably because of the differential expansion stress.

The approximate magnetic stresses can be calculated by considering the equivalent magnetic pressure and assuming it acts at the inner bore of the coil. The shear stress can then be calculated by assuming the coil to be a thick-walled pressure vessel. The magnetic pressure in an infinitely long solenoid is written

$$P = \frac{B^2}{2\mu} \tag{7}$$

$$P(\text{psi}) = B^2 \quad 57.5 \times 10^{-8} \quad \text{gauss}^2$$

where μ is the magnetic permeability.

The magnetic pressure in a finite solenoid is always higher since more ampere-turns per centimeter must always be used than for the infinitely long case. This ratio K lies between 1 and 2 and is written:

$$K = \frac{4\pi(\alpha - 1)}{10\, F(\alpha, \beta)} \tag{8}$$

where $F(\alpha, \beta)$ is given in Fig. 3.

The shear stress, which is the vector sum of the radial and tangential stress will be a maximum at the inside diameter and is given by

$$\delta = KP\left(\frac{\alpha^2}{\alpha^2 - 1}\right) \tag{9}$$

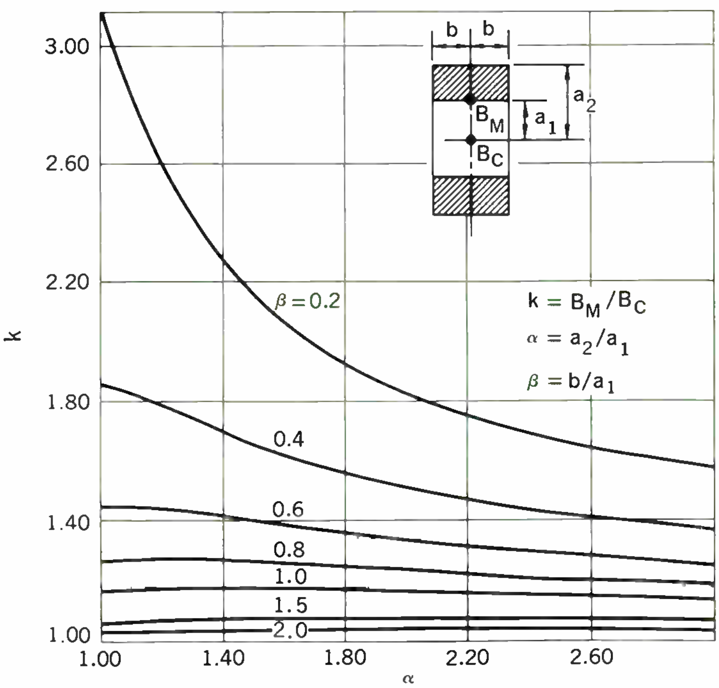
The stress is independent of the size of the coil for geometrically similar coils. However, the larger the coil bore for a given field, the lower the necessary α, and the higher the resultant stress. Take for example a family of coils at 50 kilogauss: for a coil of $a_1 = 1$ cm, $\delta = 1.04\, KP$; for $a_1 = 5$ cm, $\delta = 1.3\, KP$; for $a_1 = 10$ cm, $\delta = 1.65\, KP$. With large coils it will probably be necessary to band the coil with a retaining structure.

Stresses can also be higher than indicated by Eq. (9) in coils that are short compared to their outer diameters. Figure 7 shows the ratio k of the maximum field at the winding to that at the coil center for short coils. The stress must be scaled up by the ratio k if the pressure P is calculated from the central field.

The axial pressures operating on a coil arise from the radial component of the field. The axial pressure ranges from 20 to 40 per cent of the bursting pressure.

When there is only a single magnet system and no nearby fields or magnetic materials, the magnetic forces are self-contained. However, when neighboring coils are used, as with Helmholtz mirror coils, the forces acting between coils can be very large. These forces can be calculated exactly by considering the mutual inductance

Fig. 7. Ratio of field at the winding at the center of coils of low β. (These curves were prepared by C. E. Parker of the Oak Ridge National Laboratory.)



between coil systems, or approximately by considering the average of the field over the one arising from the other coil. Wherever possible, coil systems in separate Dewar vessels should be braced between each other by connecting tubular supports at helium temperature.

The problem of supporting large coils against the accidental presence of a large piece of external iron must be considered. The forces between a one-meter coil and a plane of ferromagnetic material one diameter away can amount to many tons of force. Suspension systems can expediently be designed with stops that do not make contact during normal operation but prevent damaging deflections.

Dewar vessels

Most small coils are operated in glass Dewar vessels. Magnets with as much as 3 kilojoules of energy storage have been operated in glass Dewar units and quenched without incident. Stainless steel Dewar vessels are less subject to breakage than those made of glass, but they restrict visual access and have somewhat higher losses. Large systems would, of course, utilize metal units, such as the one shown in Fig. 8.

There are four principal heat leaks in superconducting magnet Dewar vessels: radiation from the helium vessel walls, conduction through the supports, conduction through the electrical leads, and joule heating in the electrical leads. When operating at constant fields, magnets can be run in a persistent current mode through a superconducting short circuit to eliminate joule heating. Theoretically the leads could be disconnected during persistent mode operation, but removable contacts at helium temperature are troublesome and the switch would require careful design. Conduction through leads does not represent a large loss with 20-ampere current requirements unless there are many separate coils with individual leads. With higher currents, however, it represents the major heat leak. Use of a counterflow heat-exchange principle, where the boil-off gas is exhausted coaxially around the leads, is a great help in reducing joule heating and conduction losses. The lead cross section and lengths can be optimized by considering the conduction and resistivity of the lead material.[21]

Some typical figures for conduction and radiation losses can be given. If supports are long, the loss can be kept to about 1 milliwatt per pound of suspended weight. If the supports are grounded to nitrogen temperature this can be further reduced. A typical value for radiation loss from helium to nitrogen shields is 1 microwatt per cm^2 of helium surface. It is useful to remember that 1 watt of loss requires about 1.5 liters per hour of helium consumption.

As discussed in the section on coil protection, Dewar vessels can be constructed to immerse the coils, or to operate the coils in a vacuum space with nitrogen-cooled walls and contain the helium in tubes. Both methods have advantages and disadvantages, and further work must be done before the superiority of either can be established.

Power supplies

Rather simple power supplies can be used for energizing superconducting coils. The ac ripple content must be low and they should be suitably controlled so that dI/dt can be slow. Power supplied by means of remote programming and a crossover network between voltage con-

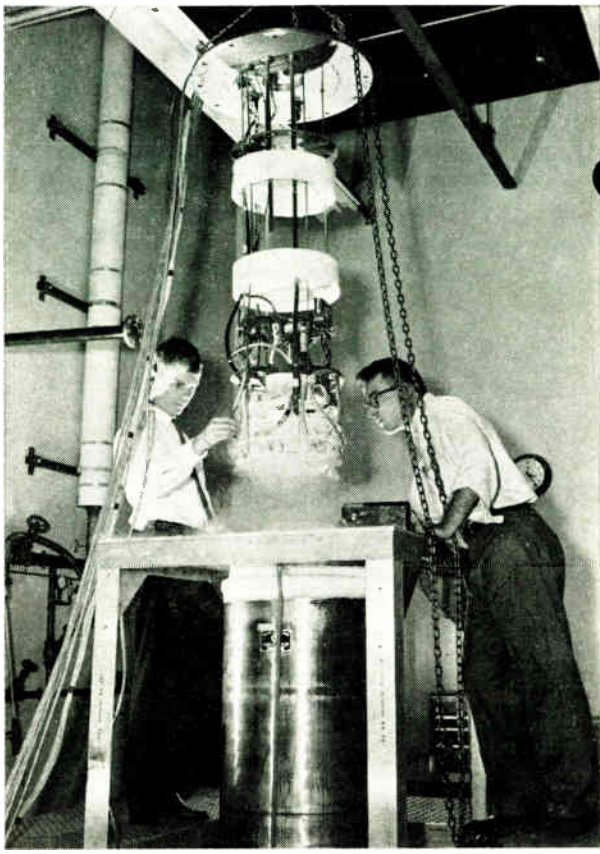
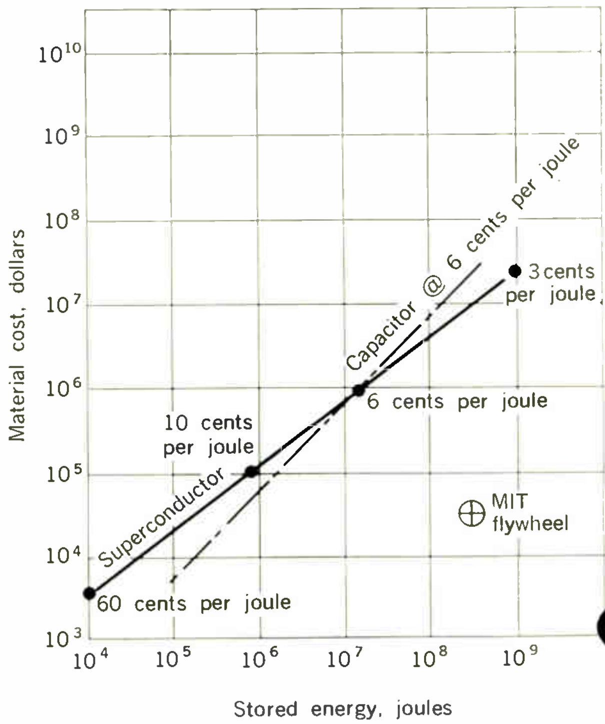


Fig. 8. A 5-inch-ID superconducting solenoid being removed from test Dewar. The coil has generated fields in excess of 30 kilogauss. (Photograph courtesy of Thomas R. Brogan of Avco-Everett Research Laboratory, whose article on MHD power generation appears on pp. 58 65 of this issue.)

Fig. 9. Stored energy in superconducting coils as a function of wire cost. Assumptions: I = 20 amperes. $\lambda = 0.3$, cost of niobium-zirconium wire = $400 per pound.



trol and current control has been found to be highly useful. For example, the maximum current can be preset and a small voltage applied to the remote error-sensing terminals. This technique will cause a small dI/dt to be applied the coil by the power supply. The inductance is generally so large that a drive voltage of $^1/_2$ volt will supply a dI/dt of only a few amperes per minute. When the set point is reached, the power supply switches to current control and holds the current constant. The coil can then be put on persistent current. The switching to persistent current is effected by introducing a superconducting short circuit across the coil and then reducing the current in the power supply to zero. The current will transfer to the short circuit and the flux linkages will remain constant.

The superconducting short circuit is generally made by wrapping several feet of superconducting wire with some heater wire and potting it. When a small amount of heater power (100 mW, typically) is applied, the short circuit develops its normal resistance of 1.4 ohms per foot and the current flows through the coil. Care must be taken to use sufficient wire in the short circuit so that its normal impedance is higher than the impedance developed by the coil inductance and required rate of change of current. If it is desirable to remove the persistent current switch without disturbing the superconducting junctions, a low-resistance removable block of copper can make up part of the short circuit. The inductance of a coil is generally so large that the introduction of a small amount of normal material in the loop will not cause an appreciable decay in a practical time period.

The allowable rate of change of field in superconducting coils is rather low, and the larger the system the lower it must be. A typical 5-pound magnet assembly can be energized at about 0.5 ampere per second up to about 80 per cent of the critical current, and beyond that should be limited to about 0.1 ampere per second.

Superconducting coils will sometimes exhibit a low critical current the first time they are energized. After each successive quench, the current will increase or "train" toward a final value, usually 10 to 15 per cent higher than the first quench. There is much conflicting data about this phenomenon of training. Many coils do not exhibit training, others only train the first time they are operated and not after subsequent cool-downs, and some coils show severe training over a two-to-one ratio. It appears that the larger a coil, the less likely it is to train, but the lower its critical current. Training has been attributed to various mechanisms—from motion of the wires to interaction of the diamagnetic currents; some combination of effects is probable. It is informative to return to Fig. 2 and postulate the training effect in terms of diamagnetic currents. As discussed in the section on wire characteristics, we assume that the degradation effect is caused by flux jumps and that the largest flux jumps take place where the circulating currents are highest—that is, at point b in Fig. 2. If a quench occurs when some part of the wire is at point b, the field will suddenly decrease, and any part of the wire that was not normal during the change will now be at point e. As the field is again raised, the currents will increase along a line ec (as shown in the other quadrant by the line ending at j). Point c represents considerably reduced circulating currents and, consequently, less probable flux jumps. This could be a contributing factor in the training that was observed. It would also explain the often-noted early quench that accompanies the reversing of the field, since in this case the currents are at their highest and pass through the point f.

Other practical applications

Energy storage. By virtue of their ability to produce a large field–volume product without loss, superconductors might be considered an attractive energy-storage medium. A large coil could be built in a toroidal form and could store considerable energy without producing external fringing fields.[22] To compete with other forms of energy storage, however, the superconducting storage must be either lighter, smaller, cheaper, more efficient, or faster operating.

If energy must be removed rapidly, it must be removed from the system in the same fashion as discussed in the section on coil protection. This requires either closely coupled copper discharge windings or a high-voltage high-resistance discharge resistor. Neither method is entirely satisfactory from an energy storage point of view, since loads are often not high-resistance loads, and copper windings are difficult to couple closer than perhaps 90 per cent and will in themselves be resistive during discharge. If energy is to be stored and removed only slowly, there is little to recommend superconductors over a battery bank, except perhaps for the ambient environment in space applications. Batteries are efficient and store considerably more joules per pound than do superconductors.

It is interesting to compare the cost of energy storage by superconductor versus a capacitor bank. Using Fig. 6, a curve showing approximate cost of superconducting energy storage can be derived; see Fig. 9. Using a cost of 6 cents per joule for capacitor storage,[23] we note that superconductors are more expensive in banks less than 10 megajoules and range from 60 cents per joule at 10^4 joules to 3 cents per joule at 10^9 joules.

Alternating-current devices. The ac performance of hard superconductors has been disappointing. Jones[24] indicates the following approximate relation for a coil made from 0.010-inch Nb-Zr:

$$\frac{P}{W} = 1.0 \times 10^{-2} \, (H^2\sqrt{f})^3 \text{ watts per gram} \qquad (10)$$

where

H = magnetic field, gauss

f = frequency, c/s

This gives a loss figure of about 1 mW per gram at 60 c/s and 500 gauss. If a transformer system could be built with such intimate coupling that the uncanceled fields were only 500 gauss, helium would still need to be supplied at a rate of 1.5 liters per hour for each kilogram of winding weight. Standard refrigerators require about 5 kW power input for each 1.5 liters per hour produced. This is many times less efficient than a standard copper transformer.

In ac or dc machinery, the outlook for superconductors is even blacker, since much higher fields must be used. However, superconducting windings can be used if they are not exposed to an alternating field. One device, for example, produces a dc field from a superconducting blocked rotor and rotates a stator with nonsuperconducting cryogenic windings.[25] The higher-gap fields available from superconductors allow such a machine to be built in a smaller package (exclusive of the refrigerator).

The transmission of alternating current will also be accompanied by losses. For example, a 0.010-inch wire carrying 200 amp rms would have a self-field of 250 gauss and would need to dissipate about 15 μW per gram. Discussions of further losses are given by Atherton[26] and McFee.[27] The attractiveness of ac transmission by superconductors is further reduced by the bleak picture for superconducting transformers.

It is possible to carry rather large direct currents without loss by means of superconductors, but the input–output devices from a cryogenic to a normal atmosphere will have losses. Although dc transformers[28] can be built if the entire secondary winding is also superconducting, these applications are rare.

Flux trapping, flux pumping, and flux concentration. Relatively simple magnets could be built if a superconducting cylinder were immersed in an external field and the currents transferred to the cylinder by reducing the external field to zero. The field can even be increased by proper shaping of the bulk cylinder.[29] In the experiments to date, flux jumps have limited trapped fields to rather low values. Most cylinders of Nb_3-Sn or Nb-Zr are limited to trapped fields of less than 20 kilogauss, although proper annealing and cooling can increase this limit to nearly 40 kilogauss. The larger the cross section of the cylinder, the more pronounced the flux jumping; hence, it is hoped that higher fields can be trapped by properly subdividing cylinders into concentric tubes with copper between sections.

Use of bulk materials need not be limited to induction from an external field. Several schemes have been derived to pump flux into a cylinder from a much smaller external field by either a mechanical pump cycle or a thermal switching cycle using two holes connected by a heat switch.[30,31] Either of these schemes is limited by the field that can be trapped in the bulk conductor.

Flux can also be pumped up by means of a rotating superconducting disk.[32] This device, referred to as a superconducting dynamo, is also useful for generating high currents for feeding a magnet with wires having large cross sections, where the currents are high enough to cause problems with the lead wires.

Superconducting shields. Superconductors can be used to block fringing fields and thus homogenize a field.[33] They can also be used in the form of a closed cylinder or short-circuited coil to cancel a time variation in the magnetic field produced by a conventional magnet.[34,35] The current-carrying capacity need be only large enough to cancel the variation in field, and in most conventionally stabilized magnet systems very little additional material is required to achieve absolute time stability. The superconducting circuit would be opened by means of a thermal switch while the field is adjusted, and then cooled to achieve a stable field.

Homogeneous magnetic fields. The problem of producing homogeneous fields is not limited to superconducting coils but, since superconductors operated in the persistent mode offer absolute time stability, it is natural to use superconductors to achieve highly homogeneous fields. A convenient and efficient way to homogenize a field is to remove some windings in an appropriate place. If the field from the uncompensated coil is written as an expansion it is relatively easy to find a "subtractive" coil that will have expansion terms of the proper magnitude to cancel one or more of the expansion terms of the main coil. This technique was suggested by Garrett[36] and has been widely

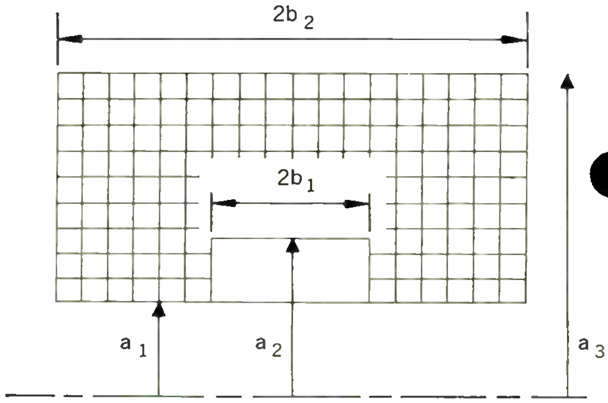


Fig. 10. Parameters to be used in conjunction with Table I.

$$\alpha = a_3/a_1 \qquad \alpha_c = a_2/a_1$$
$$\beta = b_2/a_1 \qquad \beta_c = b_1/a_1$$

used. A discussion of the method and several tables are given by Montgomery.[37] Windings can be subtracted at the outside or inside diameter, or any point in between. The nearer the subtracted windings are to the inside diameter, however, the more efficient the coil (that is, the higher the field for the weight of wire used). An infinite family of these coils is possible. Table I gives a few representative inside diameter notched coils, known as sixth-order solenoids, where the first two terms of the expansion have been canceled. The parameters are defined in Fig. 10. The axial field component can be written as in (11) if $x/a_1 \leqslant 0.5$:

$$H_x(0, x) = ja_1\lambda F(\alpha, \beta)\left[1 + E_6\left(\frac{x}{a_1}\right)^6 + \dots\right] \qquad (11)$$

I. Parameters of sixth-order solenoids with uniform current densities

α	β	α_c	β_c	E_6	F	W
2.00	2.00	1.392 0	1.263 9	−0.006 13	0.647 6	30.25
2.00	3.00	1.095 1	1.041 4	−0.003 15	1.038 0	55.24
2.00	4.00	1.034 1	0.967 4	−0.001 43	1.145 4	74.98
2.00	5.00	1.015 0	0.932 3	−0.000 70	1.189 5	94.07
2.00	6.00	1.007 6	0.912 8	−0.000 37	1.211 6	113.01
3.00	2.00	1.871 3	1.332 7	−0.004 34	1.044 8	79.58
3.00	3.00	1.233 2	1.054 7	−0.003 41	1.886 9	147.35
3.00	4.00	1.093 8	0.970 8	−0.001 89	2.160 1	199.86
3.00	5.00	1.044 7	0.933 2	−0.001 03	2.287 8	250.79
3.00	6.00	1.023 8	0.913 0	−0.000 58	2.357 4	301.32
4.00	2.00	2.385 7	1.417 7	−0.002 82	1.287 6	146.71
4.00	3.00	1.376 1	1.081 3	−0.003 09	2.588 6	276.67
4.00	4.00	1.164 2	0.983 0	−0.002 04	3.047 6	374.80
4.00	5.00	1.084 0	0.939 4	−0.001 24	3.284 4	470.21
4.00	6.00	1.047 1	0.916 4	−0.000 76	3.423 1	564.93
5.00	2.00	2.909 4	1.495 8	−0.001 89	1.441 7	231.44
5.00	3.00	1.504 8	1.108 9	−0.002 67	3.183 7	443.58
5.00	4.00	1.233 7	0.998 6	−0.002 00	3.826 0	599.9[illegible]
5.00	5.00	1.126 5	0.948 8	−0.001 34	4.183 6	752.38
5.00	6.00	1.074 4	0.922 2	−0.000 88	4.405 8	903.88

where E_6 is the sixth-order error coefficient.

As x/a_1 approaches unity, more expansion terms must be calculated. Table I also gives $F(\alpha, \beta)$ for the compensated coils, and a volume term W, which allows comparison of wire volumes necessary:

$$V = a_1^3 \lambda W \text{ cm}^3$$

Knowledge of the exact homogeneity in superconducting coils is complicated by the residual fields produced by the persistent diamagnetic currents in the superconducting wires. These extraneous fields can amount to several hundred gauss and can be troublesome unless care is taken in magnetic cycling. If the coil has not been quenched, the extraneous fields are in such a direction as to make the field more uniform. The normal residual field profile has zero gradient at the magnet center and a magnitude of less than a hundred gauss. The field peaks up symmetrically to perhaps 500 gauss near the ends of the coil.[38-41] After a quench, the end peaks are usually asymmetric, with one positive and one negative. This gives rise to gradients as high as 200 gauss per centimeter at the center of the coil. Fortunately, an excursion to nearly the critical current and back will reproduce the symmetrical peaks. The magnitude of the residual fields is easily reproduced if the same magnetic cycle is followed each time. The residual fields can be eliminated by the standard demagnetizing method of cycling the magnet between decreasing values of positive and negative current.

REFERENCES

1. Donadieu, L. J., and D. J. Rose, "Conception and Design of Large-Volume Superconducting Solenoids," in *High Magnetic Fields*, H. Kolm, B. Lax, F. Bitter, and R. Mills, eds. Cambridge, Mass.: MIT Press, and New York: John Wiley & Sons, Inc., 1962, pp. 358–369.

2. Martin, D. L., M. G. Benz, C. A. Bruch, and C. H. Rosner, "A 101,000-Gauss Niobium-Tin Superconducting Solenoid," *Cryogenics*, vol. 3, Sept. 1963, p. 161.

3. Kunzler, J. E., "Superconductivity in High Magnetic Fields at High Current Densities," in *High Magnetic Fields, op. cit.*, pp. 574–579.

4. Salter, L. C., Jr., *et al.*, "A Niobium-Tin Superconducting Magnet," in *High Magnetic Fields, op. cit.*, pp. 344–347.

5. Montgomery, D. B., "Behavior of Superconducting Wire When Wound into Coils," presented at the High Magnetic Fields Conference, Oxford, England, July 1963.

6. Riemersma, H., J. K. Hulm, and B. S. Chandrasekhar, "Flux Jumping and Degradation in Superconducting Solenoids," presented at the High Magnetic Fields Conference, Oxford, England, July 1963.

7. Lubell, M. S., B. S. Chandrasekhar, and G. T. Mallick, "Degradation and Flux Jumping in Solenoids of Heat-Treated Nb–25% Zr Wire," *Appl. Phys. Letters*, vol. 3, no. 5, Sept. 1, 1963, pp. 79–80.

8. LeBlanc, M. A. R., "Behavior of the Magnetization and Transport Current in Nb-Zr Wire," Preprint, Aerospace Corp., Los Angeles, Calif., Aug. 1963.

9. Kim, Y. B., C. F. Hempstead, and C. F. Strnad, "Magnetization and Critical Supercurrents," *Phys. Rev.*, vol. 129, Jan. 15, 1963, pp. 528–535.

10. Boom, R., Private communication, Atomics International, Canoga Park, Calif.

11. Ralls, K. M., A. L. Donlevy, R. M. Rose, and J. Wulff, "Joining Current Contacts to Superconductivity Wire," Preprint, Metallurgy Dept., Massachusetts Institute of Technology, Cambridge, Mass.

12. Betterton, J. O., *et al.*, "Size Effect and Interstitial Impurities in Nb-Zr Superconductors," in *Superconductors*, M. Tanenbaum and W. V. Wright, eds. New York: Interscience Publishers, Div. John Wiley & Sons, Inc., 1962.

13. Saur, E. J., Private communication, Institute of Applied Physics, University of Giessen, Germany.

14. Gauster, W. F., and C. E. Parker, "Some Concepts for the Design of Superconducting Solenoids," in *High Magnetic Fields, op. cit.*, pp. 3–13.

15. Riemersma, H., *et al.*, "A Variable-Composition, High-Field Superconducting Solenoid," *J. Appl. Phys.*, vol. 33, Dec. 1962, pp. 3499–3504.

16. Dowley, M. W., "Field Energy Removal from a Superconducting Solenoid," presented at the High Magnetic Fields Conference, Oxford, England, July 1963.

17. Stekly, J., "Theoretical and Experimental Study of an Unprotected Superconducting Coil Going Normal," Research Rept. 135, Avco-Everett, Everett, Mass., May 1962.

18. Smith, P. F., "Protection of Superconducting Coils," *Rev. Sci. Instr.*, vol. 34, Apr. 1963, pp. 368–373.

19. Minnich, S. H., "Protection of Superconducting Coils by Means of Damper Windings," General Electric Lab. Rept. No. 62GL165, Schenectady, N.Y., Nov. 1962.

20. *Niostan Ribbon Bulletin*, National Research Corp., Cambridge, Mass.

21. McFee, R., "Optimum Input Leads for Cryogenic Apparatus," *Rev. Sci. Instr.*, vol. 30, Feb. 1959, pp. 98–102.

22. Wiederhold, P. R., and D. L. Ameen, "Energy Storage in High Magnetic Fields Using Superconductors," presented at the High Magnetic Fields Conference, Oxford, England, July 1963.

23. Carruthers, R., "The Storage and Transfer of Energy," in *High Magnetic Fields, op. cit.*, pp. 307–318.

24. Jones, C. H., and H. L. Schenk, "A-C Losses in Hard Superconductors," presented at the Cryogenics Conference, Boulder, Colo., 1962.

25. Aigrain, P., Private communication, Ecole Normale Superieure, Paris, France.

26. Atherton, D., "Power Transmission Systems Using Superconductors," presented at the Canadian Electronics Conference, Toronto, Ont., Sept. 30–Oct. 2, 1963.

27. McFee, R., "Applications of Superconductivity to the Generation and Distribution of Electric Power," *Elec. Eng.*, vol. 81, Feb. 1962, pp. 122–129.

28. McFerran, J. B., "The Direct Current Transformer," in *Advances in Cryogenic Engineering*, vol. 6, K. D. Timmerhaus, ed. New York: Plenum Press, Inc., 1962, pp. 136–142.

29. Hempstead, C. F., Y. B. Kim, and A. R. Strand, "Inductive Behavior of Superconducting Magnets," Preprint, Bell Telephone Laboratories, Murray Hill, N.J.

30. Hildebrandt, A. F., *et al.*, "Some Experimental Consequences of Flux Conservation Within Multiply Connected Superconductors," *J. Appl. Phys.*, vol. 33, July 1962, pp. 2375–2378.

31. Swartz, P. S., and C. H. Rosner, "Characteristics and a New Application of High-Field Superconductors," *Ibid.*, pp. 2292–2300.

32. Volger, J., and J. Van Suchtelen, "Induction of Heavy Persistent Currents," presented at the High Magnetic Fields Conference, Oxford, England, July 1962.

33. Cioffi, P. P., "Approach to the Ideal Magnetic Circuit Concept through Superconductivity," *J. Appl. Phys.*, vol. 33, Mar. 1962, pp. 875–879.

34. Dunlap, R. D., C. F. Hempstead, and Y. B. Kim, "Magnetic Field Stabilization with a Superconducting Tube," Preprint, Bell Telephone Laboratories, Murray Hill, N.J.

35. Foner, S., "Stabilization and Control of Magnetic Fields with Superconducting Solenoids," *Rev. Sci. Instr.*, vol. 34, Mar. 1963, pp. 293–295.

36. Garrett, M. W., "Axially Symmetric Systems for Generating and Measuring Magnetic Fields, Part I," *J. Appl. Phys.*, vol. 22, Sept. 1951, pp. 1091–1107.

37. Montgomery, D. B., and J. Terrel, "Some Useful Information for the Design of Air-Core Solenoids," Rept. AFOSR 1525, National Magnetic Laboratory, MIT, Cambridge, Mass.

38. Anderson, A. C., and R. J. Sarwinski, "Remanent Magnetic Fields in Nb-Zr Superconducting Solenoids," *Rev. Sci. Instr.*, vol. 34, Mar. 1963, pp. 298–299.

39. Hitchcock, H. C., and P. R. Aron, "Magnetic Fields of a Nb-Zr Superconducting Solenoid," *Bull. Am. Phys. Soc.*, vol. 7, 1962, pp. 408–409.

40. Montgomery, D. B., "Current-Carrying Capacity of Superconducting Nb-Zr Solenoids," *Appl. Phys. Letters*, vol. 1, Oct. 1962, pp. 41–43.

41. Klotz, M. W., "Residual Magnetic Fields in a Nb-Zr Superconducting Solenoid," *Unpublished B.S. thesis*, Massachusetts Institute of Technology, May 1963.

This article is an expanded version of a paper presented at the IEEE Canadian Electronics Conference, Toronto, Ont., Sept. 30–Oct. 3, 1963. The National Magnet Laboratory is supported by the U.S. Air Force Office of Scientific Research.

Authors

Thomas R. Brogan received the degree of Bachelor of Engineering Physics from Cornell University in 1955. During his undergraduate years his summers were devoted to rocket engine instability studies at Boeing Airplane Company and the development of induction heating apparatus for Westinghouse Corporation. He received the Master of Aeronautical Engineering degree from Cornell in 1956, after which he joined the staff of the Avco-Everett Research Laboratory. His first effort there involved a study of the one-dimensional flow of an ionized gas in a magnetic field by the use of a shock tube.

Subsequently he was engaged in development of the high-power air-stabilized electric-arc wind tunnel (plasma generator) for hypersonic aerodynamic simulation and electric propulsion applications. He has directed the MHD Generator Project at AERL since late 1959, when the project was inaugurated by the inception of the joint Avco–private utility program for investigation of the use of the MHD concept for central station power generation. The project has been concerned primarily with the development of a large MHD generator fired by combustion gases. The work has included the investigation of the electrical properties of seeded combustion gases, detailed experimental and analytical performance studies of a large MHD generator, development of long-duration MHD ducts and combustion chambers, field coil development, seed recovery studies, and economic analyses. A prototype self-excited MHD generator with a net design output of 20 000 kW for demand usage is currently undergoing initial tests.

L. V. Berkner (F) is president of the Graduate Research Center of the Southwest, an institution founded recently to stimulate the rapid advancement of graduate education in the southwestern section of the United States. He was formerly president of Associated Universities, Inc., which organized and directs the Brookhaven National Laboratory and the National Radio Astronomy Observatory.

He received the B.S. degree in electrical engineering in 1927 from the University of Minnesota, which honored him with the Distinguished Alumni Award in 1952, and studied physics for two years, 1933–1935, at George Washington University. He was subsequently awarded honorary doctoral degrees by Brooklyn Polytechnic Institute, Dartmouth College, and the Universities of Notre Dame, Calcutta, Edinburgh, and Uppsala (Sweden).

An aviator in the U.S. 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 Material Branch from 1943 to 1945, and served on the *U.S.S. Enterprise* in 1945. He has held the rank of Rear Admiral, USNR, since 1955. He was an early leader in the Navy's development of air-borne uhf radio, radar, electronic countermeasures, and scatter propagation, as well as in space programs.

Dr. Berkner has held responsible posts in the U.S. Government, as engineer, National Bureau of Standards (on the first Byrd Antarctic Expedition); first executive secretary in the organization of the Research and Development Board, Department of Defense (now Office of Research and Engineering, DOD); special assistant to the Secretary of State in organizing NATO mutual defense assistance; and consultant to the Secretary of State on science. His report "Science and Foreign Relations" initiated the program for overseas science attachés. He was a member of and is now a consultant to the President's Science Advisory Committee.

He is a past president of IRE, the American Geophysical Union, and International Council of Scientific Unions. He is retiring vice-president of the Special Committee on the International Geophysical Year (for which he acquired the IGY program on rockets and satellites) and retiring president of the International Scientific Radio Union. He is an officer of the National Academy of Sciences and is past chairman of its Space Science Board. Among the many societies of which he is a member or Fellow are the American Philosophical Society, the American Academy of Arts and Sciences, and the Royal Swedish Academy of Sciences. He is also a former member of the Executive Committee of the International Union of Geodesy and Geophysics.

Among his nearly 100 published scientific papers and volumes in the fields of electronics, geophysics, and space science are the original space textbook, *Rockets and Satellites* (Pergamon Press, 1958), and *Science and Space* (McGraw-Hill Book Company, 1961).

rescan characters not initially identifiable; it does not need to scan a complete line and thus slow down the reading rate.

The instructions for programming the reader are done by a system called Auto-Load, which employs a preprinted sheet containing coded instructions that the print reader recognizes and automatically inserts into the internal memory. The system does not need plugboards and connecting leads. While Auto-Load sheets for frequently used documents are preprinted, an operator can make up a sheet for nonroutine batches by manually inserting the information.

Physicians aided by use of various devices

Two diverse items from the medical front consist of news of help by a transmitter in asthma research, and accomplishment by an electronic monitor in a heart ailment crisis.

Asthma a target in research. That emotional problems and environment have an intimate relation to asthma is the basis of a development that may give clues to its control. A miniature radio transmitter will soon be used to capture the conversation of children at the Children's Asthma Research Institute and Hospital in Denver. Kenneth Purcell, director of the Behavior Science Division, hopes that, by listening in on the talk of the children while they are at leisure, he will be able to pick up an indication of what triggers their attacks.

The transmitter is about the size of a package of cigarettes and weighs six ounces. Its production sprang from technology developed for Project Apollo by North American Aviation's Space & Information Systems Division. It was accomplished during research into telemetry systems perfected to monitor man and animals in space.

Two of the small devices, each with a microphone the size of a half dollar, and two dummy systems of identical appearance and weight will be used in the study.

The transmitter will be worn by the children during their leisure time, which is spent in cottage living quarters. The instrument will be carried on a belt, with the microphone attached to the blouse or shirt. A short antenna will be fastened over the shoulder. The two dummy devices will be used to help condition the wearers to carry the equipment.

As the patients talk, an ordinary commercial FM receiver in an adjoining room will pick up the conversation and transmit it by telephone to a tape recorder in the research center office for later study by specialists.

The device has some 60 components, which are packaged in an aluminum case 2 by $3\frac{1}{2}$ by $\frac{7}{8}$ inches. It operates in the 100-Mc/s range and is powered by three batteries that supply 11.6 volts. Conversation broadcast by it can be picked up clearly up to 150 feet away. It can be operated continuously for 24 hours on one set of batteries.

The transmitter operates on a single channel, but others may be added so that chest sounds and temperature also may be monitored.

Transmitters like these, modified, could have far wider application, such as for military battlefield use, for use as a space monitor in protective suit communications, and in the field of athletics.

Electronics monitors heart beat. Though her heart had stopped just a week before, a 46-year old woman was safely discharged recently from Muhlenberg Hospital in Plainfield, N.J., because of a skilled intensive-care team and a new electronic heart monitor.

The patient, suffering from a chronic cardiac condition, had been admitted to the Intensive Care Unit a few days before. In the Unit she was under constant surveillance by the electronic monitor that radio broadcasts heart rates from patient's rooms to a central nursing station. When her heart rate suddenly slowed, the monitor alerted the nursing staff and ICU house officers, and a nurse and an intern were at her bedside seconds later when she suffered the cardiac arrest; the intern had been watching the remote oscilloscope when the alarm device sounded.

The nurse (specially trained) and house officers immediately instituted external cardiac massage and after several seconds' cessation, a heart beat was restored. The next step was to restore a regular rhythm with the aid of a pacemaker. A complete EKG tape of the episode resulted, the monitor having also automatically started an electrocardiograph. The monitor, called the Guardian 6000, qualified as a life saver in this case.

The instrument's importance is in warning of changes in a patient's condition. Since those being given intensive care are in serious condition, the saving of seconds is vital; in cardiac arrest only a few minutes may elapse before irreparable damage is done to the brain because of lack of oxygen.

The data are broadcast from miniaturized radio transmitters placed near the patient and displayed numerically on the nursing station control. When a patient's heart rate exceeds upper or lower limits prescribed by a physician, the monitor rechecks the rate and, if still outside limits, sounds the alarm. The Guardian automatically activates an electrocardiograph, giving the physician a record of the episode.

The use of the unit is said to mark the first time radio telemetry—the technique of broadcasting data—has been used in a regular hospital installation for continuous heart monitoring of patients. The instrument is capable of monitoring the heart rate of six patients in sequence. It was developed by the Telemedics Division of Vector Manufacturing Company, Southampton, Pa.

X-ray technique described for measuring of thin films

An extraordinary observation during X-ray interference studies is the basis of a new method of determining the thickness of extremely thin films. The development was reported to a meeting of Federal research agencies representatives at the Polytechnic Institute of Brooklyn Graduate Center in Farmingdale, Long Island.

The phenomenon manifests itself visually as a structure in the scattered radiation accompanying the specularly reflected X-ray beam. It is attributed to an interference between the components of scattered radiation originating at two surfaces of the thin film.

This interference structure in the scattered radiation has been analyzed and used to determine the thickness of vacuum-deposited films ranging in thickness from 250 Å to 1000 Å. The new method is viewed as a significant contribution to thin film technology—the basis of all microminiaturization—and applicable to improvement of microwave equipment, infrared sensors, and lasers.

Professor Nathan Wainfran, of the Brooklyn Polytechnic physics department, presented the unusual report.

System developed that synchronizes telecasts

A system called Audlok, designed to synchronize the broadcasting of television pictures from widely separated locations is now in action.

The Audlok system consists of taking

a submultiple (below 5000 c/s) of the 31, 500-c/s synchronizing generator signal and transmitting this phase-controlled signal over an audio circuit to a remote city, and controlling the remote pictures so that they return to the point of origin in precisely the same time phase. Thus, the pictures obtained can be switched or dissolved without picture tear or roll-over. If the program is being taped, the tape machines are not subject to servo hits.

The system has been used regularly on the "Huntley-Brinkley Report" since September 9, 1963, when that program was expanded to 30 minutes. This use of Audlok is particularly significant since this program originates "live" to the eastern United States from New York and Washington.

Switching is accomplished between the two cities as if the cameras were located in one city. The Washington synchronizing generators are phased by New York so that the Washington picture (which travels via Atlanta, St. Louis and Chicago) arrives in New York in precisely the same time phase as the picture originated in New York. The circuit over which Audlok is operated—in this instance—is more than 2800 miles.

The system was made operational by the engineering department of the National Broadcasting Company.

Computers pass tests as 'diagnosticians'

Can a computer aid a doctor in diagnosing disease? The answer is yes, according to Dr. A. L. Norins, who is assistant professor of dermatology at Stanford University School of Medicine, and whose answer is based on preliminary studies with a computer system designed to accomplish this feat. The system was exhibited recently at the annual meeting of the American Academy of Dermatology in Chicago.

The Stanford system is designed to discriminate among several hundred skin diseases. For each ailment, there are 200 "attributes" stored in the computer's memory unit; they include 100 items of information about typical symptoms, physicians' observations, and patients' histories. There are also 100 items about microscopic observations of tissue samples from patients.

To make a diagnosis with the computer, a physician lists the symptoms, observations, and history of the patient involved, and also the results of the microscopic examination. The computer compares the specific case with its "experience" and chooses the most likely disease description. Samples of diseased tissue are obtained from the skin by a simple biopsy, and the skin can easily be examined by the physician.

Other groups are using computers for diagnosis of other diseases. This system is the first for skin ailments. It is the first to incorporate data from microscopic examinations, and it is being used to accumulate statistical information on dermatologic problems and in teaching students. A Control Data Corporation model 160A is the computer used in the system and in the Chicago exhibit. Studies so far have utilized a 160A located at the firm's branch in Palo Alto.

At this stage of the system's development, Dr. Norins would prefer to let a medical student evaluate a patient's condition. However, trial runs have been successful, and within a few years physicians may be depending quite heavily on computers for diagnosis. Eventually the computer center may be linked with physicians' offices by teletype, making the system easily accessible to practitioners.

A mathematical heart. Another diagnostic job is credited to a computer, this time in confirming heart defects already diagnosed by the physician. Dr. J. J. Osborn of the Presbyterian Medical Center, San Francisco, and Dr. J. G. Defares, of the University of Leyden, Holland, have cooperated with research engineers of the Berkeley Division of Beckman Instruments, at the company's analog computer facility in Richmond, Calif. They report that the analog computer has successfully confirmed diagnoses, and that it has another accomplishment: with the use of mathematics for simulation techniques, it was utilized to set up a mathematical model of the complete human cardiovascular system. The study was part of a five-year program concerned with analog computer simulation.

Dr. Osborn said the computer was used to simulate the complex mechanical functions of the heart chambers, blood volume, the elastic properties of arteries, pumping characteristics of heart chambers, and cardiac output, or blood flow. The studies are expected to lead to advanced simulation techniques in medical research and to future clinical applications.

A new Beckman analog computer was utilized that permits the results of a computation to be used simultaneously in two or more problems. This repetition enables the computer to simulate the cyclic behavior of human blood as it flows through the heart and circulatory system.

The computer model is useful in helping the investigator understand how various parts of the circulation work, both mechanically and mathematically.

The potential clinical applications of analog computer simulation include:

1. The planning of operations and the reduction of surgical time by simulating various phases of the operation, giving surgeons an advance look at the operation before it begins.
2. The study of pulse alternans (alternating pulses), a condition that often occurs with patients who are in severe heart failure.
3. Help in the determination of causes for the abnormal distribution of blood in heart disease of children.
4. Aiding in diagnosing congenital heart disease in infants and children.

A little signal goes a long way overseas

Reception of signals over long distances from "flea-power" transmitters is the subject of experimentation involving the transmission of oceanographic measurements to New York from radio-equipped buoys anchored off Bermuda.

Working with the Woods Hole, Mass., Oceanographic Institute, ITT radio engineers succeeded in receiving digital telemetry data over 600 miles from radio-equipped buoys transmitting less than one watt of power.

The strong signals were recorded at ITT's Mackay Radio and Telegraph Company's facilities at Southampton, Long Island, from a network of special ocean buoys off the shores of Bermuda. The experiments were designed to test the reliability of long-range digital telemetry transmissions employing very-low-power data stations.

More than 280 successful experiments were conducted in which the buoys radioed oceanographic data. Information received revealed the magnitude, direction, and depth of ocean currents. The state of the sea, temperature, and salinity of the water and weather conditions are typical of the data that can be collected.

The buoys, designed by the Woods Hole Oceanographic Institute under a program sponsored by the office of Naval Research, are equipped with very-low-power transmitters, batteries, and special oceanographic data-sensing gear. A command signal sent from a station at reader through the dynamism of Rome as it engulfed Etruria, Sicily, Carthage, Gaul, Britain and Asia Minor. The very long era of peace that followed these conquests contained the seeds of downfall for that mighty empire.

Even though the book covers the pre-electrical era, two items of interest are mentioned. The first is the discovery in 1936 of a small pottery jar near Baghdad that contained a thin copper cylinder and a rusted iron rod sealed by an asphalt plug. Only the lead-in wires and the electrolyte were missing to provide the evidence that some small-time electroplater was catering to the bazaars of Baghdad about 1500 years ago. But, a hundred pages hence (325), we learn that earlier than the 13th Century, chemists had had no acids stronger than vinegar with which to work. Firmer proof of electrodeposition must be awaited. The second reference concerns the discovery and the extended use of the magnetic needle. Whether it was of Chinese origin, or was used first by the Vikings, or originated in the eastern Mediterranean, is still an item of conjecture.

It is hoped that this fascinating book will quickly pay for itself so that, like the Wolf and Finch volumes, it can appear as a paperback. This work is a valuable source of information and inspiration to the practicing engineer, and it is a splendid example of a bridge between C. P. Snow's two cultures. The book indicates that there is a better chance for the scientist and technician to pierce the high plateaus of the older cultures, than to expect the historian or classicist to invade the contemporary shops and laboratories of science.

Bern Dibner, Chairman
Burndy Corporation
Norwalk, Conn.

Reviewed this month in Proceedings of the IEEE

Digital computer technology and design, vol. I—mathematical topics, principles of operation, and programming; vol. II—circuits and machine design, W. H. Ware—*John Wiley & Sons, Inc., New York, N. Y.*

Foundations of electricity and magnetism, Thomas G. Barnes—*Texas Western College Press, El Paso, Tex.*

Theory of networks and lines, James L. Potter and S. J. Fich—*Prentice-Hall, Inc., Englewood Cliffs, N. J.*

Correspondence

Engineer or scientist?—We have followed with interest and concern recent discussion in the IEEE PROCEEDINGS relating to the training, function, and image of electronics engineers.

There are two common ways of classifying occupations. One method groups people who have similar abilities and training, and who have similar goals and responsibilities. This method uses terms such as "engineer"; "scientist"; and "tradesman." The second method groups people into the field in which they work, regardless of their exact role in that field. This method uses terms such as "electronics"; "mechanical"; and "medical."

Thus each person has two affinities; one to people performing the same function; and one to people working in the same field. A motor mechanic and a plumber are both tradesmen, and have a relationship because of this. However, they also have a relationship to a mechanical engineer and civil engineer respectively because of the field in which they work.

For this discussion we have ignored one common misuse of the title "engineer," and have assumed it is applicable only to people at a professional level. As such it is applied to people in many fields; and it is obviously desirable that it be applied to the people who perform similar functions in those fields.

Historically the term has been associated with people who design and construct things rather than those who ponder the fundamental nature of things. An old definition sees an engineer "using men, money, and materials to advance the welfare of mankind."

On the other hand, people who are primarily interested in "finding out why" rather than "achieving a result" have generally been classed as scientists and philosophers.

Thus it is our contention that the people who do fundamental research are scientists, and in the case of the field of electronics, are usually physicists. Those who take the results of this research and apply it, at a professional level, to produce a desired result are engineers. The engineer therefore requires sufficient scientific training to enable him to understand the work of the scientist. Also he needs to be trained in management so that he can control and supervise those who perform the work necessary to transform laboratory results into reliable hardware. But above all else he needs the ingenuity to visualize the way in which scientific discoveries can be used to solve engineering problems.

We regret the tendency in your publication, and in others, to suggest that an engineer must be engaged in research, and that research is the only occupation requiring intelligence, initiative, and ingenuity; whereas in fact these qualities are required to as high a degree in the application of the results of research to practical problems. An analogy exists in the medical field; the majority of "doctors" do not engage in research, but they certainly keep up to date with the results of those who do, and use their abilities and training in applying these results.

From this we conclude that two professions are involved in the field of electronics. Scientists are engaged in research work which is constantly updating our understanding of fundamentals. Engineers are using the newest and the oldest scientific theories, to develop, manufacture, and utilize new devices. Both professions require men of outstanding intelligence and inventiveness, but there is a different emphasis both in outlook and training. Although there must be considerable overlap between the two roles, they should not be confused. We must get our own image of "scientist" and "engineer" clear before we can hope to project a satisfactory image to the general public.

Peter H. Griffin
Eric G. Warren
Dept. of Civil Aviation
North Sydney, Australia

Publications policy. I am very glad to see the Institute of Electrical and Electronics Engineers publishing a technical magazine for all its members. Since the merger of IRE and AIEE, the former magazines ELECTRICAL ENGINEERING of AIEE and PROCEEDINGS OF THE IRE would not have satisfied the needs of the

combined memberships of the two professional societies.

I am glad that the PROCEEDINGS OF THE IEEE is being continued for the purpose of publishing tutorial papers at the forefront of the art. With the advent of IEEE SPECTRUM you have a unique opportunity for advancing the knowledge and worth of our multifarious electrical engineers.

Technical people have recently become very much concerned over the rapid obsolescence of engineering knowledge. We have the strangely contradictory situation of the simultaneous existence of both an engineering shortage and a surplus of engineers. I have long felt that the PROCEEDINGS OF THE IRE was useful to the radio engineer who has attained the Ph.D. level, but relatively unreadable on the average to those who are on a B.S. level. This would be even more true if the PROCEEDINGS were to be sent to the power engineers making up part of our membership.

What is needed is a magazine to educate the membership in areas which are not their specialty but in which they should be somewhat knowledgeable.

I would like to see published in IEEE SPECTRUM articles written in the teaching style of the *Scientific American* magazine. By this I mean that the articles should be understandable to those outside the field of work under discussion and should not be expected to advance the knowledge of those members who are at present specializing in this work. For example, as a computer engineer, I would like to read articles written to instruct me on the construction of masers and lasers, the properties of microwave ferrites, or the design considerations for a 200 000-volt transmission line. As a computer engineer, however, I would not be fascinated by articles on binary arithmetic or elementary switching logic since I am already aware of these subjects. But I would expect that the radar and power engineers among our membership would welcome material in this category.

I hope I have made my point clear. IEEE SPECTRUM has the opportunity of raising the general knowledge and consequently the usefulness of our engineering membership by publishing instructional articles of broad interest which are designed not to go over the heads of engineers unfamiliar with the particular discipline covered.

J. D. Fogarty
Univac Div. of Sperry Rand Corp.
Blue Bell, Pa.