VOLUME 3 □ NUMBER 5 MAY 1966

IEEE spectrum

features

Departments: *Please turn to the next page*

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

departments

the cover

Considerable gains have been made recently in the high-density fabrication of cryotron circuit arrays, as described in an article by J. Paul Pritchard, Jr., beginning on page 46. A single cell module, part of a 750-cell array, is shown on this month's cover.

Spectral lines

We have many masters. I suppose that those of you who noticed the April editorial are wondering, as I am, what a new editor and a new editorial board means for SPECTRUM, a journal already run efficiently by a skilled professional staff, including several different types of editors. There are no obvious simple answers to such wondering, but some things are evident and worth rehearsing, so hear this:

As a *core* journal, which goes to all of the IEEE membership and to most of the electrical engineering profession, SPECTRUM is (or at least should be) the primary avenue of mass communication between electrical engineers. To be most effective, such communication should be bistatic, in real time, and herein lies a problem. To make it bistatic for 130 000 communicants is an almost frightening prospect; but make no mistake, SPECTRUM has missed its purpose if it does not speak *by* and *for*, as well as *to*, the individuals who make up the IEEE membership. The prospect would be more frightening if it were not for the fact that most of us are pretty quiet.

One leg of this communication path is relatively simple and well in hand. The process of getting a magazine printed and into the hands of the readers is pretty well understood, and the business of getting it done quickly is in good hands. Getting the journal to be read and seeing that it speaks effectively for those whom it serves is another story and is always subject to improvement. To this end, Headquarters has provided you with 22 official ears to bend (i.e., an editor and an editorial board), with more to be added in due course.

This is expected to be a working editorial board and its primary purpose is to be a channel by which the membership (our masters) can control the contents of the journal. It is largely up to you to put the board to work. Many of you are acquainted with at least one of the members, and for most of you there is one who is a member of your branch of the profession, so it should not be difficult to make yourself heard. Talk to the board members, or write to me; let's get this avenue of communication warmed up.

What would you like from SPECTRUM? What items attract you? Which are odious? Do you want more news? History? More entertaining editorials or more erudite articles? I have my own opinions, but not too many of yours.

I will tell you what I think. I think every article should provide intellectual exercise to some fraction of the membership and, therefore, none should be written purely for entertainment. I think there is a lower limit on technical content consistent with our professional requirement for membership. There should be some articles in each issue that provide intellectual exercise for every part of the membership and which, therefore, are quite erudite. I have heard complaints that some of the articles are so technical they discourage members from reading the journal—the one on "Relativity and Electricity" in the March issue, for instance. However, I believe that if SPECTRUM is to serve all of the membership, the technical content must have something for all classes of membership.

Furthermore, I doubt if *detailed* panel discussions, meeting reports, and interviews are worthy of the space they take. They should be more severely abstracted. What do you think? I also wonder whether editorials (like this one) get read. Do they?

Tell us what you like! Although 130 000 bosses can be wrong, we would like to make at least some of them happy—particularly since they pay the bill!

C. C. Cutler

Authors

Communications and civilization (page 43)

Donald F. Hornig. A biographical sketch of Dr. Hornig appears on page 45.

Superconducting thin-film technology and applications (page 46)

J. Paul Pritchard, Jr. (M) received the B.A. and M.S.E.E. degrees from the University of Oklahoma in 1950 and 1955, respectively, and the Ph.D. degree in electrical engineering from Iowa State University in 1960. From 1950 to 1953 he served as an air electronics officer with the U.S. Air Force. From 1955 to 1959 he was head of the Applied Math Services Group of Boeing Airplane Company, where he worked on development of computer languages for aeronautical engineering design problems. Since 1960 he has been with the Central Research Laboratories at Texas Instruments Incorporated. At present he is head of the Information Components Branch of the Advanced Components Research Laboratory. His recent technical contributions include conception and early research of a photoresist process for cryotron fabrication and an original design for a fully associative cryotronic data processor.

Dr. Pritchard is a member of Phi Kappa Phi, Tau Beta Pi, Eta Kappa Nu and Pi Mu Epsilon. He has written numerous technical articles and has lectured frequently on cryotron research and development.

Electromagnetism and quantum theory (page 55)

Dale M. Grimes (M) is a professor of electrical engineering at the University of Michigan. He received the B.S. degree in 1950 and the M.S. degree in 1951, both in physics, from Iowa State University. He then joined the research staff of the University of Michigan while continuing his graduate work. In 1956 he received the Ph.D. degree in electrical engineering. From 1956 to 1960 he served as a professor and director of the Electromagnetic Materials Laboratory. He specialized in research on insulating ferrimagnets and published 15 technical papers. He also developed and taught courses in electromagnetic field theory and in the electromagnetic properties of insulating materials. He was a member of the Technical Advisory Committee of Daystrom, Inc., and is a consultant to the National Bureau of Standards. In 1960 he helped found the Conductron Corp. and held the position of chief scientist on a part-time basis for its first three formative years.

Dr. Grimes is a member of Sigma Xi, Eta Kappa Nu, American Physical Society, and American Association for the Advancement of Science.

The origin of the term "electronics" (page 72)

Charles Süsskind (F) received the B.S. degree from the California Institute of Technology in 1948 and the M.Eng. and Ph.D. degrees from Yale University in 1949 and 1951, respectively, all in electrical engineering. He served as a research associate and lecturer at Stanford University from 1951 to 1955. He joined the faculty of the University of California in Berkeley in 1955 and is now a professor of electrical engineering and assistant dean of the College of Engineering.

Dr. Süsskind has made contributions to several branches of electronics, including microwave engineering, electron optics, and biomedical electronics. He edited *The Encyclopedia of Electronics* (Reinhold, 1962) and was coauthor (with Marvin Chodorow) of *Fundamentals of Microwave Electronics* (McGraw-Hill, 1964.) His critical study of *Popov and the Beginnings of Radiotelegraphy* (San Francisco Press, 1963, based on a *Proc. IEEE* paper) created an international controversy. One of his current projects, a history of radar, is supported by a grant from the National Science Foundation.

Communications and civilization

A civilization depends to a large extent on its communication facilities, to provide not only the materials but also the incentive for learning. Thus it is up to the engineers, who are in the forefront of the "movers and shakers" of our changing world, to help communicate new knowledge to all people everywhere

Donald F. Hornig *Special Assistant to the President for Science and Technology*

This IEEE gathering brings together a group of people—from all over the earth—who constitute one of the most important driving forces for innovation and progress. The topics being discussed lie at the heart of the modern world. The arts of the generation and control of electric power, of worldwide communications, and of the electronic transmission and manipulation of information are part of the essence of 20th century life. They not only provide us with the material tool of wealth but they increasingly become part of the quality of our century. Radio, television, the computer, cheap and abundant electric power—all these form our social patterns, our intellectual patterns, and our way of life.

Less than two centuries have passed since Volta and Faraday found the relations between chemical action and the production of electricity, and less than two centuries have passed since Joseph Henry discovered the relationship between magnetism and electricity. It is only one century since the remarkable intellectual achievement of Maxwell in formulating the dynamic theory of electricity and magnetism, and the principles governing the propagation of electromagnetic waves laid the foundation for everything that has been achieved since. The first crude telephone was installed after the Civil War, and by 1900 a primitive power-generating industry had come into existence and a few electric lights flickered here and there. But the science and the industry you represent is a truly modern one. Its great thrust came after World War I; it reached its maturity—or perhaps I should say adolescence—in World War II.

It seems safe at this point to predict that any large-scale changes during the remainder of this century will be more the product of technological advance, and particularly of the forces originating in the electronics area, than of the transitory political forces that occupy our attention from day to day. After looking over the agenda of the IEEE Convention, I am convinced that you believe this also. You should be congratulated on your foresightedness. I notice topics that run the gamut from "The City of Tomorrow" and "Automatic Highway Systems" to "After Apollo—What?" These are clearly subjects of major future importance.

It has been suggested that I discuss my thoughts about the trends of the future, and this I will do. However, I am not willing to don the robes of a prophet, as have

Full text of an address presented at the IEEE International Convention Banquet, New York, N.Y., March 23, 1966.

several of my scientific colleagues recently. For those who want prophecy, I refer you to the remarkable articles in *The New Scientist* by a series of wise and farsighted men. But I will mention a number of matters that have concerned me deeply.

I have already characterized your technology as an adolescent on the threshold of an era of extremely rapid growth. It seems to me that the heart of the revolution we are experiencing lies in the handling of information, its storage and manipulation through computers, its transmission and communication, and its use in controlling everything from manufacturing processes to space systems. All of these are involved in their most sophisticated forms when we carry out a space mission such as Gemini VIII. This activity used to be the province of the human brain alone, and is the most characteristic difference between man and the animals. Some people, at least, reason a small part of the time, and some of our reasoning is based on hard information. First through speech, then through writing and the printed book, we have been able to transmit information from one brain to another. The decision-making, reasoning powers of man have made him a sophisticated control mechanism, which, of course, is why we debate the merits of manned vs. unmanned space operations.

Reading and the printed word have been at the core of education in advanced countries. They have provided the base of education, not only in general cultural and technical matters, but also as the best means for disseminating information in our society. Civilization as we know it is tied to literacy, to the ability to read and to write.

An advancing civilization requires a base from which to proceed, a base not only of literacy but of that tradition for learning that stimulates a conscious and continuous effort by its peoples to seek new ideas. Up until the present time, those societies that have had this base have progressed by leaps and bounds, and we can expect this advance to accelerate. On the other hand, in underdeveloped societies where this base does not exist, improvements will continue to be slow, and I am afraid that the technological gap that exists between them and the developed nations will widen. This is one of the problems to which we must give our closest attention. It is there that I think your advancing technology has much to contribute, and not least by its ability to stimulate social change and its possibilities for education in a broad sense.

This age of television, of computerized information systems, and of worldwide satellite communications systems suggests the possibility of revolutionary changes in the handling, presentation, and use of information and, I believe, offers an unprecedented opportunity to enhance the progress of our own civilization and that of the developing nations.

The peoples of Africa, South America, and southeast Asia need information about so many of the very basic technologies of everyday life. I wonder whether we can assist them rapidly enough by means of the traditional methods of teaching. Do we have the time or the manpower to take the successive steps of building literacy and teaching the techniques of agriculture and commerce through these means to so many people distributed over such a wide geographical area? Must we not find ways to supplement our conventional teaching and assistance methods with techniques of information distribution and presentation that make a direct impact on the individual and his way of living while we proceed as best we can with the long-term problem of building literacy through formal education? I fully believe that the technology whereby we can do this is close at hand.

No one who has raised children in the television era can be unimpressed by the tremendous impact of this medium in increasing the awareness in children of the ways of the world. Indeed, parents are often somewhat dismayed at the extent of the impact. This result has been accomplished, not through formalized instruction, because with all the sophistication of the tool itself we have had relatively little organized use of television for instructional purposes, but just by the transmission of experiences and situations in a form that is entertaining and easy to digest. If such a significant impression can be made on youngsters by information presentations that are almost completely uncoordinated to any purpose except entertainment, what advances might we expect if we applied ourselves to devising methods of presentation that preserved the entertainment appeal but had as their purpose the imparting of educational information? And if we can sensibly ask this question in connection with our own children, I suggest that an imaginative application of this technology in the underdeveloped countries could have a tremendous impact on their rate of technological growth. A start has been made in Samoa, and increased availability of village television strikes me as the most straightforward way to end the isolation, both physical and cultural, of many peoples.

I urge your attention to this task, for there are many questions still to be resolved. Not the least of these is the problem that I have touched on—that of devising and organizing programs directly applicable to the needs of the people, in a form that is not only acceptable to the existing political structures but is also entertaining and intellectually digestible to people who are illiterate and who have not yet established a tradition of applying themselves to the hard work of learning from our customary educational presentations. I do not think we have spent nearly enough effort on this.

The distribution of programs can, of course, be accomplished by land-based broadcast facilities. But for a dispersed population in a developing country, these facilities must also be dispersed and may be expensive to build, difficult to maintain by people with few technological skills, and hard to coordinate for educational purposes. Here the satellite offers a hopeful alternative, if we can overcome political and technological difficulties. It is reasonable to assume that by the late 1960s it will be technologically possible for a solar-powered satellite to broadcast television signals that can be received by very small (perhaps six-foot-diameter) antennas with high-quality receivers on the ground. Although it would be too costly for use with home equipment, such an antenna would be appropriate for a community or school receiver through which the programs could be rebroadcast locally. We can also hope that in another decade we will be able to build a satellite that can transmit a television signal to be received over a reasonable area by high-quality home receivers with simple directive antennas. However, such a satellite would need a power supply of 20 kW or more. This is a tremendous amount of power to develop for long periods in space and will no doubt require nuclear reactor sources. Such power supplies are under development, but the technology is difficult and it is not clear when results will materialize. Even when success is achieved, the spacecraft involved may be too expensive for some applications. New techniques for information handling and processing, or other alternatives to the requirement for such high power, may be necessary to achieve a truly practical system. Nevertheless, I am convinced that it is critically important that the peoples of the world be brought into closer contact. The communications industry has been doing this, but I think bigger steps will be needed beyond those afforded by the usual commercial channels.

In another vein, the new revolution tied to electronic information handling is being led by the modern computer. I have already referred to the use of written language and the printed book as the principal base for culture and technology. Through them man has been able to accumulate his experience and to record it for the use of future generations. I don't doubt that the computer will have as great an impact on human progress. Certainly it is already exerting revolutionary changes in our methods of recording, organizing, and manipulating information—and these unique capabilities will continue to grow. But what may be more important than the computers themselves is the development of new ways to use them not only for improving government and business and advancing science but also for advancing the social sciences. We are about to do for man's brain what machine tools did for his hands.

I once checked, for what purpose I am not sure, the cost of doing 100 million multiplications or elementary logical operations. At 20 per minute and $4 per hour, this cost would be about $30 million. The same operation can be done with a modern computer for 30 cents, a gain of a hundred million. And this is just the beginning. With this development we can attack today's enormously complex problems—the problems of our transportation systems, of our cities, and of our society. The computer is not just a mechanized monitor; it is a liberating force, a force that is going to generate new insights into many of the difficulties that obstruct progress.

We are also concerned about the techniques of storing and processing information, not only to be able to present that information for educational purposes, but also as a basis for controlling our environment and our machines. In the field of control systems we must make vast strides to keep up with our other expanding technologies and needs. I can think of any number of examples of this, ranging from the requirement to control air traffic in a future airspace that includes commercial supersonic transports to the demands that will be placed upon us by our rapidly expanding power generation and distribution systems. Let us take the latter as a case in point. It seems to me a matter of simple economics that the trend of the future is toward the consolidation of power-generating facilities with larger and larger units. Most of the projected new nuclear plants are sized above 500 MW, and we are planning for several thousand megawatts in the near future. The same trend shows up in conventional plants, which means longer distribution lines and more interconnections. The Northeast power failure in November has already emphasized the need for better controls, for better communications between control points, and for automatic devices for load shedding or the control of standby power generation equipment. The trends of the future will accentuate these requirements and make it mandatory for us to find ways to control and integrate the power network in a much more effective way than we have done so far.

The technology of automatic or remote control devices has reached a very advanced state in space systems. The assessment of the limits of automatic controls has a central bearing on the planning of the future space program. In the context of our present technology there are clearly situations in which a man makes a unique contribution to a system. Generally speaking, this usually occurs when decisions or actions are needed in circumstances in which the data are too complex to be mechanized or in which the parameters of the situation could not be well enough defined beforehand. Coping with a rapidly changing environment or dealing with new and unpredicted situations are activities that are at present done better by man than by automatic systems. But the ultimate trade-offs between man and machine are not yet clear and certainly are rapidly changing.

Incidentally, in the manned vs. unmanned debates it is not always recognized that the man need not be physically present at the location where the action arises. Optical or other information can be communicated from the scene of action to men elsewhere who transmit instructions to a remote control mechanism. Man's unique capabilities can be used with requiring his actual presence. The remote manipulators that have been developed to handle radioactive materials are examples of devices that might be used in some space operations situations. A man in a Lunar Excursion Module on the moon's surface could operate a vehicle without leaving the module.

Man will probably continue to be necessary on many space missions for a long time to come. Even so, the problems he faces will be so complex that he will need every automatic assistance that we can give him to insure that his unique capabilities are not dissipated in performing unnecessary and routine functions.

This is also our problem on earth. As computers, communications, and control devices become more capable, one of our biggest tasks is to relate them to man. They can relieve tedium and boredom, eliminate both physical and mental drudgery, and open undreamed-of avenues for achievement.

In closing, I want to say again that we cannot do these things only for the rich and the powerful nations. The new possibilities will become realities with a certainty in some of our countries; but the world cannot survive with a few islands of plenty amidst oceans of misery and starvation. If we cannot use the new technologies to help the peoples of the entire world, we will have failed. These new technologies are bringing the nations of the world closer together—but this makes more of the world's problems our problems. There is no escape from this fact, but there is an escape from the facts of overpopulation, disease, and ignorance if people will devote themselves to the task of eliminating them.

Donald F. Hornig is a graduate of Harvard University, where he received the B.Sc. degree in 1940 and Ph.D. degree in chemistry three years later. After a year as a research associate at the Woods Hole Oceanographic Institution and three years as a group leader at the Los Alamos Laboratory, he joined Brown University as assistant professor. He became a professor in 1951, associate dean of the Graduate School in 1952, and later acting dean. He joined Princeton University in 1957, becoming chairman of the Chemistry Department in 1958.

Dr. Hornig was appointed Special Assistant to President Johnson for Science and Technology in 1964, when he was also named as chairman of the Federal Council for Science and Technology. Previously he had served as a member of the Advisory Committee, Office of Scientific Research, U.S.A.F., and on the Space Science Board of the National Academy of Sciences. He served on the President's Science Advisory Committee for both Presidents Eisenhower and Kennedy, and at present is chairman of that committee. He is a Fellow of the American Physical Society, the American Academy of Arts and Sciences, and the Faraday Society, London.

Superconducting thin-film

Recent advances in cryogenic array fabrication have made it feasible to consider thin-film superconductive devices as being potentially competitive in the computer field, particularly for highly parallel data processors

J. Paul Pritchard, Jr. *Texas Instruments Incorporated*

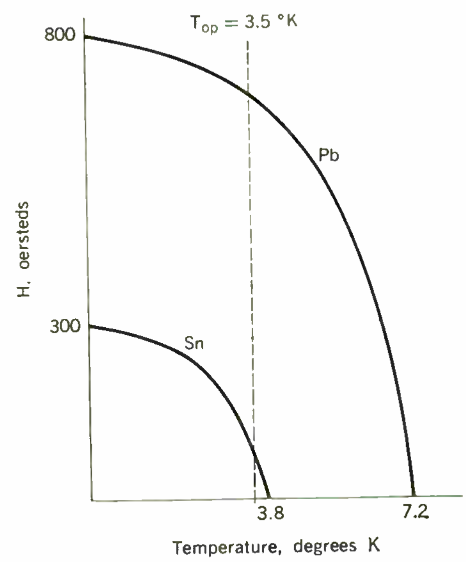


Fig. 1. Phase portraits of the superconducting boundaries of thin-film specimens of lead and tin. For cryotron application, isothermal operation is presumed, as indicated by the dashed line at 3.5°K.

In the past two years or so, significant gains have been made in cryogenic array technology and also in the necessary support areas. This survey reviews the problems peculiar to array fabrication, cites experimental evidence of recent progress, and attempts to project the likelihood of the emergence of various cryoelectronic data-processing systems. The approach is through a review of basic characteristics of superconductive devices for logic and storage functions and their implications with respect to fabrication methods.

In the decade since the superconductive switch[1] was conceived as a potential computer element, its researchers have passed through alternate periods of exhilaration and frustration. Cryoelectronic devices have been recognized as the only imminent means of economically achieving and evaluating very large capacity data-handling systems with high data-processing rates accomplished through use of highly parallel system organization and operation. Enthusiasm for these devices stems from this consideration and from the relative simplicity of fabricating and operating superconductive elements in small numbers. However, problems have resulted from the need for efficient production of device arrays at high density and count per unit of manufacture imposed by the economics of refrigerated operation near 3.5°K.

Potential notwithstanding, the time is rapidly approaching when either feasibility must be demonstrated at the prototype level or the endeavor should be abandoned in favor of more fruitful areas of research and development. In May 1964, IEEE SPECTRUM published a review of the fundamental results accrued in earlier years with respect to relevant superconductive phenomena, materials, and application-oriented device considerations.[2] The conclusions of that survey properly cited the absence of a currently suitable array fabrication means as a limiting factor to maturity of cryoelectronics. Other aspects of the proposed systems, such as array interconnections with closed-cycle refrigeration, necessarily awaited progress in device fabrication before receiving deserved research and development attention.

Since that survey, significant progress has been made in array technology[3] and, as a consequence, in the necessary support areas. The intent here is to review this progress and to predict possible areas of application for these devices, based on present problems and achievements.

Thin-film cryotron properties

For a select class of materials, there exists a superconductive phase with respect to ambient temperature and magnetic field for which the resistivity is zero. For the present discussion we are concerned with the subclass of superconductors that also display diamagnetic behavior in the superconductive phase.[4,5] These effects suggest exploitation for both small-signal and switching devices. The thin-film cryotron[6] has received perhaps the greatest attention of a variety of superconducting devices, with the continuous-sheet memory[7] a close second for mass memory applications.

The cryotron is basically a resistance switch whose conducting state may be influenced by control of ambient magnetic field levels, either self-induced or externally provided by current in an adjoining conductor. Isothermal operation is presumed. The phase portraits for

technology and applications

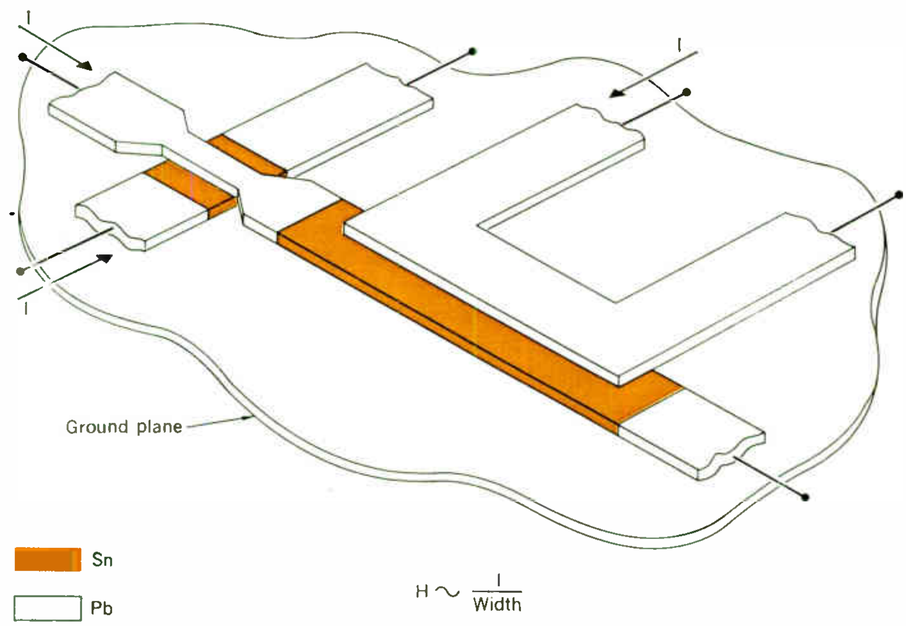
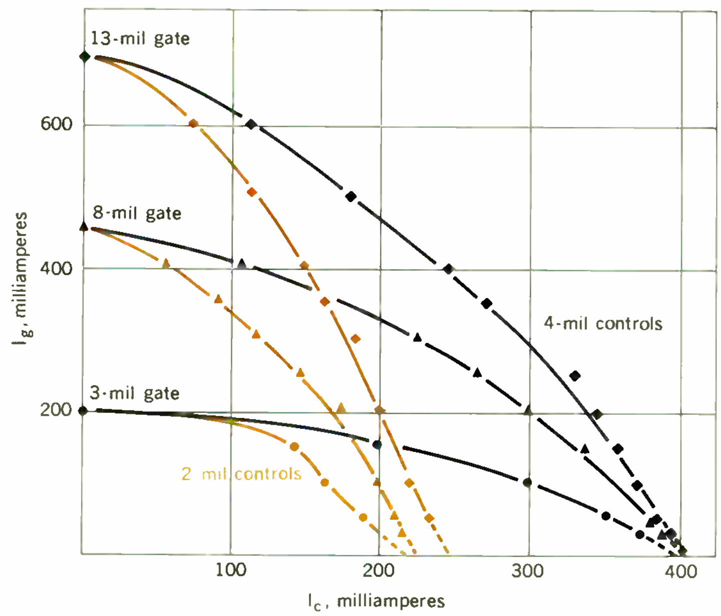


Fig. 2. Representative thin-film (≈1 μm) structures for the two basic cryotron switching mechanisms. A crossed-film cryotron is shown on the left and an in-line polarity-sensitive cryotron on the right. Insulation 1 μm thick separates the several conducting paths and the ground plane from one another.

Fig. 3. Gate current vs. control current for crossed-film cryotrons of varying crossing width ratios at 3.5°K.



the most commonly used pair of superconductors, lead and tin, are shown in Fig. 1. Assuming isothermal operation at $T = 3.5°K$, it is seen that there is an appreciable separation in critical fields for resistance restoration in these materials. Thus, one imagines the use of tin segments where a controlled resistance is desired and the use of lead elsewhere to yield nondissipative control and interconnection lines.

Figure 2 illustrates the two basic thin-film structures for the cryotron. The crossed-film cryotron on the left consists of a tin gate region sandwiched between and insulated from a diamagnetic lead shield plane beneath and a lead control line above. Lead films also provide electrical connections to the tin gate. Gate resistance can be restored by either a self-current I_g or a control current I_c, or combinations of these, as illustrated in Fig. 3 for a right-angle crossing of control and gate paths. For other than right-angle crossings, an asymmetry will exist in the gate–control current characteristic, since the two fields combine vectorially to determine the net ambient field at the gate. Thus, we have both a cryotron form that is sensitive only to control-current magnitude and a polarity-sensitive counterpart, the in-line[8,9] cryotron, as shown at the right of Fig. 2. A typical asymmetric gate–control current characteristic is shown in Fig. 4.

Magnetic fields due to current flow in the gate and control lines are essentially confined to the dielectric region in the shadow of these lines with respect to the shield plane. In this way a more uniform surface current distribution in the lines is achieved. Therefore, the critical current-carrying capability of a shielded superconductor exceeds that of its unshielded counterpart by a factor of two or three. The field confinement results also in suppression of undesired signal coupling between lines separated in the plane by distances that are at least an order of magnitude greater than the separation from the shield plane. The latter implies also a major reduction in inductance per unit length of line for a shielded circuit. Quantitatively, we are concerned with metal films having thicknesses of 0.3 to 1.0 μm and widths of 25 μm and greater. Insulation layers are necessarily of the same order of thickness in order to achieve short-circuit-free circuitry.

Crossed-film cryotrons display variable current gain by appropriate line-width control. The large width-to-thickness ratio of the shielded films implies the first-order dependence[10] of magnetic-field strength on film current

$$H \approx \frac{I}{w}$$

where w is the film width. In the absence of control current there is a critical gate current I_{cg}, for which the induced field is that necessary to restore gate resistance. Conversely, there is a critical control current I_{cc}, below which there is insufficient external field generated to restore gate resistance, in the absence of gate current. The ratio of these currents is the static gain defined as

$$G = \frac{I_{cg}}{I_{cc}}$$

It may be shown that the gain is, in principle, proportional to the ratio of gate and control widths at the crossover, called the "crossing ratio":

$$G = \frac{w_g}{w_c}$$

In practice, however, G is diminished by a small factor related to absolute geometry of the structure and the penetration depths of the materials. The gate current vs. control current characteristics of Fig. 3 illustrate these factors for cryotrons of varied line widths and crossing ratios.

For certain applications, the cryotron may be considered to be analogous to a current relay for which the conducting states are also infinitely separated. Therefore, one can visualize a current steering loop as one means of accomplishing logic, as in Fig. 5. In the absence of control currents (I_{c1} and I_{c2}), a supply current I_s divides between the two superconducting paths inversely according to their inductance. If the gate in the left branch is switched resistive by current I_{c1}, current redistribution will take place in the loop. The governing time constant is the quotient of total loop inductance by restored resistance. Representative switching time for a loop inductance of 20 squares (a few picohenries) and a restored resistance of 0.3 milliohm is approximately 25 nanoseconds.

In Fig. 5, control current I_{c2} has a complementary effect to that of I_{c1} on loop current distribution. An important aspect of the flip-flop action lies in the circuit's ability to remain in, or latch at, the current distribution that exists at the time the controlling branch reverts to the superconductive state. The lead extensions of the gates X_1 and X_2 generally become controls for cryotrons (i.e., X_3, X_4) of $G > 1$ residing in other superconductive circuitry; thus, cooperative action among circuit functions is possible.

Dissipationless storage may also be accomplished in the current steering loop. Consider I_{c1} to be holding X_1 resistive, thus diverting I_s to the right-hand path. If I_{c1} is removed, X_1 reverts to the superconductive state. If the current I_s is then also removed, a circulating supercurrent will result in the loop. Its magnitude is

$$I = \frac{I_s L_2}{L_t}$$

where L_2 is the inductance of the right-hand path and L_t that of the loop. Presence or absence of this nondissipative stored current may be detected nondestructively by observing the resistive state of the gate X_3 located in another superconducting circuit. Alternatively, the use of a bipolar current source I_s permits the stage of storage to be alternatively represented by the rotational sense of stored current flow. In this case a portion of the loop may serve as a control for an in-line gate in an associated circuit (i.e., X_4). The control and gate widths of X_4 may be so designed that resistance appears only if the field contributions aid. It is interesting to note that the latter mode of loop operation permits storage of a three-state variable, since absence of a stored current is also a sensible condition.

This applicability of a common device and circuit or-

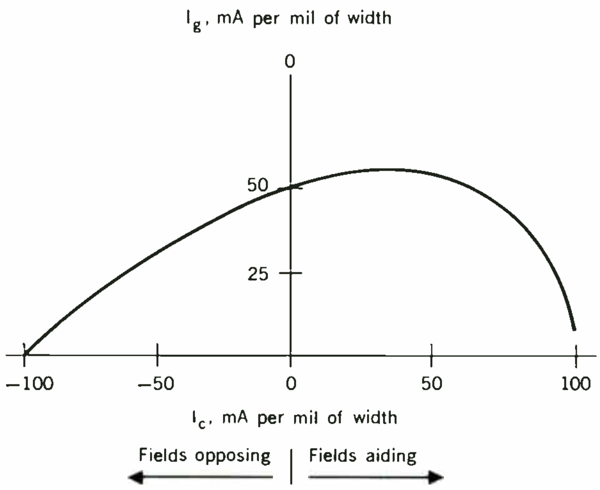
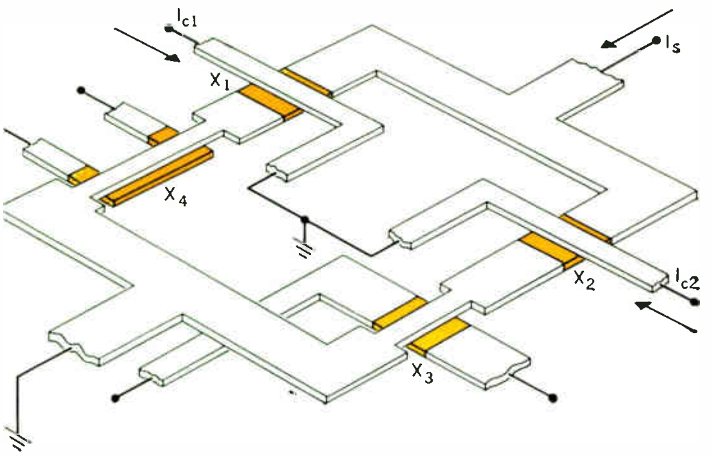


Fig. 4. Idealized gate current vs. control current for an in-line cryotron. The asymmetry evidenced with respect to gate current reflects the desired polarity sensitivity of this device geometry. A typical characteristic is shown for a device of equal gate and control widths in the range of 0.002 by 0.005 inch.

Fig. 5. A typical cryotron current steering loop illustrating the basic means of controlling and sensing current distribution. Configuration is equally useful for performance of logic and/or storage functions.



ganization for logic and storage functions provides a basis for compatible logic and memory circuitry. This and the seemingly simple fabrication requirements, with the projected applicability to device arrays of high density and count per unit, constitute the unique potential of cryoelectronic devices.

Potential applications

As opposed to conventional computers, which are organized for the sequential processing of data, there exists a class of deterministic machine organizations that permit a high degree of parallel processing. The implication of these organizations is not that of extended logical or mathematical capability, but rather of improved cost effectiveness in performing certain classes of data manipulation. The parallelism is achieved by imbedding logic within the memory in such a manner that stored data may be logically processed *in situ* and thus, potentially, in parallel by word and bit. This is the major basis for vastly increased data-processing rates.

The requirement for logic-in-memory reflects a technology requirement for integrated devices wherein arrays of functional blocks are the unit of manufacture, interconnection, maintenance, and replacement. The active semiconductor memory approaches may afford solutions at small storage capacity levels in the one- to five-year time interval. For large capacity (greater than 10^5 bits), problems of cost or power density, or both, appear to necessitate consideration of the thin-film cryotron, assuming its successful reduction to manufacturing practice in the near future.

Perhaps the most publicized parallel data-processing organization is that of the associative or content addressable memory,[11] which will be discussed in some detail in the next section. The essential feature of this system is the ability to search stored data in parallel by word and by bit, simultaneously marking those words that match the interrogation criteria for subsequent search or input/output action. The field of search within the storage word and length is arbitrarily specified in a peripheral register, as is the argument value(s) relevant to the search desired. A variety of search criteria may be accomplished through use of this method.

The applications benefiting most from these data manipulation features are variously referred to as multidimensional table look-up, set partitioning, pattern classification, and list processing, to name a few.

Less publicized forms of parallel, logic-in-memory organizations lend themselves to variable-word-length applications. In these a character organization is adopted; thus, search remains parallel by word, but drops to parallel by character.

Yet another parallel organization provides associative addressing of multiple data processors according to a processor state descriptor unique to each.[12] In this organization all processors share a common input/output data link and accept control instruction, provided in parallel by a processor, if associatively identified for such action. It is likely that major improvements in data-processing and computation rates will accrue in cases where vector manipulation is a basic mathematical tool of the problem.

Only slight reflection upon the implied count of switching functions and required degree of integration of logic and memory is necessary to recognize the requirement of batch-fabricated-array technology in these processors.

It should be apparent that a cryoelectronic random-access memory, either of discrete-cryotron or continuous-sheet form, is also a potential application. However, this is an established and highly competitive market, which is already served by ferromagnetic devices. The cryoelectronic memory should compete favorably on an economic basis where the storage capacity requirement exceeds 10^7 bits at a one-microsecond cycle time. At capacities in excess of 10^8 bits, cryoelectronic memories appear to be the only feasible medium retaining few-microsecond cycle-time capability at an acceptable total system cost. One advantage in favor of cryoelectronics is the fact that refrigeration and selection logic (integrated within the memory arrays) costs are relatively insensitive to increases in storage capacity. Further, the lossless character of superconducting selection lines permits substantial increases in memory capacity served by a single set of drive electronics without severe degradation of cycle time. Nondestructive readout is inherent in the discrete cryotron memory form, and storage is nonvolatile with respect to signal power loss.

A representative application

The associative or content addressable memory[11] is an excellent example of the functional integration possible with cryotrons.* It serves equally well to introduce a discussion of fabrication requirements. Figure 6 shows a single associative cell structure (film structure caricatured) in context with typical word and bit circuitry for the application. In actuality, the circuitry represented here by line segments is also produced in thin films. (A semicircle represents a tin gate included serially at the point shown; all other line segments are of lead and serve as controls only where they intersect a semicircle.) The basic principles of application described in connection with Fig. 5 should be noted as these appear, subtly varied, in the performance of the several functions of this cell and its related circuitry.

The direction of circulation of current in the loop powered by the "bit write" line (BW) reflects the binary storage state at the site. To produce the trapped current, BW current of polarity appropriate to the desired binary state is diverted through the left path of the storage loop by rendering gate W resistive and then discontinuing the controlling current for W and the BW current in that order. The stage of storage will be affected only in a site for which a controlling current exists for gate W and for which a BW current is present. This control function is accomplished by diverting an "enable" current (EN) through the high-inductance shunt on gate E at sites for which BW current is present, and only there. Generally only one word in such a memory possesses current in its EN branch for a given writing operation; thus, selectivity of affected sites is possible. EN current is discontinued at the appropriate time by pulsing the reset line (R-ER) associated with the "enable register" (ER) flip-flop for this word.

The existence of a mismatch between the storage state at the site and an interrogating variable (BW in Fig. 6) for this bit position in all words in memory is determined at the site and evidenced as a restored resistance in the

* A complete discussion of the associative memory function is beyond the scope of this article. The interested reader is referred to the bibliography of the referenced paper for this material.

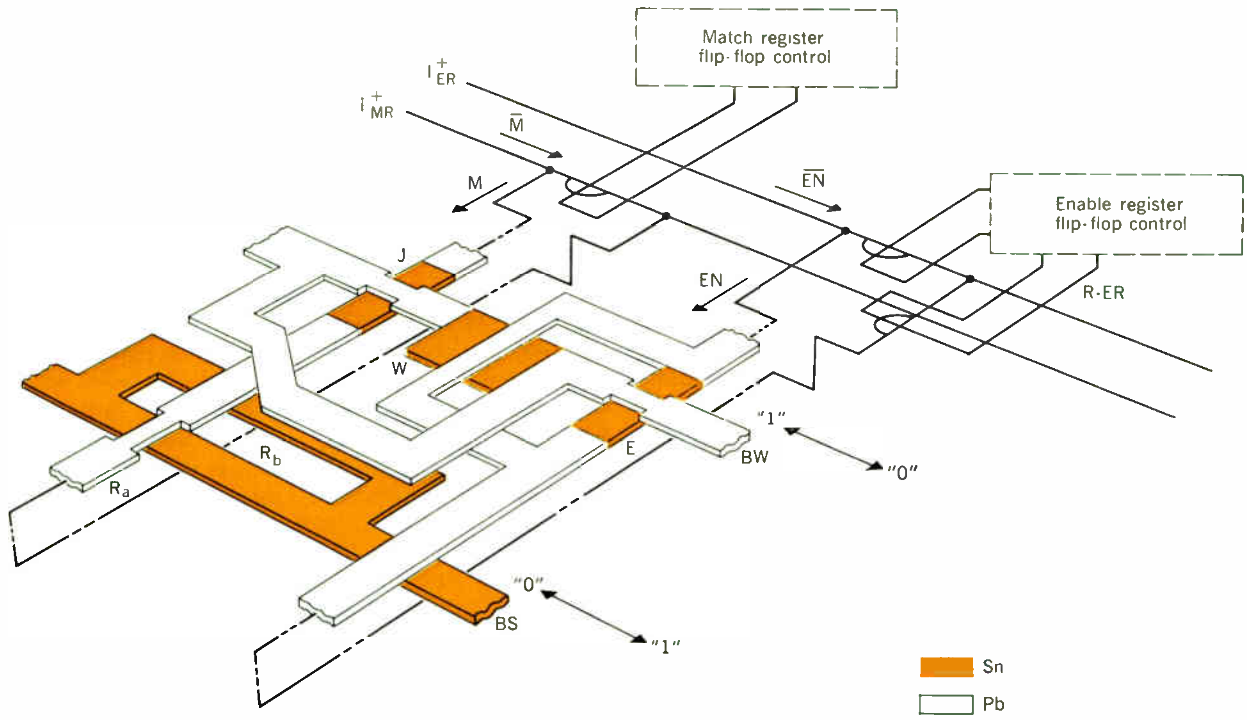


Fig. 6. Associative storage site embodying provisions for storage, read, write, and comparison functions. The film structure of the cell is caricatured; the relevant interconnection paths to word-control circuitry, simultaneously produced in the array, are shown only in line-diagram form.

gate *J*. Observe that if the interrogation and storage states are identical, the net current through the control of *J* is a minimum; conversely, if the states differ it is a maximum. A suitable choice of current levels and control geometry permits restoration of resistance in gate *J* if and only if a mismatch exists at this site. Although current in the "match register" (MR) flip-flop is initially in the *M* branch for all words that are to be compared, a mismatch at any particular site will reset the MR flip-flop for its word.

Readout of the storage state for a site involves the restoration of resistance in the in-line gate, R_b, when the fields produced by current in the "bit sense" line (BS) and the storage loop at the site aid. The uniqueness of readout to the storage sites of a given word is enabled by restoring resistance in R_a by the presence of *M* current in only that word's MR flip-flop. Thus, for a given polarity of applied BS current, the presence or absence of voltage at the BS line terminals is an indication of the storage state for the associated bit position in the selected word.

In summary, the associative site of Fig. 6 incorporates two AND functions and an EXCLUSIVE OR in addition to the accomplishment of the nondestructive data storage functions. This is obviously of academic interest only, unless the devices can be produced simply and in such a way that they will perform consistently and reliably in high-density arrays.

Fabrication requirements

An examination of the circuit of Fig. 6 illustrates the required characteristics of an acceptable fabrication means. Clearly, provisions must exist for depositing a minimum of three material types (lead, tin, and an insulator) at thicknesses of approximately 1 μm. The total number of material layers to be used in structuring a given circuit involves consideration of several process-related factors. For this discussion we will assume a photomask–photoresist[3] fabrication means wherein photoetch methods are used to produce the desired patterns in each layer following a large-area deposition of the appropriate material. Since a monolithic lead ground plane is presumed, at least five material layers are required. These include the ground plane and its insulation, followed by tin, insulation, and lead layers, which form the gates, controls, and interconnecting paths. Passive crossings of signal paths will generally be required, and although one of the paths may be formed in tin at such an intersection, it is necessary to widen both this path and the overlaying lead path in order to ensure passivity of the crossing. This procedure, however, is detrimental to efficient circuit density in the array.

It would appear that a nine-layer structure is, in general, the maximum requirement. In this case, three lead layers, with appropriate intervening insulation, are used to form the control and interconnecting signal paths with respect to a single tin layer in which all gates have been simultaneously formed. This approach permits a nearly optimum utilization of the array surface and provides flexibility for computer-aided design methods used in circuit layout. In the circuit of Fig. 6, two lead layers are used to form the control and short-run interconnectors among active tin gates. The long return paths for the ER and MR flip-flops are rendered in a third lead layer passing beneath the ground plane. Thus, they contribute only to the thickness of the circuitry and not to its surface area extent.

The emergence of the photomask–photoresist fabrication method not only has provided a means of achieving

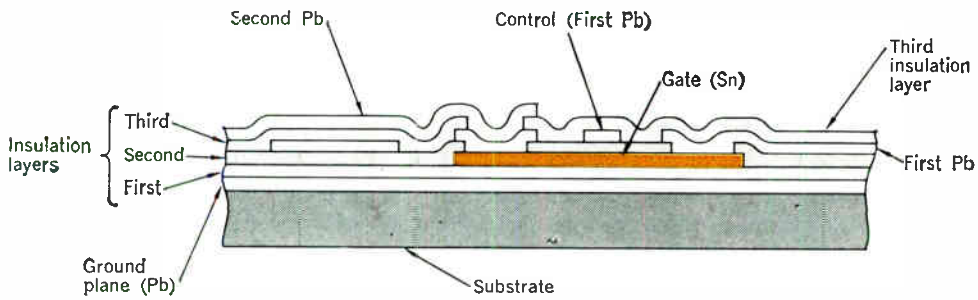


Fig. 7. Cross-sectional view of thin-film contributions to hypothetical 7-layer cryotron circuit. View is taken longitudinally through a gate whose lead connections are supplied in separately formed lead layers.

high-resolution circuitry in a practical minimum of material layers but also has permitted a vital breakthrough in low-temperature insulation practices. In order to photoetch lead selectively in the presence of prior layers of lead or tin, an etch-resistant coating must be applied to the structure as well as to the areas to be retained in the current layer. The method used is to insulate each preceding layer everywhere except at points where electrical contact is to be made from above; it is important to select an insulating material that is also unaffected by subsequent process steps. The obvious and highly effective solution is the use of a "photoresist" itself as this insulator. Photoresistive materials are photosensitive polymers that are self-patternable by virtue of intended photographic exposure and development properties. They also possess excellent surface-wetting and adhesion characteristics and do not thermally stress relieve under repeated temperature cycling. Dielectric properties of these materials[13] are more than adequate for thin-film superconductive device applications.

A representative view of the multilayer structuring is illustrated in Fig. 7. The vertical section is taken longitudinally through a controlled gate of a hypothetical seven-layer circuit. The vertical dimension has necessarily been exaggerated for clarity, since although film thicknesses are only about 1 μm, the narrowest film width, which occurs in a control segment, will typically be in excess of 25 μm. The tin gate of the illustration is serially connected to a path formed in the first deposited lead layer following insulation of the tin gates. The left-hand gate connection is assumedly not completed until a subsequent lead layer is applied. As shown, a lead pad is formed in the first lead layer to provide the necessary through-contact between the tin and second lead layers. In the example, this seemingly awkward connection practice was necessary so that a passive crossing could be provided between the left-hand gate connection and another signal path produced in the first lead layer. Such configurations are not infrequently required, however, for the purpose of meeting specific circuit functions and simultaneously achieving maximum circuit density in the array.

The interlayer contacts in the cryotron circuitry must be superconducting, inasmuch as they generally appear in a storage or switching loop. Laboratory-scale fabrication involves cycling of the circuits between the high-vacuum environment required for vapor deposition and a normal room ambient in which the photoetch steps are conducted. There is, therefore, an opportunity for the formation of oxides or other nonsuperconductive material on the surface of an intermediate metal layer, which could prevent proper interlayer contact. The solution of this problem is straightforward, but cannot be discussed in detail for proprietary reasons.

A more serious potential for device failure exists in the formation of short-circuit filaments between adjacent metal layers at other than desired points of interlayer contact. This failure mode in fabrication has been vastly diminished through incorporation of polymeric insulation. Nevertheless, because a number of cycles are required to reach room ambient conditions for photoetch processing of the several metal layers, the insulating layers are vulnerable to particle contamination. The flaws may be variously attributed to dust particles, moisture, etc., which result in voids or pinholes in the polymer layer. Photoetch steps are, therefore, conducted in a controlled environment (such as a laminar-flow, filtered air bench) to minimize exposure to dust particle incidence. Particular attention is also given to removal of adsorbed moisture from the sample surface prior to insulation. As a practical matter, short-circuiting filaments having cross-sectional areas small in comparison with that of the film structure may be cleared by conventional capacitor-discharge techniques. Neither the short-circuit nor surface-contamination possibilities make laboratory-scale fabrication impossible. Further, it is obvious that these problems would be virtually eliminated in a controlled-environment manufacturing facility in which cycling to room ambients is eliminated.

A significant factor in the cryotron technology is the absence of stringent requirements regarding control of impurity concentrations and grain size within the deposited films. It is certainly possible to degrade the superconductive properties of lead and tin films through poor practice in deposition technique; however, exercise of standard high-vacuum (10^{-6} torr) practice in vapor phase deposition of high-purity materials yields films with highly reproducible critical-temperature and critical-current capabilities. This relative insensitivity to deposition parameters can be attributed to the nonlocal interactions among paired electrons, which form the basis for the superconductive behavior exhibited. The interaction lengths among these electrons are in excess of 1000 Å for the materials of interest.

Another important factor affecting the feasibility of cryotron array fabrication concerns the means for defining film segments in each of the deposited layers so that the multilayer structure will be coherent and identically scaled at all points of the array. For the current fabrication methods, this problem is related to the means of preparing and using a well-registered set of photomasks for use in the photoetch steps. The individual photomasks are generally photographic plates developed to provide maximum contrast between the portions that transmit and block light. The artwork from which the mask pattern derives is prepared at a maximum practical magnification to minimize the effect of mechanical distortions in its preparation.

The arrays that are of concern in cryoelectronics consist of large numbers of one or more basic circuits, which are interconnected and distributed over the substrate surface in an orderly, repetitive manner. It is, therefore, desirable to produce artwork only for the basic repetitive pattern elements. The artwork is converted to a photographic reticle at a suitable reduction and successively repositioned for multiple exposure of another plate to generate an intermediate or final composite of the array. This step-and-repeat operation requires precise control of photographic and mechanical parameters, particularly in the formation of the set of photomasks for the multilayered structure.

The prior art evolved for semiconductor device fabrication placed emphasis on the ability to precisely reproduce line segments as small as 2 μm wide within a given device structure. The final array extent was restricted to one square inch or less, and the repeated structures in the array were intentionally separated to facilitate segmentation of the array into discrete devices for packaging. Thus, errors in the photographic step-and-repeat preparation of photomasks would at worst cause rejection of some, but not all, of the devices formed through misregistration of parts.

Array areas in excess of four square inches are required for cryotron circuitry from the viewpoint of economics, and rejection of failed structures because of masking errors in the array is not feasible. The repeated structures are interconnected intentionally, and the desired maximization of the device density severely limits the permissible tolerance in the mispositioning of adjacent, repetitive structures.

In short, the photomask set must yield a multilayer structure whose parts are registered within a 2-μm tolerance at all points of an array of multiple square inch area. The present art appears to permit maintenance of this tolerance for a nine-mask set involving minimum line widths of 25 μm. A variety of pattern elements may be combined to form an array of perhaps 16 square inches at this time. Several photomask fabricators are currently evaluating means of extending the existing art, with indications of success in the near future, certainly consistent with the projected time scale of cryoelectronic device development.

The degree to which high-density fabrication of cryotron circuit arrays has progressed can be illustrated in terms of an existing program, which has as its goal a cryotron associative data processor having a capacity of 250 000 bits.* The basic associative cell or bit structure is accomplished as described in the discussion of Fig. 6. A nine-layer realization of that cell is shown in Fig. 8. The cell module is 0.028 inch wide (word direction) and 0.030 inch high (bit direction). The intended use involves 50-bit word lengths, with appropriate cryotron control circuitry appended to each word. Thus, a full word with control circuitry spans approximately 1.700 inches and includes approximately 300 controlled cryotron gates and their interconnecting paths. A practical array might include 100 such words and 0.200-inch bit line extensions to facilitate interconnection of arrays. The overall dimensions of this thin-film array would be, therefore, 1.700 inches by 3.400 inches, including approximately 30 000 interconnected gates and array interconnection pads. A series connection of 50 such arrays would yield the required processor capacity.

*This effort is supported by Rome Air Development Center under Contract AF 30(602)3737.

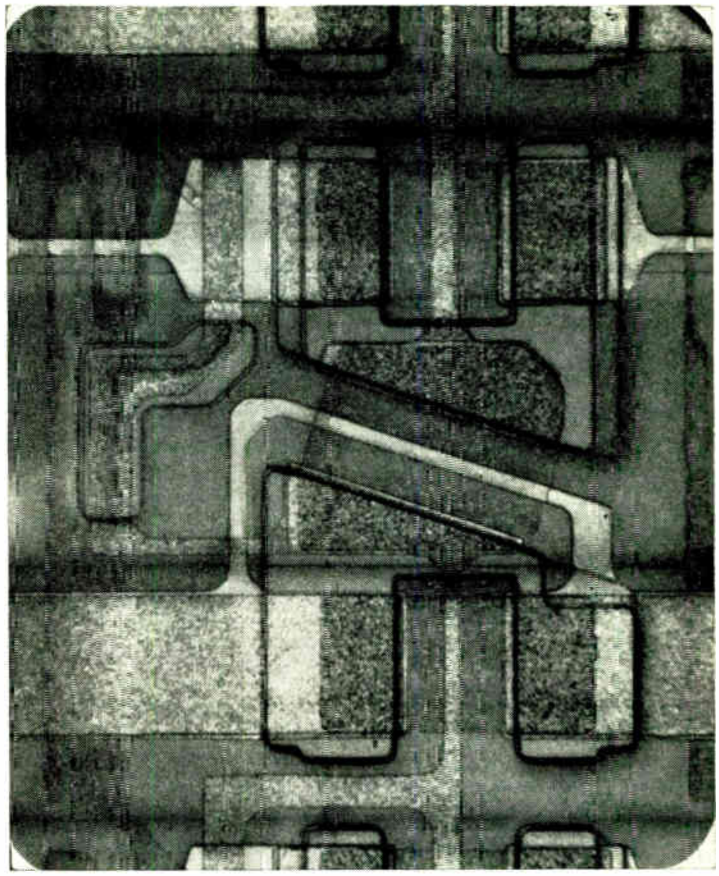
The nature of the process now being used readily permits fabrication of progressively more complex arrays for the purpose of breadboard evaluation of design performance. In the present program this approach has been used to evaluate a four-cell module, then a two-word two-cell-per-word module with control circuitry, and currently a 30-word module of 750 cells with control circuitry. Figure 9 shows one of these 750-cell arrays, a single cell of which appears in Fig. 8. Experience with this progression of modules has yielded a design with broad operating tolerances, the tightest of which appears experimentally to be 20 percent about the design center. The final module size for the 5000-word associative processor has been restructured to a seven-layer form, with an attendant loss of approximately 20 percent in cell density. Thus, the 2- by 4-inch arrays will each contain only 80 words of circuitry rather than 100, but will require approximately 20 percent fewer fabrication steps, and therefore result in a comparable reduction in exposure to the contamination characteristic of current laboratory-scale manufacture. Sixty-three of these substrates will form the full system, which leads us quite naturally into a discussion of the problems of array interconnection.

Interconnection, data linkage, and packaging

We will continue to use the cryotron associative memory to illustrate requirements for interconnection of the thin-film arrays, or substrates, to each other and to appropriate room-temperature electronics. Since word-control circuitry is produced integral to the thin-film array of word-organized cells, signal flow between arrays can be restricted to the bit direction. For each bit position within a word length there exists an input and an output path. Twenty additional paths are required to power and control the word-control circuitry. In the array illustrated in Fig. 9, the interconnection requirement results in 70 interconnector pads along each of two opposing sides of the array. For reference these thin-film pads are 0.007 inch wide, on 0.014-inch centers, and 0.200 inch long.

Simplicity, reliability, and maintainability considerations dictate the use of flexible-strip transmission lines to continue the superconductive paths between adjacent arrays. Pressure contact is made between the multiple conductors of the cable and the thin-film pads of the joined arrays. Although no superconductive loops are required to span from array to array, a superconductive contact is desired to inhibit needless Joule heating in the liquid-helium coolant. The shield plane of the transmission line is also superconducting to provide the high degree of magnetic-flux confinement afforded by its diamagnetic properties and, therefore, a maximum decoupling among parallel signal paths. This means of interconnection is certainly amenable to repair; however, an evaluation of the reliability of large numbers of pressure contacts in series and over prolonged operating periods remains to be shown. This form of contact to individual array samples has been a standard and successful practice

Fig. 8. Top view of an associative storage site in the nine-layer processor array. The cell module is 0.030 inch high in the bit line direction and 0.028 inch in the word direction.



for the past year; thus, it appears likely to be satisfactory. Pressure contacts at room temperature are vulnerable to gradual degradation because of the formation of non-conducting compounds by chemical reaction of the contact materials with their ambient. The pressure contacts for cryoelectronic purposes are immersed in liquid helium; thus, this mode of contact degradation is unlikely, in view of the unavailability of ambient materials for reaction and of the virtually dormant state of such chemical reactions. The question of degradation upon repeated cycling to room temperature remains unanswered for large numbers of interconnections; however, no degradation has been observed in individual samples maintained under pressure contact and periodically cycled to liquid helium for re-evaluation of connectors as well as device characteristics. It is worth noting that no drift in cryotron circuit characteristics has been observed in the preliminary lifetime studies.

Figure 10 shows four 30-word associative processor planes joined by three flexible-strip transmission lines and formed into a seemingly continuous ring via another flexible interconnector whose signal paths have been flared to form expanded contacts. Each of the multiple signal paths threading the ring of interconnected substrates is provided with a pair of expanded contacts in this region for adaptation to a data link with room-temperature electronics, such as current drivers and sense amplifiers. The diamagnetic shield or ground plane is continued under the data-link contacts, thereby completing the high-frequency decoupling properties of the ring of interconnected arrays.

Data and control information are sent in parallel to the cryotron circuitry as fractional-ampere current pulses; output data from the memory appear as fractional-millivolt pulses on separate and parallel paths. Approximately 120 input current paths and 50 output voltage paths constitute the data linkage for an associative processor having a 50-bit word length. The design of suitable cabling requires judicious compromises between desires based on electrical, mechanical, and thermal considerations. Since the data link presents a high-thermal-conductivity path between the temperature extremes, the individual conductor cross sections should be minimized. The data link is divided into two distinct shielded cables immediately upon departure from the superconductive interconnector to reduce crosstalk between high- and low-level signal paths. Common-mode interference between paths in each of these cables can be reduced by a variety of

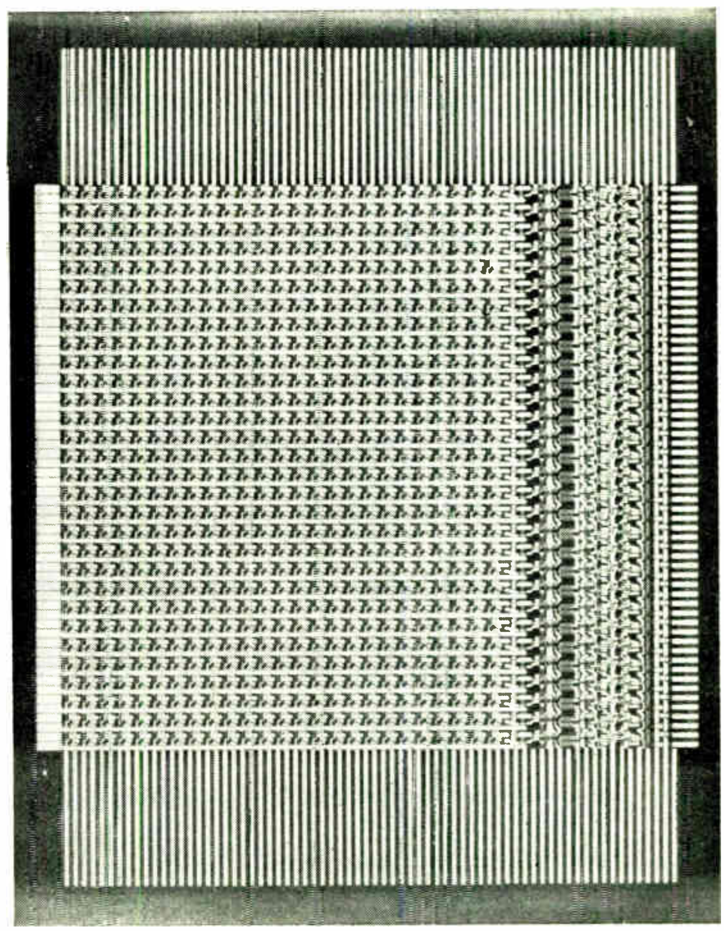


Fig. 9. Top view of a nine-layer associative processor array with 30 words of 25 cells each; word-control circuitry appears on the right. Lands for 70 interconnection points appear at top and bottom of this 1.5- by 1.5-inch specimen.

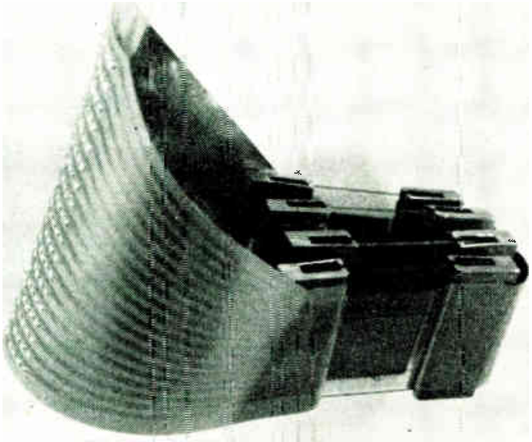


Fig. 10. Signal-path interconnection means for a multiple-array assemblage. Seventy signal paths are serially continued between four arrays via flexible, superconducting-strip transmission lines. The terminals are rendered suitable for connection to system data link via expanded contact pattern visible on left-hand interconnector.

standard practices. At present, very serviceable performance is achieved through use of twisted pairs of No. 38 insulated wire for both high- and low-level paths. Both balanced- and unbalanced-mode transmission can readily be accommodated by this means. In addition, thermal load to the refrigerant is nominal and the vacuum seal required to isolate the reduced pressure ambient in the liquid-helium chamber is readily formed. There will undoubtedly be numerous engineering improvements in data linkage as experience is gained in the operation of prototype cryoelectronic equipment, particularly as operation is transferred from the open-cycle refrigerators currently used to closed-cycle units specifically designed for these applications.

The closed-cycle cooling power required at 3.5°K for significant cryoelectronic data processing will be in excess of one watt. The heat load stems principally from thermal leakage through the refrigerator walls and along the data-link conductor, assuming an application in which only very small fractions of the total cryotron complement are simultaneously switched. A minimum working volume of 14 liters of liquid helium is anticipated, and this presents a substantial surface area for radiative heating, even with the most efficient of known insulation techniques. This fact, coupled with the necessarily large number of data-link conductors to effectively exploit the device potential in parallel data-processing applications, establishes the base cooling requirement. With successful development of the total technology, cooling requirements may well expand toward the five-watt level.

But where are the refrigerators, what are their operating characteristics, and what are their initial and sustaining costs? There is, unfortunately, little additional information to report with respect to this aspect of the total technology. There appears to be nothing fundamentally prohibitive to the development of suitable closed-cycle units, however. What is obviously lacking is an assured demand for multiwatt units to motivate this development, and a sufficient volume and term of prototype applications to yield significant operational statistics. Although the ultimate cost of refrigeration remains similarly vague, it will necessarily restrict cryoelectronic applications to large-scale data processing.

Summary

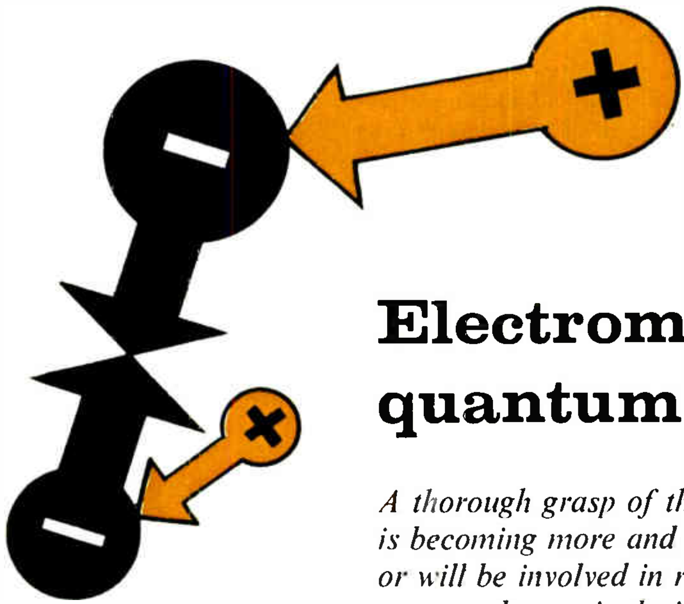
The technological advances in array fabrication of cryoelectronic circuitry described in this survey appear to have overcome the first major hurdles in bringing superconductive devices into contention for computer applications. Problems relating to reliable and repairable means of array interconnection and data linkage for highly parallel, large-capacity data processors appear also to have been adequately resolved. However, the efficiency of these solutions remains to be established in multiple-array assemblages representative of economically competitive systems.

The base cost of liquid-helium refrigeration will principally determine the lower economic bound on information-handling capacity of such systems. In the existing random-access-memory market, cryoelectronics affords a complementary means of providing moderate-speed storage at capacities well in excess of 10^7 bits. The greater potential is for highly parallel data processors not otherwise feasible by virtue of their inherent distributed logic and memory characteristics, particularly for systems requiring a minimum of 10^6 gates.

All remarks pertaining to potential applications are necessarily contingent on the satisfactory completion of a prototype cryoelectronic system in the near future. In the writer's opinion, the obstacles to be overcome are in the development and practice of manufacturing methods based on the techniques described. These obstacles, as opposed to those placed by nature, are subject to solution with sufficient perseverance.

REFERENCES

1. Buck, D. A., "The cryotron—a superconductive computer component," *Proc. IRE*, vol. 44, pp. 482–493, Apr. 1956.
2. Caswell, H. L., "Thin-film circuit technology. Part II—Thin-film superconducting devices," *IEEE Spectrum*, vol. 1, pp. 84–99, May 1964.
3. Pritchard, J. P., Jr., Pierce, J. T., and Slay, B. G., Jr., "Photomask-photoresist techniques for cryotron fabrication," *Proc. IEEE*, vol. 52, pp. 1207–1215, Oct. 1964.
4. Bremer, J. W., *Superconductive Devices*. New York: McGraw-Hill, 1962.
5. Newhouse, V. L., *Applied Superconductivity*. New York: Wiley, 1964.
6. Smallman, C. R., Slade, A. E., and Cohen, M. L., "Thin-film cryotrons," *Proc. IRE*, vol. 48, pp. 1562–1582, Sept. 1960.
7. Burns, L. L., "Cryoelectric memories," *Proc. IEEE*, vol. 52, pp. 1164–1176, Oct. 1964.
8. Brennemann, A. E., "The in-line cryotron," *Proc. IEEE*, vol. 51, pp. 442–451, Mar. 1963.
9. Brennemann, A. E., Seki, H., and Seraphim, D. P., "Dynamic operation of the in-line cryotron in bistable circuits," *Proc. IEEE*, vol. 52, pp. 228–238, Mar. 1964.
10. Slay, B. G., Jr., and Pritchard, J. P., Jr., "Switching characteristics of small-geometry thin-film superconductors," presented at AIME Conference on Integrated Electronics, Sept. 7–9, 1965, San Francisco, Calif.
11. Barnard, J. D., Behnke, F. A., Lindquist, A. B., and Seeber, R. R., "Structure of a cryogenic associative processor," *Proc. IEEE*, vol. 52, pp. 1182–1190, Oct. 1964. (Note appended bibliography.)
12. Crane, B. A., and Githens, J. A., "Bulk processing in distributed logic memory," *IEEE Trans. on Electronic Computers*, vol. EC-14, pp. 186–196, Apr. 1965.
13. Pierce, J. T., and Pritchard, J. P., Jr., "Dielectric properties of thin insulating film of photoresist material," *IEEE Trans. on Component Parts*, vol. CP-12, pp. 8–11, Mar. 1965.



Electromagnetism and quantum theory

A thorough grasp of the fundamentals of quantum theory is becoming more and more of a necessity for engineers who are or will be involved in research and development work on new electronic devices and components

Dale M. Grimes *The University of Michigan*

Devices dependent upon such results of quantization as plasmas, semiconduction, or individual energy-level transitions now occupy a diverse and important position within the broad scope of electrical engineering. Their further development requires a thorough understanding of quantum theory on the part of device research and development engineers. To meet this need, many departments of electrical engineering may soon find it necessary to offer their own courses in quantum theory, just as they did in electromagnetic field theory after World War II.

The science of electricity and magnetism is in several ways all that a science should be. There is available a vast complex of pertinent experimental information that can be understood both qualitatively and quantitatively. Through use of this information it has become commonplace to predict events by solving appropriate differential equations and matching boundary conditions. Furthermore, the basic equations are firmly founded on a few general postulates: that a scalar invariant charge exists, that a linear transformation correctly describes the relationship between inertial frames, and that the velocity of propagation of electromagnetic energy is constant when measured at two or more locations and times. Even more ideally, the logical development of the laws is general enough to permit the a priori assessment of the limits of their validity: they should be valid for cases ranging from the largest linear dimension down to the diameter of the effective "point" charge, so long as the charge is not deformed.

It is well known, however, that this ideal structure appears to break down for an electron in the vicinity of a nucleus. Two charges cannot be held in equilibrium by electrostatic forces alone; so for stability there must be relative motion. But if there is relative motion, then the accelerated charges must radiate energy; and if energy is radiated, the electron will spiral inexorably toward the nucleus. Since atoms are observed to be stable, it is usually concluded that the electrodynamic laws do not apply under atomic conditions. Thus a critical result of the laws of electromagnetism appears to be invalid well within the expected range of their validity.

In the studies of the structure of the atom that dominated physics in the first third of this century, this conclusion produced much controversy. Many physicists tried to avoid or refute it, with universal failure. Feeling against it was so strong that only great imagination and courage enabled Bohr[1] to proceed on the assumption that accelerated electrons in an equilibrated atomic environment do not radiate, and thus to make his first major contribution. Under critical prodding, primarily by Einstein,[2,3] he was chiefly responsible for interpreting modern quantum theory. His interpretations were able to take into account not only observed phenomena but also the flood of new experimental results pertaining to atomic structure; hence they have come to be the orthodox view.

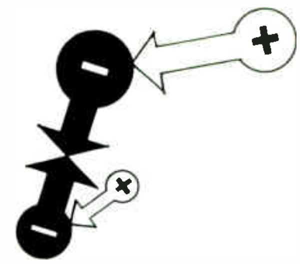
But accepting this view requires more than the denial of the validity of the electromagnetic equations. It requires that statistics be applied to the performance of individual particles in such a way that the deterministic concept that every event is the unique result of preceding ones is rejected. This rejection is without precedent, to say the least; but most of the physicists active in the development of the theory have come to accept it as a necessary consequence of the equations. Many others, however, have taken exception, including Bohm,[4] de Broglie,[5] and Feynman.[6] Indeed, Bohr is quoted as saying that only after two years of argument and propaganda was he able to convince Heisenberg and Bloch that the physical results made it mandatory to accept the philosophical interpretation.[7] And surely this objection was what made Einstein write to Born,[8] shortly before his death, ". . . your quantum theory is not the real McCoy."

The typical graduate course in quantum theory, for students in either engineering or the physical sciences, simply ignores the philosophical difficulty. It begins with the observed wave character of matter and the de Broglie relation, goes into duality and uncertainty principles, and arrives at the Schrödinger equation, which, from that point on, is the basis for calculating events on the atomic scale of dimensions. The equation is given a role on a par with relativity and supersedes the Maxwell equations as necessary. The rest of the year-long sequence is spent in solving the equation under a variety of conditions and extracting conclusions from the solutions. One's expertise in quantum theory is measured by his cleverness in substituting approximate mathematical models for particular physical situations.

This approach is in a sense a very reasonable one. It is logically self-consistent, it is pedagogically convenient for one trained in classical mechanics, and it conforms to the mainstream of thought for the development of quantum theory. And the Schrödinger equation serves as a firm base, for it shares with the Maxwell equations the property that, when it is appropriately modified to meet relativistic requirements, the quantitative conclusions derived from it are found to match a wide range of experimental data. No reasonable doubt of its correctness remains. But there are difficulties. One is that since the equation has no logical basis, there is no way of determining the limits of its validity on an a priori basis. Further, the fact that many years of work based upon it have brought neither an adequate particle theory[9] nor a unified field theory seems to indicate that the logical structure may be inadequate to support these extrapolations (although it is difficult to recognize lack of progress to an unknown objective).

In describing orthodox quantum theory, Born[10] states that a mechanical process is accompanied by a wave process—i.e., the guiding wave—which is described by the Schrödinger equation and whose significance is that it gives the probability of a definite course for a mechanical process. He finds that the wave mechanical interpretation accounts for the apparent absence of radiation from an electron in a nuclear electrostatic potential well. He says that the interpretation implies that elements of radiation from different moving parts annul each other by interference. His argument does eliminate, or at least evade, the difficulty with the electromagnetic field equations, but does nothing toward solving the philosophical problem; and it is so metaphysical that it may well be worse than the disease it is designed to cure. Moreover, it certainly is not an obvious conclusion from the equations.

Two critical questions emerge. First, how important is the interpretation of a thing as long as the mathematics is correct? Second, is Born's interpretation correct, i.e., does it correspond in some sense to physical reality? With regard to the first question, Einstein is reputed to have said that one should study what physicists do, not what they say. As for the second, who knows the interpretation of anything? In our culture an interpretation is judged by its consistency, its simplicity, and its agreement with observation. The Babylonians[11] provide an extreme example of the application of such rules. They developed a system to predict quantitatively the movements of the stars. All but seven—the sun, the moon, Mercury, Venus, Mars, Jupiter, and Saturn—were fixed in the firmament, which revolved about the earth-mountain once a day, under the guidance of the gods; the seven vagabonds revolved too, but had additional, predictable motions of their own. This theory met all the requirements of modern science. The difference is that the part of it not susceptible to experimental test does not satisfy modern standards: when we discuss mechanical processes, we prefer guiding waves to guiding gods.



The importance of a good physical model is obvious. A very famous example is Maxwell's eminently successful formulation of electrodynamics,[12] which, in the main, is distinguished from earlier mathematical formulations by being based upon Faraday's model of lines of force. More mundanely, it is often the case that meaningful approximations are made possible only by the use of physical reasoning to guide the selection of idealized models for analysis, to motivate the choice of approximation methods when necessary, and to give meaning to the coefficients in the power-series expansions. But new, perhaps more useful, interpretations or ideas are not always welcomed by those who accept the old ones. The cases of Robert Goddard and Frank Lloyd Wright, recently discussed by Bronwell,[13] make particularly interesting reading. Goddard was recognized only posthumously, and Wright was recognized during his lifetime primarily because he lived to be an old man.

This article grew from an attempt to answer the question of how best to approach quantum theory for graduate students in electrical engineering. It submits a possible alternative to the orthodox concept of the wave functions. It meets all consistency requirements; it retains the validity of the electromagnetic field equations down through atomic dimensions; it retains determinacy in physical events, thus satisfying Einstein's objection; and it places quantum theory as a derived result of electromagnetism, thus supplying an a priori limit to its range of validity. Somewhere, however, there is always a weakest spot. In this case it is necessary to consider a very intense, high-frequency electromagnetic radiation field to exist throughout the universe. Whether this interpretation corresponds to physical reality and whether it will bring the physical insight necessary for extrapolating beyond the curtain of unsolved problems are yet to be seen.

Construction of the Schrödinger equation

The interpretation described here follows from the analysis of a particular model by use of the Maxwell equations and Newton's laws. The model has four parts.[14] The first is a particle with charge q and mass m whose physical extent is much smaller than other distances of interest; it has often been considered a point charge. The second is a nucleus, which has an infinite mass and an attractive electrostatic force field that traps the electron. The third is an intense electromagnetic radiation field in which the two charges are immersed, and the fourth is a very large box whose walls are perfectly conducting at all frequencies and which contains the ensemble.

The analysis of this model, for small velocities and accelerations, should fall within the range of validity of the Maxwell equations. But there are two difficulties. Although the special theory of relativity is general enough to serve as a basis for calculating effects by and on charges at all velocities, the results for large velocities are nonlinear.[15] Furthermore, two additional conditions must be met. The electron must not deform under acceleration, and the time rate of change of acceleration must satisfy the inequality

$$\dddot{v} << \frac{c}{d}\ddot{v}$$

where d is the diameter of the charge, or else the structure of the particle must be taken into account. It is not adequate to treat the electron as a point charge, since the result depends upon the detailed structure of the charge and its deformation during acceleration. As a result, the flight of an electron next to the nucleus remains an unsolved problem. Direct methods cannot be used without detailed knowledge of electron structure, and even with such knowledge the equations would be nonlinear.

Fortunately, for many purposes a detailed solution is neither necessary nor desirable. Since the electron moves so rapidly that all observations must be made over time intervals very much longer than the average orbital period, average values over fairly long periods are the only concern. This is very similar to the case for thermodynamics,[16] where complete microscopic information would be so complex as to be almost useless on the macroscopic scale of dimensions, but time averages of the different quantities are of fundamental significance. For example, detailed positional and momentum coordinates of each molecule in a gas contain far more information than a single number giving the temperature, but are hardly as meaningful to a man interested in his comfort. The question, then, is how to extract the average information from the model in the absence of a detailed solution. As is usual in such problems, one must resort to statistics.

An electron in a nuclear electrostatic well will produce radiation having a certain pattern and frequency, which will in turn produce a radiation reaction. The details of this process are still unknown, but the occurrence of radiation is predictable from Maxwell's equations. The emitted radiation adds to that originally present and is stored inside the box. The radiation field, in turn, drives the electron. Let us postulate that it is possible to establish a dynamic equilibrium between the particle and the radiation field. Thus the model originally posited becomes one of an electron, in the vicinity of the nucleus, whose detailed motion is unknown but which is known to be driven by a stochastic electromagnetic radiation field in a dynamic equilibrium. The electron will then have a fixed time-average value of kinetic and potential energy, determined by the properties of the charges and of the driving field, and will produce a negative-real time-average charge density at each point within the box, and a corresponding time-average charge density at each point in momentum space. Because of the electrostatic force field, the average momentum of the electron will be a function of distance from the nucleus; so the charge densities will be functionally related.

The problem is formally to find possible ergodic solutions to a stochastic force field driving a bound charge. The usual procedure in such cases is first to obtain and then to solve the differential equation of motion. The difficulty is that only part of the continuous portion of the differential equation is known in detail, and that part includes nonlinear terms. In the absence of the equation and its solution, the procedure is to consider a specific related problem that can be solved, solve it, and attempt to generalize the solution. Here the solvable problem is one of a damped, stochastically driven, harmonic oscillator, for which the differential equation of motion is

$$F(t) = \dot{p}(t) + \beta p(t) + \omega_0^2 x(t) \quad (1)$$

where $F(t)$ is the stochastic driving force, ω_0 is the oscillating frequency of the undamped oscillator, β is the damping coefficient, $p(t)$ is the momentum of the electron, and $x(t)$ is the position of the electron at times between t and $t + dt$. If $F(t)$ is independent of and varies much more rapidly than $p(t)$, then $\rho(x, p)\,dx\,dp$, the resulting fraction of the time that the charge is at a position between x and $x + dx$ with momentum between p and $p + dp$, will be described by a Gaussian distribution in both x and p:

$$\rho(x, p) \sim \exp\left\{-\frac{1}{(1-\sigma^2)^{1/2}}\left[\left(\frac{x}{a}\right)^2 - \frac{2\sigma xp}{ab} + \left(\frac{p}{b}\right)^2\right]\right\} \quad (2)$$

In the limit of measurement times long enough to obtain an ergodic solution, σ goes to zero.[17] The normalized charge density at each point in either coordinate or momentum space can be found by integrating Eq. (2). For example, the total charge density at position x is the sum of charge densities due to all possible values of momentum,

$$\rho(x) \sim \frac{1/a}{(2\pi)^{1/2}} \exp\left(-\frac{x^2}{a^2}\right)$$
$$\rho(p) \sim \frac{1/b}{(2\pi)^{1/2}} \exp\left(-\frac{p^2}{b^2}\right) \quad (3)$$

where

$$a = \left(\frac{kT}{m\omega_0^2}\right)^{1/2}$$
$$b = (mkT)^{1/2} \quad (4)$$

and where m is the mass associated with the charge and kT depends upon the energy associated with the stochastic force field. With these results, the specific problem is considered to be solved.

However, the main problem is how to proceed from the solution found to the unknown desired solution. Equation (2) cannot be correct for all velocities, since the nonlinearities occurring are proportional to the square of the velocity.[15] (Also, the damping term is proportional to the time rate of change of acceleration and not to the velocity.) Therefore, Eq. (2) must be incomplete. As for the requirements of $F(t)$, the model chosen appears to satisfy directly the condition that $F(t)$ not be directly correlated with $p(t)$.

It is not clear, however, that the condition that $F(t)$ must vary rapidly compared with $p(t)$ is met. Indeed, if $F(t)$ varies at the mean radian frequency of the electron the condition is violated. However, for relativistic velocities the radiated field is not at the radian frequency but extends throughout the frequency range, scaled up by a factor of γ^3, where $\gamma = (1 - v^2/c^2)^{-1/2}$. An observer on the equatorial plane, then, would see pulses of radiation at a repetition rate of ω_0 including frequencies through $\gamma^3\omega_0$.

For nearly periodic motion and for large distances from the potential source, Eq. (1) must be nearly correct; so solutions are necessary that contain Eqs. (3) as asymptotic forms for large values of the variables. Since for a negative charge the result must always be a negative-real charge density, the solution must have the form

$$\rho(x) \sim f^2(x) \exp(-x^2/a^2)$$

$$\rho(p) \sim g^2(p) \exp(-p^2/b^2)$$

where the functions $f(x)$ and $g(p)$ vary much more slowly than the exponentials for large values of the arguments. It is convenient to work with the normalized square root of the equations as

$$U(x) = f(x) \exp(-x^2/2a^2)$$
$$\phi(p) = g(p) \exp(-p^2/2b^2) \quad (5)$$

By the formula for the dipole moment of a charged volume, the dipole moment of the time-average charge distribution is zero only if $f^2(x)$ and $g^2(p)$ are even functions. As a result, $U(x)$ and $\phi(p)$ may be either even or odd functions of the respective variable, but not an arbitrary linear combination thereof.

As mentioned earlier, the two charge densities, and therefore $U(x)$ and $\phi(p)$, are functionally related. It is desirable to determine the conditions under which a linear transformation from one to the other can be made. The Fourier integral transformations form a very general set of unitary transformations.[18] These are

$$U(x) = K \int_{-\infty}^{\infty} \phi(p) \exp(-ixp/\hbar)\, dp$$
$$\phi(p) = K \int_{-\infty}^{\infty} U(x) \exp(+ixp/\hbar)\, dx \quad (6)$$

where K and $\hbar$ are constants to be determined. Equations (3) and (6) can be simultaneously satisfied only if $K = (2\pi\hbar)^{-1/2}$ and $\hbar = 2ab$, or, from Eq. (4), if

$$\hbar = \frac{kT}{2\omega_0} \quad (7)$$

That is, the metric determining constant $\hbar$ is a measure of the ratio of energy in the radiation field at the interacting frequency to the resonant frequency of the oscillator.

Once the asymptotic form has been used to determine the constants of (6), the next step is to determine the possible functions $f(x)$ that satisfy (5) and (6). Those functions form an infinite array, the first few terms of which are proportional to 1, x, $(2x^2/a^2 - 1)$, etc. Each successive term produces an increased value of average distance from the center and an increased average momentum magnitude.

The foregoing solution, as extended to include nonlinearities, is complete. It is instructive, however, to reformulate the problem in terms of energies. The time-average energy W of the electron is the sum of average kinetic and potential energies. The kinetic energy is proportional to the square of the momentum, and the potential energy $-e\Phi(x)$ is a function of electron position. Since the functions $U(x)$ and $\phi(p)$ represent time averages at a point, the time-average value of the energy can be written as

$$W = \frac{1}{2m}\int p^2\, \phi^*(p)\, \phi(p)\, dp - e\int \Phi(x)\, U^*(x)\, U(x) dx \quad (8)$$

where the asterisk represents the complex conjugate. The first integral on the right can be evaluated if $\phi(p)$ is known. However, $\phi(p)$ need not be evaluated explicitly; instead, one can substitute the equations of (6) twice into Eq. (8) and carry out two partial integrations to obtain the resulting integrodifferential equation

$$\int U^*(x)\left[\frac{\hbar^2}{2m}\frac{d^2U(x)}{dx^2} + e\Phi(x)\, U(x) + WU(x)\right]dx = 0 \quad (9)$$

in which the integration is over the volume of the box, as the equation that the electron must satisfy. Without the integration, (9) is the one-dimensional time-independent Schrödinger equation. Of course, (9) is satisfied if the integrand is required to be zero at all points; so the Schrödinger equation appears as a nonunique result of the application of statistics to the nonlinear system. This subject is further discussed in other sources,[19-21] where (9) is derived using purely statistical methods instead of Fourier integral expansions.

In summary, a result of only classical physics is that a particle in a one-dimensional harmonic potential well in equilibrium with a stochastic radiation field will satisfy Schrödinger's equation. It is clear that the use of this model is a deterministic use of statistics; that is, statistics are not used in a way that requires the rejection of the concept of determinism in physical events, but appear in the same logical role as in conventional communication theory or statistical mechanics.

The foregoing result can be generalized to include time dependence. The momentum density is proportional to charge velocity, which, in turn, is proportional to the time-average current density. The expression

$$J_e = \frac{e\hbar}{2mi}\left[U^*(x)\frac{dU(x)}{dx} - U(x)\frac{dU^*(x)}{dx}\right] \quad (10)$$

gives the effective current density in the sense that, when it is averaged over all positions within the box, it gives correct results; but it does not necessarily indicate the current density at each point. The Schrödinger equation can be generalized by applying the continuity equation of electric current, using the charge density at each point as calculated from the function $U(x)$ and the current density at each point as calculated from the effective current density.

U^*U is, by definition, the time-average charge density at a position in space and is, therefore, independent of time. However, the time-dependent function $\psi(x, t) = U(x)e^{i\omega t}$ can be introduced with an arbitrary ω since the product $\psi\psi^*$ remains time-independent. Introducing the time into the exponential in this way will in no sense produce an actual time-sequence of events. When the effective current density and the function $\psi(x, t)$ are substituted into the continuity equation for electric current, the result, after some manipulation, is

$$\int dx \left\{\psi^* \left[\frac{\hbar^2}{2m} \nabla^2\psi + i\hbar \frac{\partial\psi}{\partial t} + e\Phi(x)\right] + \begin{matrix}\text{complex}\\ \text{conjugate}\end{matrix}\right\} = 0 \quad (11)$$

Equation (11), like (9), is integrodifferential. Comparing the two equations shows that if

$$W\psi = \frac{\hbar}{i}\frac{\partial\psi}{\partial t}$$

then the integrand of (11) is zero, and gives the Schrödinger time-dependent differential equation as a nonunique result. The average oscillating frequency of the electron is, clearly,

$$\omega = \frac{W}{\hbar} \quad (12)$$

Equation (11) results from the interaction of the charge and the background field, and therefore it should apply wherever the extent of the particle is much less than the distances of interest in the rest of the problem. In other words, it should apply to problems involving the behavior of atoms or of individual electrons, but not to those in which the wavelength of the radiation field is larger than the other dimensions of interest, such as problems of the nucleus.

Energy-level transitions

According to the orthodox view, the Schrödinger equation and the laws of electrodynamics contain the complete solution for the absorption and emission of radiation by an atomic system. The quantization of the radiation process—that is, the absorption and emission of radiation in energy packets whose energy is proportional to the frequency—follows directly. The analytical procedure is to apply first-order perturbation theory and the principle of the conservation of energy to the Schrödinger equation, then to consider the atomic system to be located inside a box (just as in the theory advanced here), and finally to evaluate the interaction of the confined radiation and the atomic system in the limit as the dimensions of the box become very large.

In recent years techniques for handling nonlinear systems have improved dramatically, mainly because of the impetus of engineering problems. The equations of Manley and Rowe[22] are of particular interest, for they relate power-flow limitations in a nonlinear system to the frequency of the power, independently of the source of the nonlinearity and subject only to the system's being lossless. These equations can be directly applied to an atomic system, without going through the artifice of introducing the average time, to describe the continuity of current flow, and ultimately the time-dependent Schrödinger equation.

The Manley–Rowe equations are as follows:

$$\sum_{m=0}^{\infty}\sum_{n=-\infty}^{\infty} \frac{mP_{mn}}{m\omega_0 + n\omega} = 0$$
$$\sum_{m=-\infty}^{\infty}\sum_{n=0}^{\infty} \frac{nP_{mn}}{m\omega_0 + n\omega} = 0 \quad (13)$$

where m and n are integers, ω_0 is the equilibrium frequency of the nonlinear oscillator, ω is a perturbing frequency, and P_{mn} is the power flow into the system at the frequency $(m\omega_0 + n\omega)$. Shortly after the first paper of Manley and Rowe appeared, Weiss[23] showed that the relations also follow from the quantization of atomic transitions.

One can also reverse the arguments to derive the quantization of emission and absorption of electromagnetic energy from the model introduced in the preceding section and the Manley–Rowe equations. To analyze the radiation–atom interaction, let us postulate a unique, though unknown, relationship between the frequency of the driving stochastic field and the electron's energy. Consider the atomic electron to be in equilibrium at frequency ω_0, and a perturbing radiation at frequency ω to fall upon the system.[24] Another possible but unoccupied equilibrated electron energy level occurs at the frequency $\omega_0 \pm \omega$. For this case the two equations of (13) become

$$P_{0\pm1} = \mp\frac{\omega_0 \pm \omega}{\omega} P$$
$$P_0 = \pm\frac{\omega_0}{\omega} P \quad (14)$$

where $P_{0\pm1}$ is the power flow at frequency $\omega_0 \pm \omega$, and P is the power flow into the system at frequency ω. The total energy absorbed or emitted at frequency ω is the time integral of the power, or

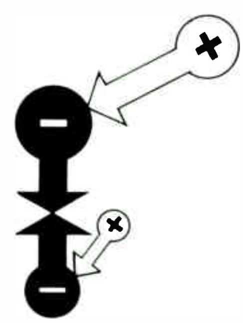
$$\Delta W = \int_0^\tau P\,dt = \pm\frac{\omega}{\omega_0}\int_0^\tau P_0\,dt$$

where τ is the time required for the energy exchange to take place. Since the second integral represents the energy stored at frequency ω_0, and it is proportional to ω_0 and independent of either P or τ, the energy exchanged by the atom at frequency ω is

$$\Delta W(\omega) = \pm\hbar\omega \quad (15)$$

or, in other words, energy can be absorbed or emitted from a continuous radiation field only in quantized units. Comparison shows that the $\hbar$'s in (12) and (15) are the same.

The question of particular interest is whether radiation itself or only its interaction with matter is quantized. The quantization of the interaction processes is a sufficient basis from which to derive the blackbody equations, as was originally done by Planck.[25] Einstein added a more convenient derivation, which also required only that radiation interact in quantized units.[26] However, certain other processes cannot be handled so simply. With regard to the photoelectric effect, for example, it was not possible to explain, using only classical theory, why an intense monochromatic radiation field at a frequency just below the critical frequency does not produce current flow, whereas a very weak radiation field just above the critical frequency does.



Other effects too, in particular the Compton effect, indicate a unidirectional radiation packet.

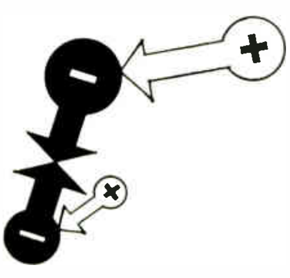
Current thought on the quantum theory has also been influenced by considerations other than experimental evidence. Since quantization is obviously present, even though there is no basis for expecting it in atomic theory, through analogy it seems reasonable to believe that radiation is quantized. In addition, Bohr and Rosenfield showed that it was consistent with the interpretation of quantum theory to ascribe quantization to the electromagnetic field per se.[27] So the evidence seems overwhelming that not only the energy transferral process but radiation itself is quantized.

The argument of the preceding section does provide a logical basis for the Schrödinger equation. Moreover the argument of this section—that is, the application of the Manley–Rowe equations—provides a basis for accepting the quantization of the interaction process; however, it contains no basis for expecting radiation itself to be quantized. So the analogical basis can be discounted. In spite of widespread opinion, then, there is no logical support for field quantization; as Rosenfield himself recently pointed out,[28] there is only an absence of arguments against it.

This leaves experimental evidence. As remarked earlier, the blackbody equations do not imply field quantization. With the use of the Manley–Rowe equations, the photoelectric effect can be understood without assuring field quantization, for the threshold field results when the incoming frequency combined with the equilibrium frequency equals the extremum driving rate for an electron in equilibrium with its nucleus.

The question is often raised as to how a single atom could selectively receive the incident energy from an area much larger than its geometrical cross-sectional area. This question is not particularly significant, however, since the concept that the ratio of the geometric area of a radiator to its cross-sectional area for signal detection should be of the order of unity is based upon physical intuition. The effective cross-sectional area of so simple a radiator as an infinitesimal dipole is 0.12 λ^2, regardless of the physical cross-sectional area; therefore, an enormous ratio of effective to geometric area is not surprising. The effective area is determined by the inductive field of the resonant radiator. Among a group of electrons, then, presumably one can obtain slightly more than an equivalent share of energy, producing a reactive field that funnels still more energy in, etc., until the transition is complete.

The absorption of radiant of a wavelength that is very large compared with the radiator is derivable on a strictly classical basis, according to both Chu[29] and Taylor.[30] If we consider the atom as a transmitter, it can emit energy into the far field at arbitrarily large wavelengths, but only if the Q is sufficiently large. Concomitantly, the bandwidth must be small, in excellent agreement with observation; and, invoking the principle of reciprocity, it must also receive energy.



It may be concluded, then, that neither the experimental nor the theoretical evidence is sufficient to lead unambiguously to a particular interpretation; for this reason, the question of field quantization remains open.

The conducting box

In the model postulated here, the conducting box contains the radiant energy. In reality, however, the question is "What kind of physical mechanism can confine radiant energy within a closed system?" Of possible mechanisms advanced in other contexts, one in particular would explain the radiant field, but as part of an open system. It is associated with the steady-state model of the universe. In that model, proposed by Hoyle[31] and by Bondi and Gold,[32] the major external features of the universe undergo no change. New matter is continuously being created from nothing, forming into stars, and expanding away indefinitely. In a steady-state universe, an argument first presented by Olbers[33] in 1826 remains valid. Olbers begins by placing the origin of coordinates at an arbitrary point in space and constructing a differential volume at an arbitrary distance from it that is small compared with the universe but large enough to contain a great many galaxies. The radiated energy reaching the origin of the coordinates is inversely proportional to the distance from the differential volume and proportional to that volume, which, in turn, is proportional to the square of the distance. The result is that each differential volume throughout space contributes equally to the intensity at the origin.

If radiation is summed over all space, the intensity at the origin should be infinite. However, radiation from some bodies farther from the origin will strike bodies nearer it; so the total radiation temperature of each point in space can be shown to be the same as the radiation temperature at the surface of each body. By this reasoning, since the number of stars and the extent of the universe are taken to be infinite, the background sky should be bright rather than dark. (The argument is that in whatever direction one looks, one should be looking directly into a star, though it may be very distant; and therefore the background sky should be bright.) Any theory of cosmology, then, must include an explanation of why the sky is dark. That advanced by Hoyle and by Bondi and Gold does so by reference to the expansion of the universe. As a result of the Doppler effect, radiant energy received from receding stars is decreased. Because the stars are receding at velocities in proportion to distance, the radiation from the very distant ones is vanishingly weak and contributes so little to the equilibrium radiation that the background sky appears black.

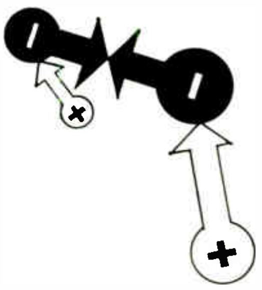
Since quantized radiation occurs only when electrons undergo transitions, and the radiation field postulated here is radiated continuously, one would expect the field resulting from quantized processes to be much less than that from continuous processes. As a result, there is a range of possible steady-state universes in which there will be a background equilibrium continuous radiation field even though the background sky is dark. All theories of cosmology seem to have set the continuous field equal to zero simply by fiat.

The possibility of there being something omnipresent throughout the universe has been discussed ever since Berkeley first criticized the writings of Newton.[34] His criticisms, based upon pure logic, have been perpetuated through the writing of Mach[35] and Einstein[36] to such present-day writers as Dicke[37] and Sciama.[38] The contention is that if Newton's laws are correct at a point in space, then that point must have some characteristic determined by the rest of the universe. But this subject, though it has many relevant ramifications, is outside the immediate scope of this article.

Such a field, if present, might be the source of the interstellar matter, which the "steady-state-universe" school has presumed to be created from nothing; and it would increase the total energy of the universe manyfold over that contained in the known matter. The entire question of the relationship between field and matter appears to be another case of "Which came first, the chicken or the egg?"

Conclusions

The postulate that a background electromagnetic radiation field exists appears to be a sufficient base upon which to build at least most of modern quantum theory. It leads to the conclusion that the problem of the interaction of two charges cannot be isolated but must be considered within its proper cosmological setting. The price of isolating it is that one must impose the wave interpretation and, concomitantly, the loss of determinacy upon what may otherwise be considered a rather conventional statistical mechanical result.



An experimental check of the theory would be difficult. The problem is that since the field is in equilibrium with all processes, an experiment must be devised to detect the particular field upon which the functioning of the detecting instruments depends. Circular reasoning here seems unavoidable.

One check is the simplicity of the logical foundations. The proposed interpretation appears to have a clear advantage over the orthodox one, in that it is deterministic. A second check is whether the physical picture given by the interpretation can provide the physical insight necessary for extrapolation into the next unknown layer of physical phenomena. Can it, for example, provide a background picture on which to construct a unified field theory, or a better picture of particle structure?

In any event, the interpretation described does provide a deterministic basis from which to construct the Schrödinger equation as an offshoot of classical electrodynamics. As such it has proved to be a convenient pedagogical device for students who have a background in electromagnetic field theory.

The author is indebted to Thomas J. Way for his assistance in editing the manuscript.

REFERENCES

1. Bohr, Niels, "On the constitution of atoms and molecules," *Phil. Mag.*, vol. 26, pp. 1–25, 1913.
2. Bohr, Niels, "Discussion with Einstein on epistemological problems in atomic physics," in *Albert Einstein, Philosopher-Scientist*, P. A. Schilpp, ed. New York: Harper & Row, 1959, pp. 201–241.
3. Einstein, Albert, "Remarks concerning the essays," *ibid.* pp. 665–688.
4. Bohm, David, *Casualty and Chance in Modern Physics.* New York: Harper & Row, 1961.
5. de Broglie, L., "L'interprétation de la mécanique ondulatoire," *J. Phys. Radium*, vol. 20, pp. 963–979, Dec. 1959.
6. Feynman, R. P., "Space–time approach to nonrelativistic quantum mechanics," *Rev. Mod. Phys.*, vol. 20, pp. 367–387, Apr. 1948.
7. Nielsen, J. R., "Memories of Niels Bohr," *Phys. Today*, vol. 16, p. 22, Oct. 1963.
8. *New York Times*, European ed., p. 1, July 2, 1965.
9. Dirac, P. A. M., "Foundations of quantum mechanics," *Nature*, vol. 203, pp. 115–116, July 11, 1964.
10. Born, Max, *Atomic Physics*, 7th ed. New York: Hefner, 1962, chap. 5.
11. Koestler, Arthur, *The Sleepwalkers.* New York: Grosset & Dunlap (Universal Library Series), 1963, pp. 20–21.
12. Maxwell, James Clerk, *A Treatise in Electricity and Magnetism*, Vol. II. New York: Dover, 1954, pp. VIII–XII and sections 528 and 529, pp. 175–177.
13. Bronwell, A. B., "Is philosophy off limits?" *IEEE Spectrum*, vol. 1, pp. 110–114, May 1964.
14. Grimes, D. M., "Point charge in a box: a new approach to an old problem," *Proc. IEEE (Correspondence)*, vol. 53, p. 737, July 1965.
15. Sommerfeld, A., *Electrodynamics.* New York: Academic Press, 1952, pp. 297–301.
16. Callen, H. B., *Thermodynamics.* New York: Wiley, 1960, pp. 3–4.
17. Chandrasekhar, S., "Stochastic problems in physics and astronomy," *Rev. Mod. Phys.*, vol. 15, pp. 20–30, Jan. 1943.
18. See, for example, Riesz, F., and Sz.-Nagy, B., *Functional Analysis.* New York: Ungar, 1955, pp. 280–295.
19. Bohm, David, "Hidden variables in the quantum theory," in *Quantum Theory, Radiation and High Energy Physics*, vol. III, D. Bates, ed. New York: Academic Press, 1962, pp. 345–387.
20. Kershaw, David, "Theory of hidden variables," *Phys. Rev.*, vol. 136B, pp. 1850–1856, Dec. 21, 1964.
21. Comisar, G. G., "Brownian-motion model of nonrelativistic quantum mechanics," *Phys. Rev.*, vol. 138B, pp. 1332–1337, June 7, 1965.
22. Manley, J., and Rowe, H. E., "Some general properties of nonlinear elements. Part I: General energy relations," *Proc. IRE*, vol. 44, pp. 904–913, July 1956. Also "General energy relations," *Proc. IRE (Correspondence)*, vol. 37, pp. 2115–2116, Dec. 1959.
23. Weiss, M. T., "Quantum derivation of energy relations analogous to those for nonlinear reactances," *Proc. IRE (Correspondence)*, vol. 45, p. 1012, July 1957.
24. Grimes, D. M., "Energy level transitions," *Proc. IEEE (Correspondence)*, vol. 53, pp. 1663–1664, Oct. 1965.
25. Planck, Max, *The Theory of Heat Radiation.* New York: Dover, 1959, chap. 4.
26. See, for example, Ruark, A. E., and Urey, H. C., *Atoms, Molecules, and Quanta.* New York: Dover, 1964, pp. 60–63.
27. Rosenfield, L., "On quantization of fields," *Nucl. Phys.*, vol. 40, pp. 353–356, Jan. 1963.
28. Rosenfield, L., "On quantum electrodynamics," in *Niels Bohr and the Development of Physics*, W. Pauli, ed. New York: McGraw-Hill, 1955, pp. 70–95.
29. Chu, L. J., "Physical limitations of omnidirectional antennas," *J. Appl. Phys.*, vol. 19, pp. 1163–1175, Dec. 1948.
30. Taylor, T. T., "A discussion of maximum directivity of an antenna," *Proc. IRE (Correspondence)*, vol. 36, p. 1135, Sept. 1948.
31. Hoyle, F., "A new model for the expanding universe," *Monthly Notices Roy. Astron. Soc.*, vol. 108, pp. 372–382, 1948.
32. Bondi, H., and Gold, T., "The steady-state theory of the expanding universe," *Monthly Notices Roy. Astron. Soc.*, vol. 108, pp. 252–270, 1948.
33. Olbers, H., and Jahrbuch, Bodes (1826), as cited in H. Bondi, *Principles of Cosmology.* New York: Cambridge, 1960.
34. Berkeley, George, "Concerning human knowledge," reprinted in *The World's Great Thinkers*, vol. I: *The Speculative Philosophers.* New York: Random House, 1947, pp. 310–315.
35. Mach, Ernst, *Science of Mechanics*, 6th ed. LaSalle, Ill.: Open Court, 1960, pp. 271–291.
36. Einstein, Albert, *Out of My Later Years.* New York: Philosophical Library, 1950. Also Einstein, Albert, *The Meaning of Relativity.* Princeton, N.J.: Princeton University Press, 1955, p. 99.
37. Dicke, R. H., "Cosmology, Mach's principle and relativity," *Am. J. Phys.*, vol. 31, pp. 500–509, July 1963.
38. Sciama, D. M., "On the origin of inertia," *Monthly Notices Roy. Astron. Soc.*, vol. 113, pp. 34–42, 1943. Also Sciama, D. M., *Unity of the Universe.* New York: Doubleday, 1961, chap. 7.

helmshaven and he ordered Lindemann to alter course and attack the shadowing cruiser, *Suffolk*. Fortunately for *Suffolk*, *Bismarck*'s turn was spotted in time and the British cruiser fled amid the water columns of several near misses from the German 15-inch guns.

Kapitän Lindemann knew that the superior naval might of the Royal Navy, the RAF, and Coastal Command would settle for nothing less than the destruction of his ship. Circumstantial evidence indicates that he pleaded with the admiral to return at once to the safety of Wilhelmshaven. He cited the loss of precious fuel oil through damage as one overriding reason for caution. Lutjens, however, was flushed by the victorious encounter against *Hood* and *Prince of Wales* and he was determined to break out into the broad Atlantic where his powerful ship could wreak the most havoc against Allied convoys and smaller warships; then, to swing in a wide arc back toward the Channel port of St.-Nazaire and relative security.

By the evening of May 25, Lindemann's predictions as to British countermeasures were rapidly evolving. As *Bismarck* plowed along at 25 knots on a course south by west, the Admiralty ordered the entire British Home Fleet to leave immediately from Scapa Flow at top speed, deployed to intercept *Bismarck*. The battleships *Rodney* and *Ramillies* were ordered to leave their convoys and proceed northeast (Fig. 2). Two fast cruisers, *London* and *Edinburgh*, were also detached from convoy duty off the west African coast. But the weakness in the marshaling of British naval forces lay in the lack of carrier-borne aircraft to attack the German battleship. The Denmark Strait was at the extreme range of Coastal Command's American-built *Catalina* flying boats and Lockheed *Hudsons*. A number of these craft were equipped with ASV Mark II, a 200-Mc/s (air-surface vessel) search radar capable of detecting the presence of large surface vessels at ranges up to 30 miles. To remedy this situation, Force "H," consisting of the aircraft carrier *Ark Royal*, battle cruiser *Renown*, heavy cruiser *Sheffield*, and six destroyers, was directed to steam immediately (Fig. 2) from Gibraltar to the North Atlantic.

The narrowing circle

The final fix on *Bismarck* by *Suffolk*, as the latter scuttled away under a cover of smoke, gave the British a 500-mile radius of search to cover all possible changes in course during the next 16 hours.

It may seem odd that the Royal Navy should have deployed so massive a naval force to intercept and destroy one enemy battleship. But after *Hood* had been sunk and *Prince of Wales* had been put out of action, the Admiralty realized that *Bismarck* was a behemoth of super-battleship proportions. And indeed she was. In addition to her main battery 15-inch rifles, with an effective range of 20 miles, she mounted a secondary armament of twelve 5.9-inch dual-purpose (capable of antiaircraft fire) guns, innumerable 3-inch antiaircraft, multiple-barrel "pom-poms," and other light, quick-firing weapons. All this, plus radar gear and a 30-knot speed, made her more than a match for any battleship of the British line.

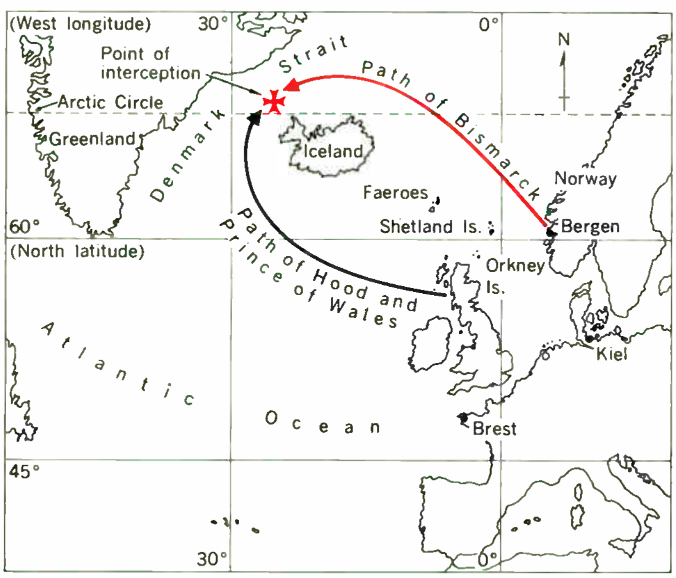
With the exception of *KG-5* and *Prince of Wales*, all of the Royal Navy battleships were 20—or more—years old. *Prince of Wales* had already felt the weight of *Bismarck*'s wallop. *Ramillies* was a veteran of the Battle

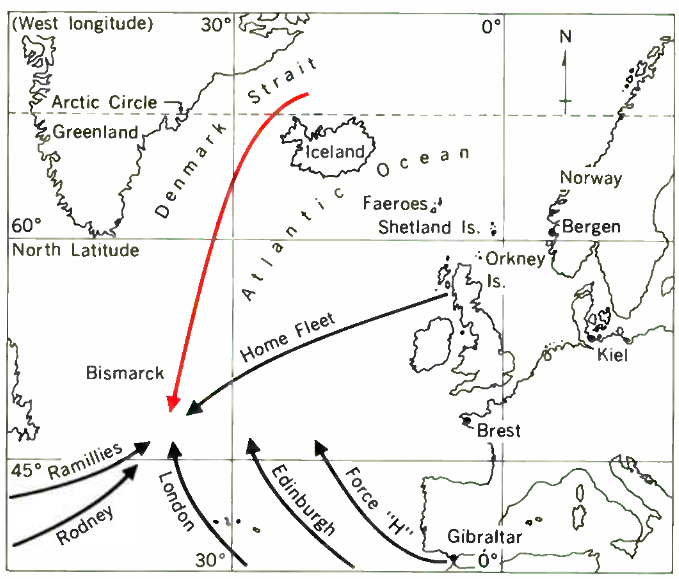
Equivalent naval ranks:

Kriegsmarine	Royal Navy
Admiral	admiral
Vizeadmiral	vice-admiral
Konteradmiral	rear-admiral
Kapitan zur See	commodore, or flag-captain
Fregattenkapitän	captain
Korvettenkapitän	commander

Fig. 1 (left). Map of the North Atlantic showing intercept courses taken by *Hood* and *Prince of Wales* in the pursuit of *Bismarck*. Maltese cross indicates the location in the Denmark Strait of the naval action of May 24, 1941.

Fig. 2 (below). After the sinking of *Hood*, *Bismarck* attempted to break out into the open sea to attack Allied convoys, but fast striking forces of the Royal Navy were being marshaled for the hunter-killer mission.





of Jutland in 1916; *Rodney*, although she mounted eight 16-inch guns, was too slow; *Renown*, a vintage battle cruiser, had the speed, but her armament and armor were inferior to that of *Hood*. Even *KG-5*, the newest British battleship, could not match *Bismarck*'s specifications.

Successive sweeps by the RAF and Coastal Command on May 24–25, through squalls and overcast skies, made no contact with *Bismarck* within the circle of air search.

A significant incident occurred, however, on the 25th. Nine short-range *Swordfish* torpedo bombers from the small aircraft carrier *Victorious* of the Home Fleet, dropping through a hole in the cloud cover, spotted *Bismarck* and launched a torpedo attack. The light, old-fashioned biplanes were met by a withering antiaircraft fire, but they managed to score one hit against the German's starboard bow. The damage, although insignificant, was enough to shake Lutjens' confidence in his ability to elude his British pursuers. The *Swordfish* attackers gave a fix on *Bismarck*'s position.

On the basis of the latest report, radar-equipped *Catalinas* from Loch Erne in Northern Ireland were making systematic search sweeps in the tracking operation. On the morning of the 26th, one of the search planes, flying in thick weather, received a blip on its ASV Mark II radar. Dropping to within a few hundred feet of the surface of the sea, the plane's crew verified the radar contact by visual identification. Bursting flak left no doubt that the object was *Bismarck*.

The net is drawn

Force "H," speeding north–northwestward from Gibraltar, encountered unusually heavy seas, and *KG-5* and *Rodney* were also having a problem—the fuel consumption figures, which had doubled since the vessels were put at flank speed, indicated that they would run out of bunker oil before they could hope to catch *Bismarck*.

The Admiralty knew that the only hope of blocking the enemy's escape lay in a powerful air strike by *Swordfish* torpedo bombers from *Ark Royal*. The carrier, leading Force "H," was now within extreme range. The *Swordfish* took off with instructions to attack a solitary warship "bearing on an azimuthal course of 183 degrees" at a fixed distance in nautical miles from the carrier. These minimal instructions almost caused a tragic and fatal blunder— *H.M.S. Sheffield* was also proceeding alone on a parallel shadowing course to *Bismarck*! The first *Swordfish*, flying in the overcast, picked up a single object radar blip. The plane swooped down in an attack run against the indicated warship. Just as the pilot released his torpedo, he heard the radio warning: "Look out for the *Sheffield*!" The second and third *Swordfish* bombers also released their missiles. Miraculously, two of the three torpedoes, equipped with magnetic exploders, detonated in great geysers of water after running but a short distance. The third torpedo missed its mark. The mistaken target was *Sheffield*!

But the main flight of *Swordfish* found the right target and attacked from every quadrant of the compass. *Bismarck* twisted and turned in a violent evasive action as the helm was swung hard over from port to starboard and back again. One torpedo found its mark near the bow, while another struck the wildly careening stern near the portside propellers. A tremendous vibration racked the entire vessel and she heeled over sharply in a tight turn. The second torpedo had done its job well, hopelessly jamming the rudder hard over, obstructing the port propellers, and forcing the portside engines to be stopped. The damage also flooded the steering telemotor flat and knocked out the steering gear in the compartment.

Bismarck, now thoroughly crippled and steering by her engines, was attempting to proceed northward at a speed reduced to 10 knots. Although at bay, she could still bite; her armament was intact.

The hunters move in for the kill

Rodney and *KG-5*, by now only 50 miles distant, were racing in for the kill. Five destroyers from Force "H" were also within range for the final action. The confident British, no longer transmitting coded messages, were sending their signals in plain English, and the wireless operators aboard *Bismarck* had no difficulty in interpreting the intent of their enemies.

Lutjens expected the Luftwaffe to come to his assistance, and he was very dismayed when a radio message from the Reichsführer exhorted him and his gallant crew "to fight valiantly to the end."

Thus, in the early morning hours of May 27, nine days after she had left Gdynia, *Bismarck* approached her final rendezvous with destiny. *Bismarck*'s radar, before dawn, had detected the stealthy approach of the destroyer flotilla, and when the "tin cans" were within range, the German gunners gave them the full greeting with salvos from both the main battery and 5.9-inch secondaries. The startled destroyer captains, believing themselves safe under cover of darkness, sheered off sharply. The Germans' radar-controlled gunnery had been so accurate that two of the destroyers were struck by shell fragments.

In this action, *Bismarck* utilized her search radar, which was sufficient to give an instantaneous fix on a target, but its overall capabilities were insufficient to provide enough "on-line" data as to the course and speed of enemy warships for an accurate fire-control plot.

Shortly after dawn, *Rodney* came up, followed by *KG-5*. The two battleships deployed to the left and right of the German vessel, thereby effectively splitting *Bismarck*'s firepower so that only two of her four turrets could be brought to bear against either British warship.

The British vessels opened fire at about 20 000 yards distance. The second salvo from *Rodney*'s 16 inchers scored, knocking out both of *Bismarck*'s radar masts and three of her 5.9-inch turrets. The German return fire had hardly commenced when a broadside of 14-inch shells from *KG-5* blasted *Bismarck*'s no. 2, 15-inch turret; one shell striking the barbette while another rolled back the armored glacis plate like the top of an opened sardine tin. Shell fragments and shrapnel swept the bridge. From the later accounts of survivors, it is believed that Lutjens and Lindemann were killed in that inferno. The German giant was being systematically reduced to a shambles. Successive salvos knocked out the no. 1 forward turret, and all electrical communications between central fire control and the gun batteries were obliterated. The remaining turrets operated by local fire control.

High overhead, *Swordfish* observers from *Ark Royal* monitored the fire from the British ships and reported on the damage to the German caused by direct hits. In an hour's time, *Bismarck* was reduced to a battered hulk. *Rodney* and *KG-5* broke off action while the heavy cruisers *Norfolk* and *Dorsetshire* delivered the *coup de grâce* with long-range torpedoes. Just before noon, *Bismarck* rolled over and sank. The scene of the final engagement is pinpointed by the Fig. 3 map.

Radar had played an important role on both sides. The British edge in airborne radar was offset by the German superiority in surface vessel radar gear. But the lesson was well learned by the British and their "boffins" knew that R & D efforts in military electronics would have to be greatly accelerated to close the gap.

The 'Channel dash'—'Scharnhorst' and 'Gneisenau'

If the *Bismarck* affair had a humbling effect on the Royal Navy, a further embarrassment would soon occur in the daring dash by a German armada through the English Channel from the French port of Brest to the North Sea. This run was spearheaded by the twin battle cruisers, *Scharnhorst* and *Gneisenau*, with the cruiser *Prinz Eugen* and a flotilla of destroyers and minesweepers.

Korvettenkapitän Fritz-Otto Busch, a principal chronicler of the war history of *Scharnhorst*; Sir Robert Watson-Watt, one of Britain's foremost boffins of World War II; and Fregattenkapitän Helmuth Giessler, the former first lieutenant and navigating officer of *Scharnhorst* (Fig. 4), are the leading sources for the sequence of startling events surrounding the Channel breakthrough and the last action of the big battle cruiser (Fig. 5) in December 1943.

In an exchange of correspondence with the author, Giessler gives a graphic description of the radar equipment installed aboard both *Scharnhorst* and *Bismarck*, plus some significant historical background:

...The navy was in no particular hurry with the manufacture of radar units because our commander-in-chief, [Admiral] Raeder, had received repeated assurances that no war was to be expected before 1944... Thus we were completely surprised aboard the Scharnhorst *when war broke out in September 1939... During the first days of the war we received a radio telemetering device which was mounted on top of the rotating hood above the optical rangefinder. Like other devices, it had been made by the naval firm GEMA. It operated at the frequency of 368 megacycles (82 cm), had an output first of 14 kW and later of 100 kW. The device was rotatable 360 degrees. The antenna was mounted directly in front of the operator's stand and measured 3 by 6 meters, with six rows of 16 dipoles each... About 1941, in Brest, a second device was mounted on the rear gun turret....*

All German battleships and cruisers were equipped with this type of device. No other type was developed during the war. Only after the discovery of the Allied devices [Rotterdam apparatus] that worked on 9 cm, in February 1943, were these centimeter devices copied...

In my opinion, the reason for omitting the study of the ultrashort waves was that the Germans had been on the offensive until 1942, and thus were able to force the action on the enemy. Therefore, we "did not have to" start an entirely new development at a time when all industry made supreme efforts.... Even in December 1942, the scientists advised... that centimeter waves were useless for reflecting devices because the reflections would be mirror-like!

But now back to the Scharnhorst... *Since I had a hand in the military side from the very beginning, I was able to use the device for guiding the ship tactically, but likewise for artillery purposes. During the period from January to March 1941, when the battleships* Scharnhorst *and* Gneisenau *operated in the Atlantic, we did not observe a single English warship with a radar device...*

The Bismarck *was the first to have a radar search receiver on board which worked in the frequency of the English ASV devices. By means of this device, the two cruisers* Norfolk *and* Suffolk *which kept in contact in the Denmark Strait were discovered...[after the sinking of* Hood*] the* Bismarck *kept hearing the beams of the English devices, and it was assumed that the cruisers were still keeping contact. Actually, contact had been disrupted for some time. However, on* Bismarck, *the technicians did not take into account that the search receiver hears the wave directly, but the echo, too, must return to the sender!*

By February 1, 1942, British naval intelligence reports indicated that the Germans might attempt to move their heavy naval units from the camouflaged confinement of Brest to engage in some marauding on the open seas. The British coastal radar networks had done remarkable work in winning the Battle of Britain, and airborne radar was making life very hazardous for German U-boats. The big coastal batteries at Dover, with radar assistance, were already capable of making the German-held port of Boulogne, at the southern mouth of the Straits of Dover, a risky place for enemy shipping.

On February 3, an elaborate system of three reconnaissance air patrols was established to provide early warning of any enemy activity in the Channel. The patrol lines, named *Stopper*, *Line SE*, and *Habo* (an acronym for "Havre-Boulogne"), covered the bottleneck of the Brest harbor roadstead, the airspace between Cape Ushant and the Channel islands of Jersey and Guernsey, and the surveillance of the two principal ports of Havre and Boulogne, respectively. All of the patrol aircraft—Lockheed *Hudsons*—were equipped with ASV Mark II search radar. These patrols flew routine, uneventful missions for the following week.

Then, on the evening of February 11, the first plane on the Stopper run, picked up a German *Junkers 88* on its radar. The *Hudson* switched off its radar and took evasive action. Later, the RAF crew switched on the radar and found it unserviceable. The plane returned to its base and its crew took off in another *Hudson*, arriving at the Stopper patrol line three hours late.

Meanwhile, the Line SE patrol, on a synchronized schedule, was experiencing the same difficulties from unresponsive airborne radar. However, this aircraft did not immediately return to its base. When the failure was finally reported, the balance of the patrol mission was "scrubbed."

Habo, the backstop patrol run, apparently experienced no difficulty with its ASV equipment. In retrospect, the logical thing to have done was to fill in the Stopper and Line SE gaps with reserve aircraft from Habo, but human beings, as well as hardware, are subject to failure.

Circumstantial evidence, gathered in the postwar period, seems to indicate that the Germans were effectively jamming the British airborne radar. In fact, Winston Churchill, in his later account of the total situation, wrote: "...when the enemy ships were spotted, the radar of our patrolling craft broke down. Our shore radar also failed to detect them. At the time we thought this an unlucky accident. We have learned since the war that General Martini, the Chief of the German radar, had made a careful plan...the jamming had grown so strong that our sea watching radar was in fact useless...To allay complaints an official inquiry was held which reported the publishable facts."

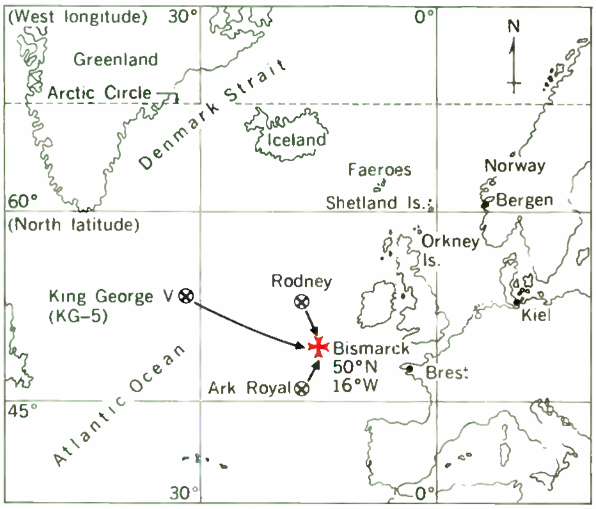


Fig. 3. Closing for the kill. On the morning of May 27, the big guns of the battleships *Rodney* and *KG-5*, and planes from the carrier *Ark Royal*, sank *Bismarck* at the coordinate bearings indicated at the Maltese cross.

Fig. 4 (below). Captain Helmuth Giessler (right), navigating officer of *Scharnhorst* until his transfer in 1943, is shown on the bridge of the battle cruiser with the commanding officer, Captain Hoffmann. Giessler, now retired, furnished the author with valuable technical information.

Fig. 5 (bottom). Overall view of *Scharnhorst*. Note the camouflage pattern intended to break up the profile into the illusion of a much smaller vessel when seen at long distance.

With this background of German boldness and technical ingenuity, we now come to the historical events shown in Fig. 6.

Late in 1941, *Scharnhorst* and *Gneisenau* put into Brest for refitting and repairs. The British were confident that the RAF would seek them out and destroy them.[2] But the two battle cruisers were too well camouflaged to suffer more than minor damage from British aerial bombs. Nevertheless, they were in a hot spot, and, by the end of January 1942, secret orders from Berlin directed Vizeadmiral Ciliax to remove all naval units under his command to the German base at Wilhelmshaven.

Steam-up was ordered on the night of February 11. At about 9:30 P.M., the routine reconnaissance RAF aircraft (Stopper) appeared over the port and dropped flares. Both the port and the ships had to be quickly concealed by a smokescreen. Shortly thereafter, the German naval force was under way to begin the great adventure. Within an hour the fleet was steaming up the Channel at 27 knots. Weather conditions were ideal; the night was exceptionally dark and moonless.

By 8:00 A.M., the German vanguard was about 55 nautical miles due west of Dieppe. During the change of course up the dog leg (from 9:14 to 10:23 A.M.) the advance minesweepers discovered a newly laid minefield, but the warships eased through it at reduced speed.

Meanwhile, the British shore-based radar stations continued to plot surface vessel movements with various degrees of impairment caused by jamming throughout the morning hours.

British defenses come to life

At about 10:00 A.M., the German night fighters, which had been flying very low to avoid radar detection, were replaced by *Messerschmidt 109* day fighters. As the German warship squadron approached the Straits, a single RAF plane appeared. German AA direction shots marked it by flak bursts and the Luftwaffe planes sent the interloper crashing into the Channel.

By noon, the Germans had passed the narrowest part of the Straits between Dover and Cap Gris Nez. According to Busch, the Germans were amazed. There was still no sign of determined British countermeasures.

But as the German force approached abreast of the Thames estuary, the British defenses sprang into action. A squadron of *Swordfish* torpedo bombers and a flight of *Beaufighters* attacked the German ships. All of these craft were either shot down or driven off by the Luftwaffe and antiaircraft guns.

Meanwhile, the minesweepers were still at work. At about 3:30 P.M., a heavy explosion shook *Scharnhorst* from stem to stern. She had struck a mine and there was flooding in the machinery spaces, which required the engines to be stopped for repairs. *Gneisenau* and *Prinz Eugen* proceeded on their course while *Scharnhorst* lay dead in the water for more than half an hour. Fortunately for the German vessel, the weather closed in and there was a light drizzle. No British aircraft were sighted during this interval. But damage control reported that the electronic echo-sounder and radio direction-finder had been knocked out of action by the mine explosion.

By 4:04 P.M., engine repairs were completed and *Scharnhorst*, screened by flanking E-boats, again sped up the Channel at 27 knots. *Gneisenau*, *Prinz Eugen*, and the other German warships, now far ahead of the damaged

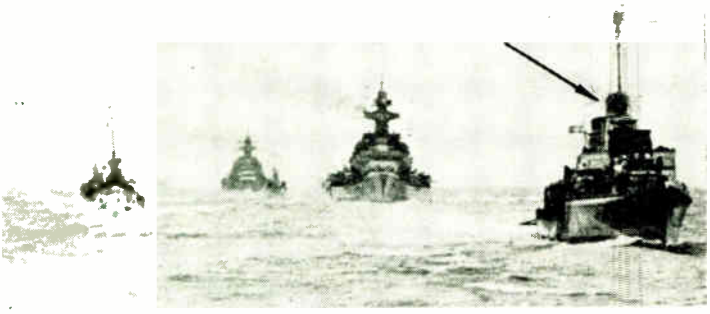


Fig. 6. Battle cruiser *Gneisenau* (second from right) leading its sister ship, *Scharnhorst*, during the Channel breakthrough of February 12. German destroyers formed the flanking screen to protect the battle cruisers from attack during the high-speed run. Arrow indicates the radar antenna grid on the foremast of the lead destroyer.

flagship, were encountering hostile action from a British cruiser and destroyers. Also, several groups of RAF bombers were thrown into the action. Wave after wave of these attacking aircraft were unsuccessful in their efforts to sink the heavy German ships. *Scharnhorst* now had to contend with a serious navigation problem. Since the radio direction-finder and echo-sounding gear were out of service, the ship had to proceed with utmost caution to avoid the treacherous sandbanks off the Dutch coast to starboard and the mine fields on the port side of the ship channel.

Mission accomplished. At slow speed, on the morning of February 13, *Scharnhorst* was escorted into Wilhelmshaven harbor. The other vessels of the naval force had already arrived. The impossible had occurred. *The Times* of London admitted that the Royal Navy had suffered its most humiliating experience in 300 years.

Airborne radar—the Luftwaffe has its problems

Germany had lost the Battle of Britain because of the efficiency of the British radar network. In Germany, each branch of the armed services—army, navy, and air force—had worked independently on radar development, with no obligation to assist each other or to exchange information.

The German scientists and electronics experts were competent enough, but the Wehrmacht high command neglected to exploit some excellent R & D discoveries. And many promising devices and techniques, which could not be put to immediate operational use, were rejected by the military as "technically immature."

The British radar apparatus—1943

In January 1943, a British heavy bomber was shot down near Rotterdam, Holland. Some strange equipment, bearing a chalk inscription, "Experimental 6," was found in the smashed aircraft. The device was sent to the Telefunken Company in Berlin, where scientists determined that it functioned on a 9-cm wavelength. Until this time, German researchers had been convinced that such a wavelength was unsuitable for radar. The Telefunken people called the salvaged device the "Rotterdam apparatus."

The Germans attempted to reconstruct Experimental 6, but, unfortunately, a number of essential parts were missing from the damaged prototype. It was not until August 1943, when another British plane carrying the same device was shot down, that the equipment could be entirely rebuilt.

The reconstructed radar gear was finally installed on the top platform of a flak tower, and the sight that the scientists beheld provided one of the major surprises of World War II. Etched on the screen (the "yellow-green eye"), in fluorescent light, was a complete picture of the city of Berlin. Beyond the city's perimeter, woods, lakes, canals, and houses could be seen within a radius of 40 miles. The startled German scientists realized that darkness and weather had been conquered, and that the RAF and the Royal Navy were capable of observing every target that was optically invisible.

Coming events cast their shadows

The efficacy of the new British radar device, in its naval applications, soon confirmed the Germans' worst fears. In September 1943, two German destroyers, *Hermann Schoemann* and *Z26*, were lost in the Arctic Ocean under strange circumstances. These vessels had participated in numerous actions against Allied convoys, and their crews were accustomed to operations under adverse weather conditions. The destroyers were suddenly confronted by an unwelcome intruder—a British heavy cruiser was spotted emerging from the icy mists as a vague silhouette. The two German ships were immediately straddled by a salvo of 8-inch shells. And the second broadside finished the business.

It was inconceivable that the brief visual contact could have been sufficient to give the British fire control enough data for such accurate shelling by optical rangefinders. Actually, the yellow-green eye was doing its work and the cruiser's fire control had "zeroed in" on the enemy targets minutes before the visual sighting.

'Scharnhorst's' last cruise—December 1943

Background to disaster. To obtain accurate and authentic technical information and historical background on the details of *Scharnhorst*'s final sortie, the author also contacted the Militärgeschichtliches Forschungsamt (Military History Research Institute) in Freiburg, Germany. The following significant passage from the German reply may be of interest:

...The radar equipment of the Scharnhorst *was probably technically inferior to that on the opposite side. Since the pursuit of the* Bismarck, *it had become clear that Great Britain had overtaken the previous German advance and had herself succeeded in obtaining a lead....*

By December 1943, the German military situation on the Eastern Front had deteriorated and a powerful Russian offensive was pushing back the Wehrmacht.

At Stalin's insistence, Prime Minister Churchill agreed to send four convoys, of about 35 ships each, to the Russian ports of Murmansk and Archangel during the winter of 1943–1944. Previous convoys had suffered heavy losses from German naval forays and Luftwaffe strikes. The arctic routes to Russia were the most hazardous and difficult sea lanes in World War II.

Persistent Allied air and naval action had reduced the striking power of both the Luftwaffe and U-boats. A successful attack by British midget submarines had crippled the giant battleship *Tirpitz* at her anchorage in Alta Fjord, Norway. The battle cruiser *Gneisenau* had been severely damaged by mines and aerial bombardment and was out of commission at Gotenhaven. The other major surface units of the Kriegsmarine—*Prinz Eugen*, *Hipper*, *Lützow*, and *Admiral Scheer*—were required for service in the Baltic Sea.

The German High Command advised Oberkommando Gruppe/Nord (Naval Command Group—North) that pressure on the Russian front had to be quickly relieved. *Scharnhorst* was the only heavy naval unit of the First Battle Group available for action against the resumed Russian convoys. It was a desperate move, for German intelligence reports indicated that the British warships were equipped with radar fire control. Since the German naval ordnance could direct fire electronically only to a very limited extent, a night encounter would render the German units practically blind against the enemy's all-seeing yellow-green eye.

Grand Admiral Dönitz knew that an Allied northbound convoy (designated JW.55B) was due to sail late in December from Loch Ewe, Scotland, to Murmansk. What he did *not* know was that: an unladen convoy (RA.55A) would leave Murmansk a few days later; Admiral Sir Bruce Fraser, in the battleship *Duke of York*, accompanied by the cruiser *Jamaica* and the destroyers *Savage*, *Saumarez*, *Scorpion*, and *Stord*, would cover JW.55B; Vice Admiral R. L. Burnett, in the light cruiser *Belfast*, with the cruisers *Norfolk* and *Sheffield* and the destroyers *Musketeer*, *Matchless*, *Opportune*, and *Virago*, would protect RA.55A.

The Royal Navy had anticipated Dönitz' objectives, and Admirals Fraser and Burnett scheduled a series of complex tactical maneuvers to intercept any German raiders before they could attack the two convoys in the vicinity of Bear Island.

Vizeadmiral Oskar Kummetz (chief of Oberkommando Gruppe/Nord) did not like the idea of risking *Scharnhorst*, but Dönitz, determined to convince Hitler that capital ships were still a decisive factor in naval warfare, ordered her to depart with a task force group.

Naval log: 23–25 December 1943. In this description of the grim battle fought far north of the Arctic Circle, the writer will attempt to intersperse chronologically the historical accounts from Royal Navy sources with excerpts from Kapitän Giessler's letter (in italics). There are discrepancies in the two chronicles in reference to the actual sinking of the German battle cruiser.

For the mission which was ordered shortly before Christmas, no special practice was possible. . .the commander [Konteradmiral Bey], together with a small staff, came aboard very shortly before sailing time. The commander had not yet been in charge of a group of such units during winter in the North Sea before.

I do not believe that the First Battle Group had received any message warning them that the British ships would probably use the new centimeter devices. The Scharnhorst *at that time did not yet have a search radar receiver for these wavelengths aboard. . .*

A Luftwaffe pilot spotted convoy JW.55B just north of

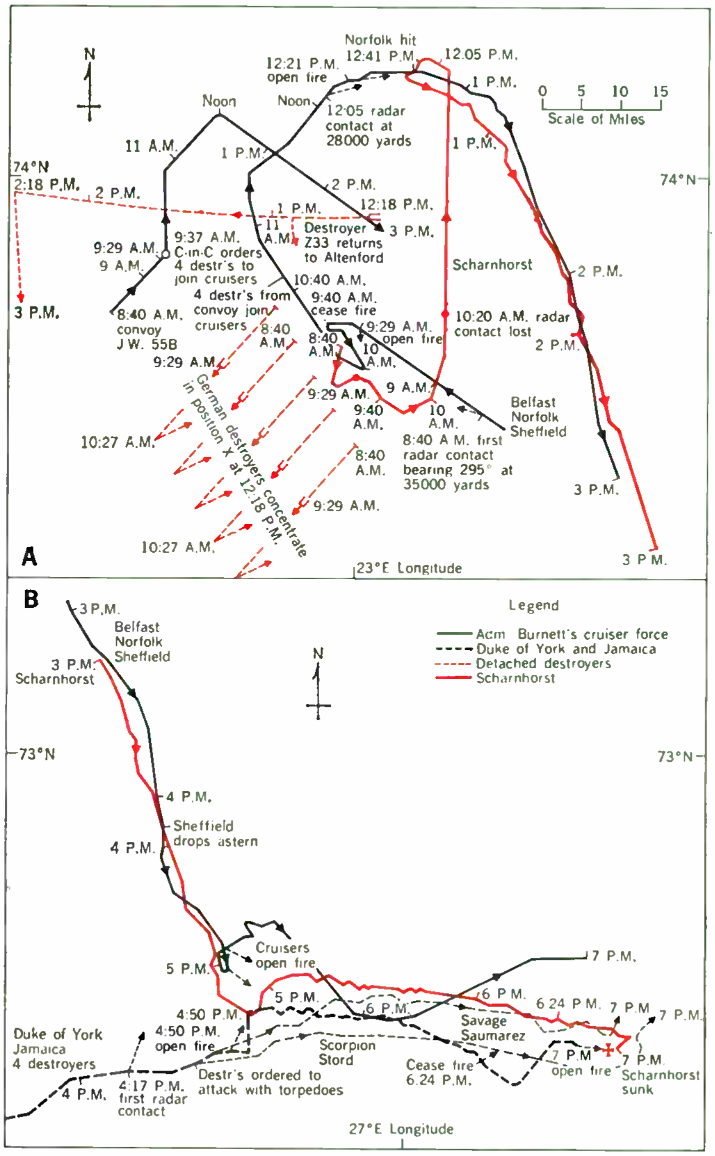


Fig. 7. A—Sinking of *Scharnhorst*, phase I. Chart shows tactical maneuvers of British and German warships in the waters off Spitzbergen between 8:40 A.M. and 3:00 P.M., December 26, 1943. B—Sinking of *Scharnhorst*, phase II. Chart indicates the action between 3:00 P.M. and 7:00 P.M., during which time the heavy guns of battleship *Duke of York* ensured the ultimate British victory.

the Faeroes, on the morning of December 23, and he delivered the first attack. And, before noon on that day, five German dive bombers pressed home an attack against the 19-ship convoy. Despite the limited visibility of the arctic twilight, the ships' gunners managed to down two of the enemy aircraft and drove off the rest.

Meanwhile, convoy RA.55A, consisting of 22 ships, remained undetected by the Germans. Admiral Fraser, however, knew that JW.55B was under surveillance and attack by the Luftwaffe, and he was sure that this convoy would be the prime German objective. Therefore, he ordered RA.55A to pass through an ice-choked channel north of Bear Island and to deploy four of its destroyers to augment the screening force of JW.55B.

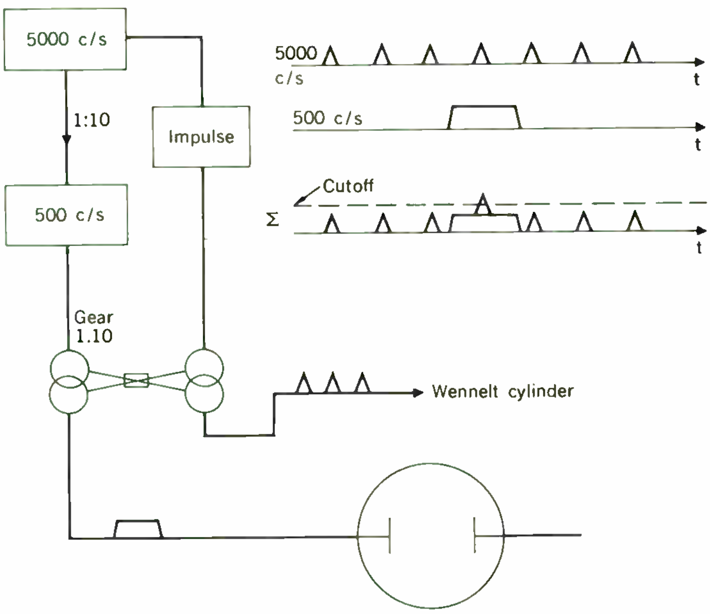


Fig. 8. Diagram of "Emilprinzip" radio detection device. The theoretical maximum range that could be indicated with a radio frequency of 500 c/s, with an accuracy of ±35 meters, was 300 km. The basis of this device was a 5-kc/s oscillator that produced the impulse. By frequency subdivisions of 1:10, 500 c/s of the total quantity were led off and also transformed into impulse. Both impulses were added and applied to the control grid (Wennelt cylinder) of the cathode-ray tube. The functioning point was so regulated that only the peak of the 5-kc/s impulse, produced by the 500-c/s impulse, caused the CRT to illuminate. Both impulses operated the phase shifters, coupled to each other over a 1:10 gear apparatus.

At 2:00 P.M. on Christmas Day, Oberkommando Gruppe/Nord ordered the operation to proceed. Konteradmiral Bey—and probably Kapitän Hintze of the *Scharnhorst*—wanted to call off the operation, but Dönitz insisted upon the sortie.

As ordered, the Scharnhorst *left port in the evening [December 25] in extremely bad weather. The five ["Narvik" class] destroyers followed, but they had only limited use of their armament because of the waves.*

The British Admiral Fraser received a report from an intelligence *agent very soon after the departure of the SH* (Scharnhorst), *so he was able to take immediate action; swing the convoy JW.55B to the north, employ the cruisers covering this convoy to attack the SH, and have his heavy battle forces close in fast. . .*

26 December—Phase I. During the night of December 25–26, Fraser's squadron steamed eastward at 17 knots through heavy seas. British intelligence had worked efficiently, for at 3:39 A.M. on the 26th, Fraser received a signal from the Admiralty that "*Scharnhorst* was probably at sea."

By 4:00 A.M., the northbound convoy JW.55B was about 50 nautical miles south of Bear Island. *Scharnhorst* and her destroyer screen were proceeding at flank speed northward on an intercept course with the convoy. Admiral Burnett's cruiser squadron was about 150 miles to the east of JW.55B and steaming at 18 knots on an azimuthal course of 236 degrees (approximately SW by W). Fraser, in *Duke of York*, was still being buffeted by gale-force winds 350 miles to the southwest. He was anxious to prevent the German task force from attacking the convoy during the short period of arctic daylight, and, therefore, broke radio silence to obtain an exact fix both on the convoy's and Burnett's position.

Despite the heavy weather, Fraser ordered propeller turns for 24 knots; Burnett also went to flank speed.

At 8:40 A.M., the search radar aboard Admiral Burnett's flagship, *Belfast*, made contact with the enemy battle group at a range of 35 000 yards on a bearing of 295 degrees relative. *Scharnhorst*, now unescorted, was about 30 miles south of the convoy; Bey had deployed five destroyers [Fig. 7(A)] to locate JW.55B and guide the big battle cruiser in for the kill. The destroyers, with inadequate radar equipment, steamed to within 18 000 yards of the convoy without detecting it in the arctic gloom. An hour later, Bey ordered the destroyers to reverse course and head back to Norway.

At 9:20 A.M., *Sheffield* made visual contact with *Scharnhorst* at a range of 13 000 yards. The cruiser *Norfolk* opened fire with her 8-inch gun batteries.

Around 10:00 o'clock, the SH was shot at from the dark with star shells and 20.3-cm (8-inch) shells. Our own radar was able to locate the cruisers Sheffield *and* Norfolk *only at a distance of 11 kilometers (about 12 000 yards), while the British devices located the SH at a distance of 33 kilometers (about 36 300 yards). According to . . . survivors, during this short battle, the antenna of the forward radar was smashed, so that thereafter only the rear radar at the lower height could be used. . .*

Norfolk scored two direct hits with 8-inch shells on the enemy's superstructure. *Scharnhorst* and the two other cruisers, *Belfast* and *Sheffield*, remained silent in this phase of the action.

Scharnhorst now attempted to open the range, by veering to the northeast, in order to give her 11-inch guns the trajectory advantage. (As the flight path of a heavy shell describes a parabola, opening the range presents a greater danger to an enemy because the projectile will be dropping almost vertically, thereby increasing the chance for an armor-piercing shell to crash through the armor deck of the target and to explode deep below decks where maximum damage can be done.)

Admiral Burnett suddenly guessed the German's intention of eluding him to attack the convoy. Here, the British cruisers were at a decided disadvantage; they could not make more than 24 knots in the heavy seas, while *Scharnhorst* was capable of a speed of almost 32 knots. To offset this setback, Burnett ordered the fast destroyers, *H.M.S. Musketeer*, *Matchless*, *Opportune*, and *Virago*, already detached from the convoy, to serve as a forward screen to attack or monitor *Scharnhorst*.

Although the SH was now "almost blind," the German admiral tried once more to get at the convoy. During her course to the north, the SH was located immediately by the radar of the cruiser [Sheffield] *and was not released again [from radar detection]. The SH did not realize that it was under constant electronic observation! Around noon, another short exchange of gunfire ensued without* [Scharnhorst's] *sighting or locating the enemy. . .*

There was a brisk 20-minute artillery duel between *Scharnhorst* and Burnett's cruisers. At ranges that varied from 10 000 to 16 000 yards, *Norfolk* was struck twice by 11-inch shells and one of her main turrets was knocked out. The four destroyers were unable to approach within torpedo range and the chase continued.

After this encounter, the SH broke off the action and tried to steam to the southeast at flank speed. During that time also, the SH was under constant radar observation so that Admiral Fraser always knew her location and could act accordingly. In the afternoon, the SH was located by the Duke of York*'s radar at a distance of 42 kilometers! This sealed the fate of the ship. . .*

26 December—Phase II. *Belfast* continued shadowing the German by search radar from a distance of about 13 500 yards as the short arctic day turned to impenetrable night. *Sheffield* was forced to drop astern because of propeller shaft trouble, but *Norfolk* and *Belfast* kept the radar fix. This information was relayed to Fraser, who then easily plotted an intercept course.

At 4:17 P.M. (see illustration on page 63), *Duke of York* made its first contact by search radar at a range of 45 000 yards, at which time the Royal Navy battleship was steaming on an azimuthal course of 90 degrees. *Scharnhorst* was bearing on a course of about 157 degrees true.

Duke of York's fire-control radar began to accumulate data at a range of 29 700 yards—a fact that the German navy later considered to be incredible!

Suddenly, at 4:50 P.M., a parachute flare from *Duke of York* revealed *Scharnhorst* in a stark and glaring outline. Bey was aghast; he had no inkling that the British battleship had been in pursuit. Too late! Before he could bring his guns to bear, the low-pitched howl of 14-inch shells from *Duke of York* [Fig. 7(B)] straddled the German and one hit was scored at a range of 12 000 yards.

Scharnhorst reacted swiftly. Veering in a tight circle, she reversed course to the east and steamed off at such high speed that it appeared for a time that she would escape. Burnett's cruisers attempted to intercept and fired a few ineffectual salvos. Intermittently, *Scharnhorst* would change course, fire a broadside at the cruisers, and then resume her former course at full speed. This made things very difficult for even the British radar fire control. *Duke of York* scored three more hits during this running action and was, herself, struck in the foremast by one of *Scharnhorst*'s 11-inch shells (illustration on page 62). Fraser's fire-control radar was knocked out temporarily by this lucky hit, and the British flagship was straddled by near misses many times.

The hot pursuit continued in pitch darkness, illuminated only by the flashes of cannon fire and the phosphorescent glow of exploding shells each time *Scharnhorst* was struck. A white squall descended, and the strange, unearthly drama continued, with the yellow-green eye of the fire-control radar relentlessly seeking out its prey.

Shortly after 5:00 P.M., *Scharnhorst*'s main batteries were put out of action, but her secondary 5.9-inch rifles continued to fire blindly at the four "S"-class destroyers [Fig. 7(B)] that were knifing in to a range of 4000 yards to launch a torpedo attack. At 6:00 P.M., Bey radioed to his superiors: "We will fight to the last shell. . ."

Duke of York, which had been firing at rate of one salvo a minute, stopped shooting to permit *Jamaica* and the destroyers to deliver the *coup de grâce*. In all, 55 torpedoes were fired, the last three of which did the most telling damage. *Scharnhorst*, by 6:45 P.M., was dead in the water. Only a dull glow from the many fires raging on board could be seen. At about 7:00 P.M., there was a terrific explosion. That was the end of the battle cruiser.

Of a complement of almost 2000 men, *Belfast* was able to rescue only 36 survivors. None of the ship's officers survived and the highest rating taken aboard the British cruiser was a petty officer.

Epilogue

In spite of all the hits by heavy artillery and torpedoes (about 18 torpedo hits alone) the ship was still going forward. When she capsized, the propellers were still turning and the electric lights were on!

This battle probably was the last classic action between battleships. Technology has made the sea battle of old, traditional from the age of sailboats, absurd. However, it took a very long time until the top leaders of all navies learned that by the invention and rapid development of radar, the basis of naval warfare had radically changed. The night and the fog no longer gave any protection, and the danger of discovery persisted. . .

(Signed) H. Giessler
Kapitän zur See

Appendix. German radar gear

The author is indebted to Dr. W. L. Rubin of the Sperry Gyroscope Company for the following interpretive and descriptive analysis of the various electronic equipment, systems, and techniques used by the German military establishment during World War II.

From all indications, the Germans, in their selection of wavelength and polarization, were well aware of the important radiation and propagation principles that enter into radar system design.

Emilprinzip. To obtain a more accurate range information capability than that which could be read off a cathode-ray-tube display, the Germans employed calibrated range-gating (Fig. 8) circuits. This permitted more accurate ranging by displaying the target only when it fell within the range gate. The development was quite significant in military range-finding applications.

Improving the minimum range. The German military literature refers to a technique for cutting off early receiver amplifying tubes to minimize receiver overload, and for "blocking" when the transmitter fires. In this way, reception is improved immediately following the radar transmission. Although fairly obvious, the idea is good. There is no description available, however, as to how the transmitter and receiver were connected to the radar antenna, but one may assume that the Germans were knowledgeable in this area.

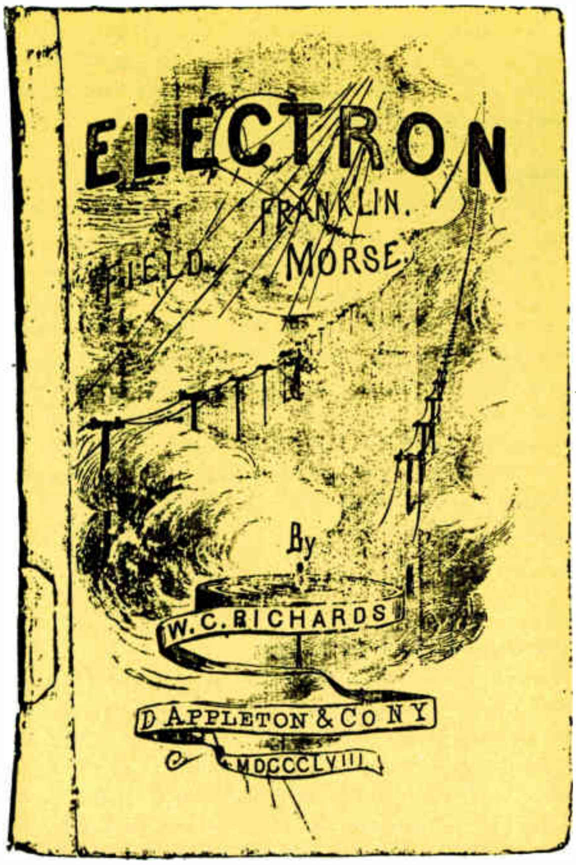
Pulse phase shifter. The title "pulse phase shifter" is misleading since the Germans utilized the characteristic of a charging *R-C* network to obtain accurate time delays. This was an important development since it probably led to greater accuracy in ranging and tracking, as well as to single-pulse and intermittent-pulse operation.

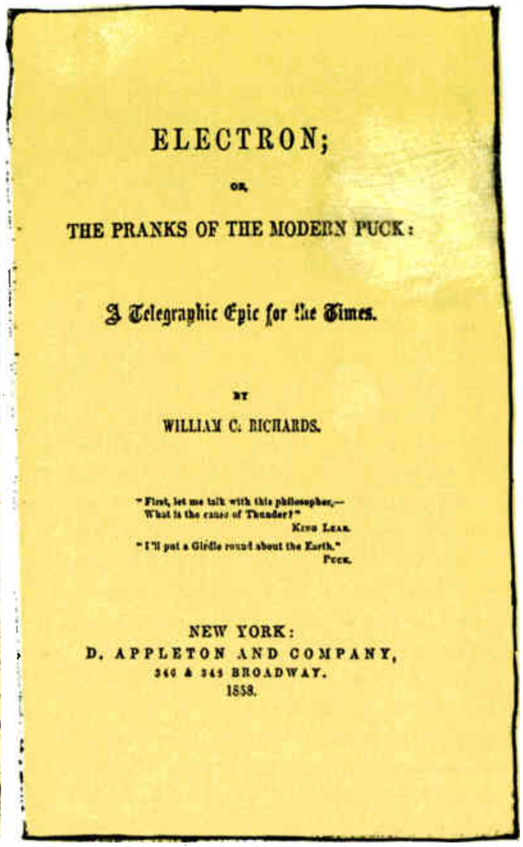
The author wishes to acknowledge the assistance and cooperation of Rear Admiral E. M. Eller, USN (Ret.), Curator for the Navy Department and Captain F. Kent Loomis, USN (Ret.), Assistant Director of Naval History.

REFERENCES

1. Ruge, Friederich, *Der Seekreig* (The German Navy's Story, 1939–1945). Annapolis: United States Naval Inst., 1950, pp. 173–174.

2. Busch, Fritz-Otto, *Holocaust at Sea*. New York: Holt, Rinehart and Winston, Inc., 1956, pp. 36–41.




ELECTRON;

OR,

THE PRANKS OF THE MODERN PUCK:

A Telegraphic Epic for the Times.

BY

WILLIAM C. RICHARDS.

"First, let me talk with this philosopher,—
What is the cause of Thunder?"
KING LEAR.

"I'll put a Girdle round about the Earth."
PUCK.

NEW YORK:
D. APPLETON AND COMPANY,
346 & 348 BROADWAY.
1858.

The origin of the term 'electronics'

As the technology has evolved, so the definition of 'electronics' has evolved—from a narrow physics application to the overall concept of electron devices and their utilization

Charles Süsskind *University of California, Berkeley*

A historian of the technology reviews the beginnings and changing meanings of the term "electronics." He points out that it goes back to the adjective "electronic," and even further to the name "electron," which was first proposed in 1891. In its modern context, the origin of the word "electronics" has been ascribed to the first appearance of the magazine of that name, in 1930, although there is some contention that it was already in use in Great Britain. In any event, its definition has broadened through the years and the future may see even further changes.

When the AIEE and the IRE were planning their merger into the IEEE, there was a great deal of internal controversy as to whether that second "E" was to stand for "electronic" or "electronics." Some of that dissent was evidenced by the response to a 1961 IRE PROCEEDINGS editorial on the subject.[1] (It is interesting to note that *The Encyclopedia of Electronics*,[2] published in 1962, does not include a definition of the word itself, as was pointed out to me as its editor by a German colleague.[3]) In the case of the IEEE, good grammatical sense finally prevailed[4] and "electronics" was recognized as the correct modifier for "engineers," although not all the world agrees. We still have an Electronic Industries Association and a Western Electronic Manufacturers Association, and the British IRE recently became the Institution of Electronic and Radio Engineers. (On the other hand, Australia's Institution of Radio and Electronics Engineers uses the final "s.")

The IRE editorial precipitated several comments,[5-7] including my own note on the origin of the term "radio."[8] Another letter[9] also went over much the same ground as mine did, and subsequent correspondence with its author, Helmut Drubba of Hannover, revealed his own interest in the origin of the term "electronics." Some of the material cited in the following was first brought to my attention by Herr Drubba.

'Electronics' and 'electronic' in science

"Electronics" (Ger. *Elektronik*) originally designated the branch of physics concerned with the particulate

(Continued on page 77)

properties of materials and the behavior of electrons—roughly what we today might call "physical electronics." The usage still survives among scientists; for instance, Britain's Institute of Physics and Physical Society has an Electronics Section concerned with physical electronics.

In that context, the word occurred in the titles of two periodic publications dating back to the first decade of the century: *Jahrbuch der Radioaktivität und Elektronik*, which appeared between 1904 and 1924, when it merged into the *Physikalische Zeitschrift*; and *Ion: A Journal of Electronics, Atomistics, Ionology, Radioactivity and Raumchemistry*, which first appeared in London in 1908 and lasted just two years, in the second of which it switched to German (*Ion: Zeitschrift für Elektronik, Atomistik, Ionologie, Radioaktivität und Raumchemie*). The name of the former publication, a respected hardcover annual, represents possibly the earliest use of the noun, as has been pointed out by Espenschied[5] and by Drubba.[10]

The scientific adjective "electronic" is even older. *The Oxford English Dictionary*[11] traces its earliest use to Fleming,[12] who used it in the title of a popular article on "the electronic theory of electricity" in 1902.[12] Still older, of course, is the name "electron" for an elementary unit of electric charge, which was first proposed in 1891, six years before Joseph John Thomson (1856–1940) actually confirmed the existence of a subatomic particle.

Introduction of the concept of displacement into the theory of electromagnetism by James Clerk Maxwell (1831–1891) led to the viewpoint, held by most of his followers, of electricity as a continuous medium. The opposite view, first promulgated by Wilhelm Eduard Weber (1804–1891), that electrodynamic phenomena could be explained in terms of discrete charges, was revived through the efforts of two men. One of them was George Johnstone Stoney (1826–1911), who read a paper on the physical units of nature at the British Association meeting in Belfast in 1874. From the laws of electrolysis established by Michael Faraday (1791–1867), Stoney postulated a "single definite quantity of electricity" that he called the "electromagnetic electrine." The paper did not appear in print until 1881.[13] The other man responsible for reviving the atomistic hypothesis was Hermann Ludwig Ferdinand von Helmholtz (1821–1894), who, in a lecture before the Chemical Society in London, also in 1881,[14] pointed out that there was no conflict between the two views, since Maxwell's theory and the theory of discrete charges were both based on Faraday's work.

In 1891, "electrine" gave way to "electron" in a paper published by Stoney in Dublin that must be regarded as the first use of the term in its modern sense.[15] Three years later, stung by a German professor's passing reference that ascribed priority in reviving the corpuscular view of electricity to Helmholtz, Stoney pointed out somewhat irascibly that his two announcements preceded Helmholtz' note in the *Philosophical Magazine*[16] that is cited by some authors[17] as the source of the term "electron." However, the Irish reference is earlier.

After J. J. Thomson's experiments in 1897 verified the existence of a corpuscle of negative charge, the name that had been proposed by Stoney earlier was generally adopted. It is possible to speculate whether the acceptance of the term "electron" was so general because the word was, so to speak, "in the air" even before Stoney proposed it. It is, of course, descended from the Greek ἤλεκτρον, or amber, whose electrostatic properties were known to the ancients. As "electric," it appears in 1600 in the *De Magnete* of William Gilbert (1540–1603). In English, the substantive "electricity" was first used in 1646 in *Pseudoxia Epidemica* by Sir Thomas Browne (1605–1682); and it appears in the title of *Mechanical Origine or Production of Electricity* published by Robert Boyle (1627–1691) in 1675. Finally, Lloyd Espenschied has turned up one instance each of the use of "electron" and "electronic" well before Stoney, though one is fictional and the other is doubtless spurious.[18]

The first instance concerns a book of poems published in New York by William Carey Richards (1818–1892) in 1858—*Electron; or, the Pranks of the Modern Puck: A Telegraphic Epic for the Times*. In this work, which was occasioned by the laying of the transatlantic cable, the history of electricity is represented in verse as the adventures of the sprite Electron (see title illustration). The second case also accompanied an interesting episode in the history of electricity. After Faraday discovered electromagnetic induction in 1831, he communicated the results to the Royal Society (on November 24); the paper later became the first in the series collected as the three-volume *Experimental Researches in Electricity*. He also sent a hasty résumé to Jean Nicolas Pierre Hachette (1769–1834), which was communicated to the Académie des Sciences and appeared in *Le Temps* before publication of the complete Royal Society paper. Two Italian scientists, Leopoldo Nobili (1784–1835) and Vincenzio Antinori (1792–1865), acting on the scanty information contained in the résumé, quickly performed some experiments that they claimed "verified, extended, and, perhaps, rectified" Faraday's results. Within a month, they had prepared a paper for *Antologia*, a journal of sciences, letters, and arts published in Florence that was forever in trouble with the censor. (In fact, it was suppressed by the Tuscan government later that year.) The resulting publication delays led to the article's appearance in a belated issue dated November 1831, which was nominally earlier than Faraday's first publication. Continental readers were thus led to believe that the Italian work had anticipated that of Faraday, who took an early opportunity to disabuse them.

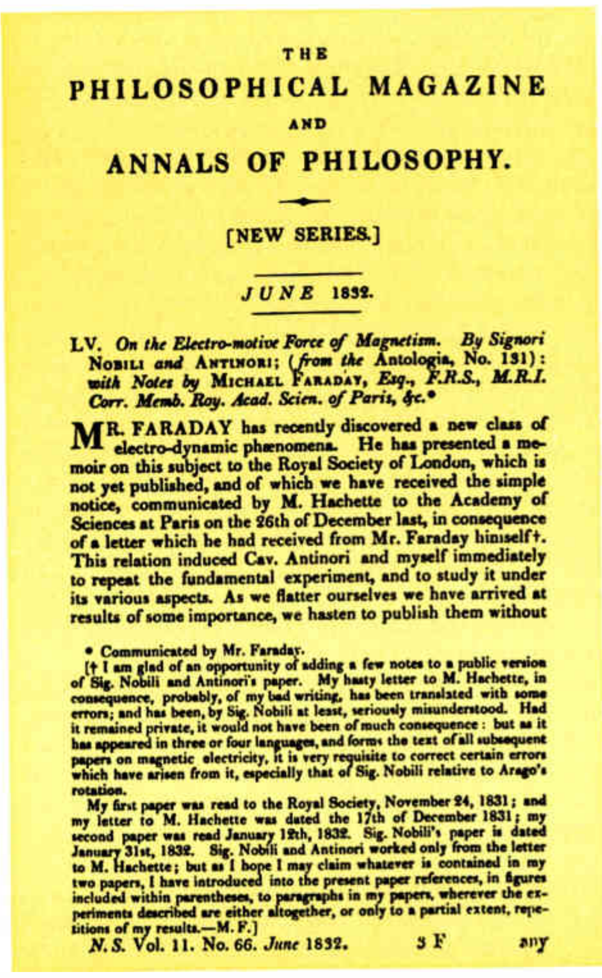
Faraday arranged to have the *Antologia* paper published in the *Philosophical Magazine* for June 1832 in English translation, to which he appended his own notes.[19] In one place, perhaps owing to his erratic handwriting in his "hasty letter to M. Hachette," there appears the phrase, *electromo state*. Here Faraday inserts a footnote, "This should be *electronic state*"—and that is very likely *the* first use of the adjective "electronic." The only trouble is that it is almost certainly spurious. Faraday doubtless meant "electrotonic" to begin with. In another résumé of this work, written two weeks earlier to a friend, the chemist Richard Phillips (1778–1851), Faraday coined the term *electrotonic*, adding exuberantly, "Am I not a bold man, ignorant as I am, to coin words. . . ?"[20] We must therefore conclude that Faraday wrote "electronic" to Hachette by mistake, that it was misread as "electromo," and that he compounded the error by correcting it back to "electronic" instead of "electrotonic."

'Electronics' as a branch of technology

In the more prevalent current context, "electronics" (Ger. *technische Elektronik*) means a branch of technology; the origin of the word has been ascribed to the first appearance, in 1930, of the McGraw-Hill magazine of that name.[21]

The question arises as to whether this usage of the word might not predate 1930; certainly the use of the adjective "electronic" in the same context does. For instance, in his two-volume history of the radio industry, Archer describes E. F. W. Alexanderson's attempts to develop telephone repeaters by employing (1) magnetic microphones before 1912 and (2) the De Forest triode in 1912–1913. Archer mentions that after the latter was perfected for such uses in the laboratories of the General Electric Company, it resulted in what Alexanderson later "called an electronic amplifier."[22] A more reliably dated use of the adjective exists in a 1919 paper on telephone repeaters by Gherardi and Jewett.[23] They quite clearly classified repeater elements as electrodynamic, electronic, and gaseous. They explained that by "electronic" they meant "thermionic," i.e., circuits employing high-vacuum De Forest triodes.

A couple of years later, we find the term *technische Elektronik* in the titles of two articles by Meyer;[24,25] in being translated into English for *Science Abstracts*, it was inverted to "electronic technology."[26] The terms "electronic devices," "electronic methods," and "electronic control" also occur.[27] It would thus appear that "electronic" was in fairly general technological use in the early 1920s. However, the only evidence that "electronics" was "in the air" before 1930 is indirect. Reminiscing on the occasion of the name change of the British IRE to IERE, the Institution's general secretary, G. D. Clifford, remarked editorially that "when the Institution was formed on 31st October 1925 there was considerable debate as to the most apt title and many founder members favoured inclusion of the term 'electronics.'"[28] Proof of this statement could be presumably found in the minutes of the founding committee, if they exist. As a matter of fact, the first name of the institution was the Institute of Wireless Technology. Its members were evidently preoccupied with semantics from the beginning: a rather testy article on nomenclature appears in one of the early issues of their journal.[29]



THE

PHILOSOPHICAL MAGAZINE

AND

ANNALS OF PHILOSOPHY.

[NEW SERIES.]

JUNE 1832.

LV. *On the Electro-motive Force of Magnetism. By Signori* Nobili *and* Antinori; (*from the* Antologia, No. 131): *with Notes by* Michael Faraday, *Esq.*, *F.R.S.*, *M.R.I. Corr. Memb. Roy. Acad. Scien. of Paris, &c.**

MR. FARADAY has recently discovered a new class of electro-dynamic phænomena. He has presented a memoir on this subject to the Royal Society of London, which is not yet published, and of which we have received the simple notice, communicated by M. Hachette to the Academy of Sciences at Paris on the 26th of December last, in consequence of a letter which he had received from Mr. Faraday himself†. This relation induced Cav. Antinori and myself immediately to repeat the fundamental experiment, and to study it under its various aspects. As we flatter ourselves we have arrived at results of some importance, we hasten to publish them without

* Communicated by Mr. Faraday.

[† I am glad of an opportunity of adding a few notes to a public version of Sig. Nobili and Antinori's paper. My hasty letter to M. Hachette, in consequence, probably, of my bad writing, has been translated with some errors; and has been, by Sig. Nobili at least, seriously misunderstood. Had it remained private, it would not have been of much consequence: but as it has appeared in three or four languages, and forms the text of all subsequent papers on magnetic electricity, it is very requisite to correct certain errors which have arisen from it, especially that of Sig. Nobili relative to Arago's rotation.

My first paper was read to the Royal Society, November 24, 1831; and my letter to M. Hachette was dated the 17th of December 1831; my second paper was read January 12th, 1832. Sig. Nobili's paper is dated January 31st, 1832. Sig. Nobili and Antinori worked only from the letter to M. Hachette; but as I hope I may claim whatever is contained in my two papers, I have introduced into the present paper references, in figures included within parentheses, to paragraphs in my papers, wherever the experiments described are either altogether, or only to a partial extent, repetitions of my results.—M. F.]

N. S. Vol. 11. No. 66. *June* 1832. 3 F any

Electronics magazine

The circumstances surrounding the birth and christening of the magazine *Electronics* are well established.[30,31] In July 1929, Maurice Clements of McGraw-Hill wrote a short memo to his superiors, Edgar Kobak and Malcolm Muir, suggesting a monthly to be called *Electrons*; he followed it three months later with a longer memo outlining the editorial concept and estimating circulation and sales. Kobak assigned Clements and O. H. Caldwell (then editor of *Radio Retailing*) to obtain the reactions of various industry leaders. One of those visited, in October 1929, was Dr. John Mills, public relations director of Bell Telephone Laboratories. As Clements recalls it, Mills "not only approved the plan of *Electrons* but remarked: 'Why not call it *Electronics*—I believe I saw this engineering term in a British scientific journal.'"[30] The name and publishing plan were adopted and the first issue came out in April 1930. An interesting sidelight is that the artist who designed the cover chose a lower-case initial "ε" and thus established a typographical style that survived until 1964, when the magazine became a semimonthly (it had been a monthly and then a weekly) and the cover was restyled with a capital "E" for "greater impact." It is not true, as some have subsequently speculated, that the lower-case initial was chosen originally because it stood for the charge on the electron.

A controversy arose when *Electronics* magazine claimed that it originated the word. In a tongue-in-cheek item, managing editor W. W. MacDonald wrote at the end of his monthly editorial column in 1949:

"Back in 1930, McGraw-Hill coined the word Elec*tronics*. The suffix rolled nicely off the tongue and recently inspired the naming of another company magazine Nucleo*nics*. Now one of our confreres is toying with the word avio*nics* and we wonder where this is going to end.

"That final suffix should, perhaps, be spelled *nix*."[21]

The challenge was taken up by none other than O. H. Caldwell, who had since left McGraw-Hill and had become editorial director of *Tele-Tech* (now *Electronic Industries*). He pointed out that McGraw-Hill had certainly not coined the word "electronics," which was "already being used in England."[32] In time, Espenschied joined the fray,[18] and later Drubba;[10] the information they discovered has been summarized in the foregoing.

The contention that "electronics" stems from Britain is also supported by a letter to the editor of *Science* from E. F. McDonald, Jr., arguing that "radionics" should be adopted to describe the industry,[33] in part because "electronics is of British origin" and partly for the curious etymological reason that "radionics" is more appropriate: it "springs from the Latin 'to radiate' and the Greek 'ion,' to wander—hence 'wandering or traveling radiations,'" whereas "electronics" (by an analogous derivation that is not very analogous) stands for "wandering amber"! Although McDonald asked for the readers' reactions he apparently received none, for the matter was evidently dropped and the term was certainly not adopted.

I. E. Mouromtseff, a knowledgeable engineer who was a score of years nearer to these developments than the author, stated unequivocally in 1945 in a review article: "It may be of interest to note that the word 'Electronics' was first coined as the name of the well known McGraw-Hill technical periodical which made its appearance in April 1930."[34] My conclusion is the following. Although earlier instances of the *word* "electronics" exist in another (the scientific) context, and although the adjective "electronic" was used earlier in the technological context, I have been unable to find an occurrence of "electronics" as describing a branch of technology or industry that would antedate the first appearance of the magazine, not even the aforementioned British usage that is said to have inspired the name.

Changing definitions of electronics

It may be of interest to note how the definition of "electronics" has changed over the years. This is how Mouromtseff defined it in 1945:

"Electronics is that branch of Physics and Electrical Engineering which deals with the phenomenon of conduction of electric currents through vacuum, gases, and vapors, with their control, and their practical applications."[34]

Seeking to redefine electronics to include information processing and control, W. L. Everitt wrote in 1952:

"Electronics is the science and technology which deals primarily with the supplementing of man's senses and his brain power by devices which collect and process information, transmit it to the point needed, and then either control machines or present the processed information to human beings for their direct use."[35]

A still more recent definition, which reached back to fundamentals, was adopted by the board of directors of the Electronic Industries Association in 1961:

"Electronics is that branch of science and technology which deals with the study, application, and control of the phenomena of conduction of electricity in a vacuum, in gases, in liquids, in semiconductors and in conducting and superconducting materials."[36]

The current IEEE definition is even more encompassing:

"Electronics is that field of science and engineering which deals with electron devices and their utilization."[37]

One might well speculate as to whether another revision will be necessary a decade from now!

REFERENCES

1. Hamburger, F., Jr., "Poles and zeros," *Proc. IRE (Editorial)*, vol. 49, p. 1373, Sept. 1961.

2. Süsskind, C., ed., *The Encyclopedia of Electronics*. New York: Reinhold, 1962.

3. Döring, H., Private communication.

4. Ryder, J. D., "Poles and zeros," *Proc. IEEE (Editorial)*, vol. 51, p. 295, Feb. 1963.

5. Espenschied, L., "Origin of the word 'radio,'" *Proc. IRE (Correspondence)*, vol. 50, pp. 327–328, Mar. 1962.

6. Goodenough, E. F., "Origin of the word 'radio,'" *Proc. IRE (Correspondence)*, vol. 50, p. 2522, Dec. 1962.

7. Egidi, C., and Leschiutta, S., "On the origin of the word 'radio,'" *Proc. IEEE (Correspondence)*, vol. 51, pp. 1156–1157, Aug. 1963.

8. Süsskind, C., "On the first use of the term 'radio,'" *Proc. IRE (Correspondence)*, vol. 50, pp. 326–327, Mar. 1962.

9. Drubba, H., "On the origin of the word 'radio,'" *Proc. IRE (Correspondence)*, vol. 50, pp. 1995–1996, Sept. 1962.

10. Drubba, H., "First uses of the term 'electronic,'" *Tele-Tech*, vol. 12, p. 30, Nov. 1953.

11. *The Oxford English Dictionary*. London: Oxford, vol. 13 (Supplement and Bibliography), 1933, p. 326.

12. Fleming, J. A., "The electronic theory of electricity," *Popular Sci. Monthly*, vol. 61, pp. 1–23, 1902.

13. Stoney, G. J., "On the physical units of nature," *Phil. Mag.*, ser. 5, vol. 11, pp. 381–390, May 1881; also in *Sci. Proc. Royal Dublin Soc.*, vol. 3, pp. 51–60.

14. von Helmholtz, H., "On the modern development of Faraday's conception of electricity," *J. Chem. Soc.*, vol. 39, pp. 277–303, 1881.

15. Stoney, G. J., "On the cause of double lines and of equidistant satellites in the spectra of gases," *Sci. Trans. Royal Dublin Soc.*, sec. 2, vol. 4, pp. 563–608, 1891. (The term "electrons" occurs first on p. 583.)

16. Stoney, G. J., "Of the 'electron,' or atom of electricity," *Phil. Mag.*, ser. 5, vol. 38, pp. 418–420, 1894.

17. See, for instance, Whittaker, Sir Edmund, *History of the Theories of Aether and Electricity*. London: Nelson, 1951, p. 365.

18. Espenschied, L., Private communication; see also *Tele-Tech*, vol. 12, pp. 60–61, May 1953.

19. "On the electro-motive force of magnetism, by Signori Nobili and Antinori, with notes by Michael Faraday," *Phil. Mag.*, ser 2, vol. 11, pp. 401–413, June 1832.

20. Cited in Thompson, S. P., *Michael Faraday, His Life and His Work*. London, Cassell, 1898, p. 116. (Silvanus Phillips Thompson, 1851–1916, was a kinsman of Richard Phillips and had the original of Faraday's letter in hand at the time that the biography was written.)

21. MacDonald, W. W., "Business briefs," *Electronics*, vol. 22, p. 63, Nov. 1949.

22. Archer, G. L., *History of Radio to 1926*. New York: The American Historical Society, 1938, p. 120.

23. Gherardi, B., and Jewett, F. B., "Telephone repeaters," *Trans. AIEE*, vol. 38, pp. 1287–1345, Oct. 1919.

24. Meyer, F. W., "Die freiströmende Elektrizität der technischen Elektronik und die entwicklung der Starkstromtechnik und des Maschinenbaues," *Elektrotech. Z.*, vol. 42, pp. 689–693, 725–728, 1921.

25. Meyer, F. W., "Die Steigerung der Empfindlichkeit in der technischen Elektronik nach den Anforderungen des Maschinenbaues," *Elektrotech. Z.*, vol. 43, pp. 981–982, 1004–1009, 1034–1038, 1922.

26. *Science Abstracts (Sec. B: Electrical Engineering)*, vol. 25, 1922, p. 162.

27. *Science Abstracts (Sec. B: Electrical Engineering)*, vol. 25, 1922, pp. 162, 598.

28. Clifford, G. D., "Change of the Institution's name," *Radio and Electron. Eng. (Editorial)*, vol. 27, p. 165, Mar. 1964.

29. Fawcett, F. T., "Some absurdities of wireless nomenclature," *J. Inst. Wireless Technol.*, vol. 1, pp. 52–61, May 1927.

30. *McGraw-Hill News-Bulletin*, vol. 2, p. 1, Feb. 9, 1961; p. 2, Feb. 23, 1961; p. 2, Mar. 11, 1961; p. 2, May 4, 1961.

31. Henney, Keith, Private communication. (Mr. Henney was associate editor, 1930–1935, and later editor, 1935–1946, of *Electronics*.)

32. *Tele-Tech (Editorial)*, vol. 8, p. 17, Dec. 1949.

33. MacDonald, E. F., Jr., *Science (Correspondence)*, vol. 97, pp. 554–555, !943; reprinted, with editorial comments, in *Wireless Engr.*, vol. 20, p. 437, 1943.

34. Mouromtseff, I. E., "Electronics and the development of electronic tubes," *J. Franklin Inst.*, vol. 240, pp. 171–192, 1945.

35. Everitt, W. L., "Let us re-define electronics," *Proc. IRE (Editorial)*, vol. 40, p. 899, Aug. 1952.

36. *Proc. IRE (Industrial Engineering Notes)*, vol. 49, p. 35A, Nov. 1961. (Excerpted from the *Weekly Report* of the Electronics Industries Association, Aug. 28 and Sept. 19, 1961.)

37. *IRE Dictionary of Electronics Terms and Symbols*. New York: IRE, 1961, p. 51. (Taken from "IRE Standards on Electron Tubes: Definition of Terms, 1957," *Proc. IRE*, vol. 46, pp. 983–1010, July 1957.)

Do you use your Engineering Societies Library as much as you might?

Nowadays there is a great deal of talk about the information revolution, information retrieval, and the mechanization of libraries, talk that might even lead you to assume that libraries as they now exist have become next to useless. Such is hardly the case. A library, even the smallest and most specialized, is only useless when its potential users don't know it exists, don't know what's in it, don't think of using it when the need arises, or don't use it because they don't know how to use it. You'd think the last reason listed would be the last to apply to engineers, who are, after all, educated men. But no, says ESL Director Ralph H. Phelps, engineers in general don't know how to use libraries effectively because they haven't been trained to use them. In fact, if you are like most engineers, you probably don't use your own Engineering Societies Library nearly as much as you might. It may pay off for you to spend a few minutes considering just what ESL is, and to find out how it is prepared to help you—no matter where you are, whether you are near ESL, which is in the same building as IEEE Headquarters in New York, whether you are on the West Coast, or in Europe. No farther away than your telephone or mailbox are many resources and services you can tap free or at moderate cost—book borrowing by mail, photocopies of articles, microfilms, literature searches, translations from any language, and so on. Although the mechanization of libraries is not yet here, you should realize that through your IEEE membership in ESL (which is in itself one of the most comprehensive engineering libraries in existence), you already have entrée to a potent information grid that links you to the library resources of the world. Even if ESL doesn't have what you need, it can probably lead you to where it is. Besides, you already pay, through your IEEE dues, $60 000 a year toward ESL's support.—N.L.

What is ESL?

To start off, Mr. Phelps, can you tell us something about the origin of the Engineering Societies Library?

Nineteen-thirteen was the year that the Engineering Societies Library was established as a unified library. The original collection grew from the libraries of four professional societies: the American Society of Civil Engineers; the American Institute of Mining, Metallurgical, and Petroleum Engineers, then known only as the American Institute of Mining Engineers; the American Institute of Electrical Engineers; and the American Society of Mechanical Engineers.

Of the four libraries, the Civil Engineers' Library was the largest, which is natural, considering that ASCE is the oldest of the civilian engineering societies.

The AIEE library was also a large one and was particularly strong in old material, having incorporated the Lattimer Clark Library, which is commonly known as the Wheeler Collection. Incidentally, Wheeler had given his collection to the AIEE library with the provision that it be housed in a fire-resistant building.

Would you say, then, that the electrical engineering collection is particularly good for historical searches?

Yes indeed. The collection is also now very broad, for over the years, other professional societies have incorporated their libraries with ESL. There are now 12 societies that are associated with ESL and that contribute to the support of its operation. This support amounts to 50 cents from each member of each associated society. By contrast, I should point out that most public libraries are supported to the extent of $3 per capita.

The question of library support itself raises an interesting historical point. Andrew Carnegie, who in 1904 gave a million dollars toward the establishment of a headquarters building for the societies and their libraries, is often regarded as the chief benefactor of libraries in America, which of course he was. But at the same time, the connotation came about that the libraries are free. Actually, Carnegie gave buildings and left it to the people to pay for the holdings, the use, the staff, and whatnot. This is still true, and such operations are part of our increasing costs.

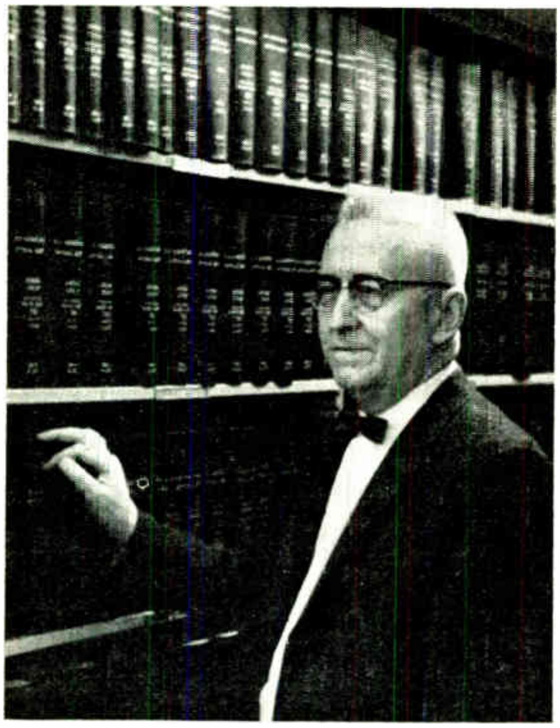
Roughly what proportion of the collection is devoted now to electrical engineering and electronics?

This is a disturbing question, because you can no longer separate fields of engineering. Perhaps you could do so once, but you certainly can't now. Things that electrical and electronics engineers didn't dream of as being part of their business some years ago are now in their domain.

Then how do you go about making new acquisitions? Aren't they collected on the basis of specialties?

No, we combine them all. We look at engineering as a package. For instance, although many of the journals used by electrical and electronics engineers are in physics, and related areas, these same journals may be of equal

If you are like most engineers, you are not using your Engineering Societies Library as much as you might. In this interview, ESL Director Ralph H. Phelps talks about ESL resources and services—in short, about what ESL can do for you



interest to mechanical engineers. There is no separation that is clear cut, to my mind at least.

Probably the test of the successful library is how well users are able to use it. What's the probability that an IEEE member coming in, phoning in, or writing in will find what he wants?

Quite high. Either we have it or we can tell them where it is. No library in this day of many thousands of scientific and engineering journals can carry them all, so we have to be selective.

Our selectivity is not arbitrary, however. In acquiring new material, we cooperate informally with various libraries, and particularly with the New York Public Library. It tends to emphasize the collection of scientific material whereas we try to select more in the applied science and engineering area.

That raises a question I was wondering about. In looking through your Special Libraries Directory, which describes the thousand libraries in the New York area, I noticed that the Science and Technology Division of The New York Public Library lists a collection of 450 000 books and periodicals, whereas you have about 190 000. Yet you say you are the world's largest engineering library. How do you make your measure?

The other libraries include much more general science than we do—more chemistry, mathematics, meteorology, and so on—while we concentrate on engineering. They have a far wider scope of coverage, whereas we are trying to center around what we think the engineer needs or what we learn that he needs from our correspondence, telephone calls, and recommendations from users

Engineers as library users

Many of our members may not be aware of the services ESL offers, and which of them are free and not free. Do you believe that IEEE members are using the library as much as you think they could?

I don't think that any engineer or any group of engineers is using it as much as they might. This fact can be tied back to the lack of training in the use of information and/or libraries. Engineers do not get the service they could get simply because they do not ask for it. Many aren't trained in library use and many are completely unaware, say, of the existence of the Engineering Societies Library. You could say that this is the fault of the library for not publicizing its services. However, we think it best to put more of our limited dollars into resources and services than into publicity. We do, of course, appreciate it when the societies make a point of informing their members about the services that are available in ESL.

Is this lack of training in library use a problem that librarians talk about?

Yes. This is a major problem that is talked about often in library circles. For years, chemists, and to a lesser degree chemical engineers, have had courses in their literature, and there are several excellent guides to the literature of chemistry because there has been a market for them in schools. You find more such guides in the sciences. There are guides to the literature of mathematics, guides to the literature of physics, and so on. However, as far as I know, there have been only a few guides to the literature of engineering, for which I understand there was little market, apparently because so few schools teach the literature of engineering. I believe the American Society for Engineering Education is working on this problem.

Is this problem becoming more acute now because of the growing raft of published material?

Undoubtedly this is the situation. There has always been a problem. However, there was a time when many engineers could get by with a few handbooks, textbooks, and catalogs; but as engineering becomes more diversified and sophisticated, it is necessary to go beyond these—in fact, often considerably beyond them.

What's where? How can I get it?

What's the best way for an IEEE member to approach ESL? What if he's a distant user? Should he phone?

If the member wants to borrow a book, the best thing he can do is write to us and say so, giving the author and/or title and date, if he knows it. Or he might simply request a recent book or an old book, depending on the circumstances, dealing with a particular subject, which we would select on the basis of the information he gives us. He may borrow books (as many as three at one time) for a period of two weeks. This means only the time the books are in his possession. We don't include mailing time. Each book costs him 50 cents a week. The charge for members who come in personally to borrow books is the same.

If a member writes in, how long will it be before he receives the books he asks for?

Usually the book will be sent out the next day. If he requests a particular book, it goes out quite promptly, assuming it is not on loan to somebody else at the moment. If we have to do some digging for a book by subject, it will take perhaps two to three days. But the search goes forward pretty fast. Of course, the books are sent by a book-rate mail delivery, which I assume is more equivalent to parcel post than to first class.

Suppose it does take three days to find a book he wants? Is the member charged for that? Or is it free?

No. He isn't charged. This service is free. In this sense, we do some free literature searching. We get a book or list a few books for a member even though he may not want to borrow them. For extensive searches, however, there must be a charge.

What about a telephone request for information? Suppose a member calls up and says, "I need to know thus and so," and it takes one of your staff people a day to find it and call him back. Would that be free?

Not for a day's work. There would be no charge for, say, up to half an hour's work.

Might a library staff person find a periodical or book, and read a paragraph or two to the member over the telephone? Would that be included in the free type of service?

We try not to read much over the phone. This not only takes time, but quite often the person at the other end doesn't get a precise record of it. We'd rather send him a photocopy, if he can wait that long, and then he has the exact text.

What procedure do you prefer on such requests?

On requests for photocopies, or the delivery of a specific item to the inquirer, we prefer to have a letter because errors do creep into telephone calls. However, a member might wish to call in first to see if we have what he needs, then verify his request by letter.

How long does it take to get a photocopy, and what does it cost?

Most orders for photocopies, and we handle 10 000 or more per year, go out in a matter of one to three days. The order is verified from the letter, the book or periodical is taken from the shelf and verified, the copies are produced, invoices are prepared, and the whole job is wrapped up. Some people order through their organization, which may have deposit accounts that provide for an automatic billing against a sum already on deposit. The charge is 40 cents per page plus 60 cents handling charge per item.

I might add that although we have a more or less established procedure, in pressing cases we sometimes do push things through nonroutinely for members. For instance, in response to a phone or letter request, we might send out a single-page photocopy free of charge. In general, though, we do charge.

Another way in which we may already be serving some of your members without their actually being aware of it is through their company libraries, which may hold a company membership in ESL. Actually, the requests they make of their company library may get routed to us, so that indirectly they are tapping the resources here.

Would it be worthwhile, then, for IEEE members to urge their companies to join ESL? Would that facilitate their own accessibility to the material?

I believe that many company libraries would benefit from a company membership in the Engineering Societies Library.

We haven't yet talked about the user who comes to ESL to do his research. He has here, of course, a large, bright, quiet reading room with lots of table space, open shelves with bound and loose recent periodicals, directories, complete card catalogs, etc., and rapid, courteous service—I am speaking from experience—but he doesn't have free access to the central stacks. Why is that?

There are several basic reasons for closed stacks. First, the stacks are built as compactly as possible, so we can get the maximum storage per cubic foot. There is no space for tables, no facilities for doing any work in the stacks. Second, when persons not well acquainted with libraries go to stacks, they are likely to misshelve books and this can cause problems; later requests can't be fulfilled because the book can't be found. The third and most important reason is that a book can only be put in one place in the stacks. The card catalog has analytic entries in it, to the extent of probably an average of five per book, so that if you go to a section of the catalog on a specific subject you may find five times as many cards representing books that are in other sections of the stack. For example, a general book on electrical engineering may have specific sections dealing with radio, radar, telephone, and telegraph, in which case we analyze all of these in the catalog so that you can get to them from whichever angle you take; but the book can be in only one place in the stack.

By way of summing up the services ESL offers, just what part of these services is covered by the 50 cents that each of our IEEE members pays?

First, this 50 cents buys a comprehensive collection of books in the field of engineering. Insofar as IEEE members are concerned, this is electrical/electronics information. It buys persons trained both in library techniques and in science and engineering who can analyze and catalog and arrange these books so that they can be gotten to. It provides for the staff that goes through many hundreds of notices of new publications, for reviews to determine which ones to acquire—a selection factor. It provides for persons who answer telephone and mail requests from members for brief information for which no charge can effectively be made.

Toward the interlibrary grid

Mr. Phelps, what can members look forward to in the future growth of the Engineering Societies Library? How fast are you actually growing, and what are your plans for further development?

Well, just statistically speaking, the rate of acquisition of books has doubled since World War II. In the same period, book loans have doubled, telephone inquiries have increased four times, and letters answered have increased fivefold. Orders for photoprint copies have gone up three times. The large increase in telephone inquiries, by the way, is a trend that you find throughout libraries generally.

Can you tell us something about your aim of developing a fully integrated engineering information center, which is mentioned in one of your reports?

With the development of newer and advanced means of communication, we anticipate that it will be easier to get information from other sources or to transmit information to other places—let's call them information centers. A center may be any major or specialized collection of information. For instance, a private company may quite properly consider its own library to be a center.

Are you talking about an interlibrary grid, like a power system, where the information would flow throughout the grid?

That is the thought. Up to now, however, there has been no very substantial agreement as to how all this is to be accomplished technically. Perhaps less is known about how it can be financed, outside of the government. People often refer to the National Library of Medicine and its computerized setup and publication program as being quite advanced and quite the thing. I've heard comments both for and against it, but let's agree that it is advanced. However, the government spends about $5 million a year to operate that library. By contrast, let's consider the operation of the Engineering Societies Library. It makes all of its publications available to the Engineering Index for indexing; EI publishes a variety of indexes (including one section on electrical and electronics engineering), and disseminates information about the ESL collection throughout the world. Although ESL and EI are separate organizations, their total combined budget is of the order of $1.2 million. Without more money, it is difficult at this time to see how we can do much more. Although engineers have a rapidly growing appreciation of technical literature, I don't believe they quite understand how rapidly the cost is growing.

You mean the cost of information handling?

Yes, the cost of acquiring and getting information. Frankly, it's been worked out in one study that the cost of scientific publications is increasing 8 percent per year. Moreover, the number of publications in science and engineering grow at the rate of 12 percent a year. Put these two figures together, and you have an idea of how much more you should spend per year for publications.

Nevertheless, it is a fact that the Engineering Societies Library and the Engineering Index actually constitute a potent information center today. We have an organization, ESL, acquiring broadly across the board of engineering, plus an organization, the EI, that is producing a detailed analytic subject and author index of the articles within the journals, and disseminating this throughout the world. Additionally, in 1964, the Engineering Societies Library, with the cooperation of a publishing firm, produced a 13-volume subject catalog of all of the books, reports, monographs, etc., in the library. These volumes are available in libraries throughout the world. Thus, an engineer, wherever he is, may go to a nearby library and through this catalog find out about our books and other types of publications, or through the Engineering Index find information on the articles within the journals, reports, proceedings, transactions, etc., in ESL. Obviously, our information dissemination system is already quite effective. We handle inquiries from individuals and from libraries all over the world.

Some ESL services

Bibliographic
Book loans
Book reviewing
Co. membership
Film loans
Information
Mail service
Microfilms
Photoprints
Published subject cat.
Reference
Referral
Retrospective search
Translating

Do you feel that it will be very long before a tightly knit and comprehensive international interlibrary grid will be developed?

I'm hopeful that it will be developed soon. However, it isn't going to be a matter of revolution but of evolution. Actually, we already have a catalog of the titles and holdings of journals in 700 large libraries in the United States and Canada. We also have a world list that tells us about some things abroad. The world list is not quite as extensive, in terms of number of journals, as is this "Union List of Serials in Libraries of the United States and Canada." It lists over 150 000 different journal titles in 700 libraries.

So that if a member comes to you, he already has a good tap on the resources of the world?

Yes, we already have this network. It isn't as mechanized now as it undoubtedly will be in the future, but the very existence of the directories we have been talking about, and many others, constitutes the framework of the system.

The great wave

Suppose now, Mr. Phelps, that as a result of this interview, you start receiving many more requests from IEEE members, say a growing wave of requests as members get into the habit of thinking of ESL when they need information. Do you welcome new requests? Are you prepared to handle them?

We always welcome new requests. We never have enough money to do all the things we would like, so that we'll probably have to charge for many of the services, but we haven't yet reached a limit—there are still more things we can do free.

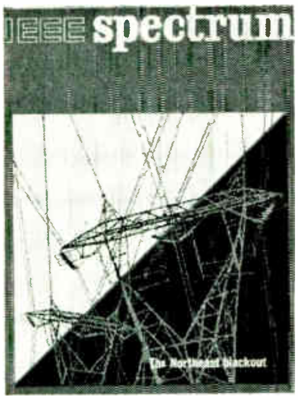
Do you have actual statistics as to what percentage of IEEE members are now using the library?

I can't state a figure because many requests come in through a person's library or his secretary, or he may not identify himself as being a member. This happens both with telephone and letter requests. When someone writes a letter asking for a book loan, without saying that he's a member of any society, we try to deduce from the subject matter of the request what society he may belong to and we look him up in the societies' membership lists. It would be helpful if each IEEE user did identify himself as a member.

Members then do have more privileges than nonmembers?

Yes, indeed.

For more information on the services of ESL, write to Ralph H. Phelps, Director, Engineering Societies Library, 345 East 47 Street, New York, N.Y., 10017.



Comments on the Northeast power failure

The open invitation to the membership to comment on Gordon D. Friedlander's article, "The Northeast power failure—a blanket of darkness" (IEEE Spectrum, pp. 54–73, Feb. 1966), produced a large volume of interesting correspondence. The following letters were selected, after careful review and consideration, as representative of the many observations on the blackout made by Spectrum's readers. The editors regret that insufficient space precludes the publication of many more of these communications

Integrated central control does not imply public ownership

I have read the article on the Northeast power failure with interest, particularly since it is the first comprehensive engineering comment I have seen that is devoid of a commercial or political slant.

In some of the numerous previous articles and dissertations on this failure, I have seen some favorable references to Britain's completely integrated and centralized control electric supply network (the world's largest single-control system), and, of course, opposition in the United States on the grounds of "creeping socialism."

May I assure Americans that an integrated, single-control system does not necessarily mean public ownership. Britain instituted central control of electric generation and distribution in 1929, and it functioned well for 18 years without any effect on the sovereignty of the numerous private generating companies...

We are not without our difficulties. For example, a sector of west London was blacked out in 1964 for about four hours because of a busbar fault and fire at the Battersea Power Station. The fault area was rapidly confined, however, and about 75 percent of the service was restored in minutes. The delay on the remainder occurred because an external underground cable burned out by reverse flow into the short circuit...

In another instance, during the winter of 1963–1964, an unusually heavy frost caused massive flashovers in switching stations in the Midlands. These effects were quickly reduced to a minimum by [system] isolation, so that each generating station fed its own service area.

To sum up, the North American answer lies in completely integrated central control. The restoration situation in CONVEX is a clue. Central control does not mean creeping public ownership—that is an ideological matter.

E. J. Dawes (SM)
London, England

When should a system break its interconnection ties?

Mr. Gordon Friedlander's article on the Northeast power failure in the February SPECTRUM was of great interest to me as a longtime electric power engineer. Please convey to Mr. Friedlander my appreciation and admiration of not only this but also of a number of other articles written by him during the past several years.

I have followed the various articles, news releases, and analyses concerning this catastrophic outage with much interest, because I feel that the occurrence deserves the careful attention of electrical engineers who are in a position to analyze electric utility system weaknesses and suggest corrective measures. Complex as our vast and interconnected electric power systems are, there are still basic and extremely important protective features that, in my opinion, can and should be incorporated for their control. If these are included, they should protect against major disturbances, such as the blackout that affected a great part of New York and New England, including all of New York City, for a long and terrifying time on November 9–10.

While the information released, including Mr. Friedlander's article, is of great interest and value, in none of the information on this subject coming to my attention is the following leading question answered:

Why did not the protective features of the Consolidated Edison system quickly and automatically disconnect the ties with the Niagara Mohawk and CONVEX systems so that its own system, with its large reserve of about 1350 MW, could have weathered the disturbance that originated outside its own system?

In Mr. Friedlander's article, the block diagram, Fig. 11, on page 63 shows the intertie power flows before and during the disturbance. The flows from Niagara Mohawk to Con Edison just before the disturbance totaled 365 MW, and the flow from Con Edison to CONVEX was 35 MW, for a net interchange of 330 MW into the Edison system. Immediately on occurrence of the disturbance, the flows were reversed from Con Edison to Niagara Mohawk in the amount of 850 MW and increased to 250 MW to CONVEX. The net change over these ties was 330 + 850, plus 250, for a total of 1430 MW. If these ties had been opened as this reversal started to build up, the Con Edison system would have been relieved of an immense outgoing surge of power and its spinning reserve could easily have absorbed the net 330 MW that it would have lost from Niagara Mohawk.

Biased tie-line load control can and should incorporate protection against excessive power flows imposed on one system by one or more neighboring systems to which it is interconnected, if such excessive power flows endanger the system. A utility system is primarily responsible to its own system customers and it should guard against anything that endangers such primary responsibility. In this case, the interconnected PJM system protected itself by opening the ties with Con Edison, but Con Edison apparently did not protect itself from Niagara Mohawk or CONVEX. Why did it not?

Not only did such failure create a dangerous situation that was filled with potential catastrophe, but Con Edison is also paying dearly because of the tremendous cost of the long outage and for the repair of three of its most modern steam-electric generating units, totaling 1500 MW capacity, that were damaged during the power failure.

W. L. Newmeyer (F, L)
Denver, Colo.

System operators and steady-state capability limits

The invitation to the IEEE membership to comment on Mr. Friedlander's informative article prompts me to make a few germane remarks about two separate, but related, parts of the article.

First, on page 56 under the heading "The incident," it is stated that a backup relay tripped out line Q29BD at Beck, thus shifting the load of this line onto the remaining four lines, which tripped out sequentially because they were loaded beyond the critical level. The implication is that the steady-state stability capability of the four-line transmission system was exceeded. If so, then it follows that the system must have been operating fairly close to the steady-state capability limit of the five lines before line Q29BD tripped. Suppose that, under the same loading conditions, a line fault had occurred in the no. 1 tripping zone of this line that would have involved the primary tripping element of the same relay; the same line would have tripped. The remaining four lines would have had to pick up the load of the tripped line and would have tripped out in the same sequential order, causing the same dire consequence. In this case, however, the evidence of a critically loaded transmission system would not have been so obvious. It seems that the backup element of the relay operated correctly at its predetermined setting, with no malfunction of the relay.

An overload alarm relay set at a maximum-safe-load operating value below the backup relay setting may have saved the day. There may be some merit in having a power overload alarm relay for each of the five lines at Beck, if heavy line loading is to be continued.

Second, on page 62, the explanation for the breakdown in area 4 is well documented and clearly explained. A case in point, of concern to me, is that 11 minutes had elapsed after the first power surge to the time, 5:27 P.M., when four boroughs of New York City lost their source of power. During this interval, voltages and frequency fluctuated. If adequate protection to meet such a contingency was operative at the time, it is paradoxical that Con Edison did not separate completely from all ties and maintain its own system, since its spinning reserve at the time was sufficient to meet the city load demand (according to quoted values).

The sequence of events leading up to "blackout" strongly suggests that Con Edison's operating policy was, "hold the tie lines, no matter what happens." Seasoned system operators have a "sixth-sense feel" for a system and they instinctively know what should be done to alleviate a serious, prolonged system disturbance. It is inconceivable that a veteran system operator would live with a seesaw megawatt surge of the magnitude that Con Edison experienced for 11 minutes without cutting loose from major tie lines—if he had the authority to take the initiative.

My compliments to Mr. Friedlander for a real contribution to the IEEE membership and to others who play a vital role in maintaining the continuance of power flow.

Paul E. Shields (SM)
University of Maine
Orono, Maine

Use of underfrequency relays for load shedding

It is evident to us that some immediate remedies can be applied to the existing pool systems to limit the affected area and to speed up load recovery during major disturbances. Well-planned use of underfrequency relays for automatic load shedding and isolation of systems can effectively and economically prevent shutdowns. Out-of-step conditions can also be detected on the first slip cycle to separate systems not running in synchronism to prevent cascading of the trouble.

Both the General Electric Company and Westinghouse Electric Corporation, two major protective relay manufacturers, have emphasized in the past the importance of underfrequency relay applicaton in load shedding when demand exceeds system generation.

C. W. Cogburn and G. C. Kelly of the Florida Power and Light Company also pointed out in their article in *Electrical World* (Nov. 4, 1957) that application of underfrequency relays can be used to maximize service continuity by preventing system shutdown during system disturbances and by speeding load recovery.

Load shedding has not been a popular word in the electric utility industry in general because it means interruption of customer service, but with all these sad and costly experiences in blackouts, it is high time for us to review this philosophy. It is simple logic that short customer interruption is far better than major system shutdowns that involve equipment damage together with loss of start-up power.

Pool interconnections expose all systems to transient power swings and these will increase as the power system grows in complexity. Transient stability characteristics in pool operations should be studied and swing separation relays can then be applied intelligently and strategically to minimize the spread of a disturbance.

No power system can be built to be 100 percent reliable, even if an infinite amount of money were available to build an infinite number of tie lines and generating plants. The philosophy of "critical load saving" needs to be developed and applied in a coordinated manner. Customers will forgive a short outage incurred in an emergency but cannot understand the extended blackout. A great deal can be done with existing, proved, low cost, and reliable relays that have been long available to the industry. Most transmission systems are extremely reliable but what excuse is there for not applying low-cost backup protection to cover the worst disaster, whether the problem is caused by transmission lines or human error?

The techniques and tools have been with us for a long time, but the reliability of lines and equipment has lulled most of us into thinking we do not need backup protection. It is unfortunate that some of the leaders of the industry still feel that protective relays, if applied in this situation, will probably do more harm than good!

Roger Lane (M)
Li Liu (M)
Iowa Electric Light and Power Co.
Cedar Rapids, Iowa

Planned load shedding and data collection for emergencies

In reading Mr. Friedlander's excellent article, one has to come to the conclusion that the CANUSE system is quite weak from the stability point of view. Thus, any disturbance may break it into several isolated areas. As long as this is the case, any power company should be prepared for a sudden generation loss or load increase. The rule that seems to be as old as the power systems themselves says that the emergency situation can be met by:

1. Planned load shedding when underfrequency develops.
2. Maximum reliability of the supply of all essential auxiliaries in the power plants.

The second conclusion from the outage development is the inefficiency of data-collecting and processing systems and, to some extent, the lack of coordination. It seems that any power pool, even one not completely integrated, should have a common dispatching center over the local energy control centers of the member companies. This may facilitate the data collecting and analyzing in the emergency case, and the coordination of the efforts to restore the normal service.

J. B. Jordan (SM)
Jean Robert
Laval University
Quebec, Canada

System reliability and stability

Somewhere along the line of continuous discussion that has followed the great power blackout, its subject was changed. At the beginning it was "reliability of electric energy supply systems," but this later turned into a controversy between proponents and opponents of strong ties. Unfortunately, a "strong" tie has not been defined. Also, how "strong" is "strong enough." So, like many other issues of our times, this turned into a matter of semantics. Is a strong tie determined by the number of lines connecting two systems? Is it the cross section or transfer capacity of these lines in relation to the sizes of the systems, or to the emergency power flow? Is it the setting of the tie breakers? Or is it several—or all—of these things? This is an important issue, but no quantitative definition has been given thus far.

The main issue, however, still remains reliability and the stability of electric systems. Reliability of a system is, above all, a measure of responsiblity which cannot be shifted to its ties with other systems. A system must be reliable even with all its ties open. If it is not, it cannot claim to be an independent system. Moreover, ties—like anything else—have their disadvantages. The most apparent disadvantage is that a disturbance in one system can spread over all interconnected systems.

Absolute reliability, including absolute stability, are goals that a system can approach asymptotically, but can never reach. Dependent upon the magnitude of the disturbance, a failure or instability can appear on any system. This does not necessarily mean an outage, for alternative sources and supply lines are used in these cases. It can be shown that all methods for increasing reliability of systems can be reduced to one: the provision of alternatives to the working components of the systems. The differences between the methods consist merely in the way the alternative components are introduced into service in an emergency. This can be done by manual exchange of the faulted component, reswitching, or automatically. The last method is the quickest and the best, and leads to automatic transmission and distribution. Although automated transmission and distribution may be viewed as an extension of the already automated generating facilities, it cannot be achieved in the immediate future because of the considerable costs involved. This barrier can be overcome by continuous research into reliability and stability of systems, an area in which much remains to be done.

The lesson to be learned from the November 9 blackout is that concatenations of faults and subsequent events leading to great power outages are possible. To prevent such chain reactions, they must be first anticipated. Clearly, continuous research into methods to foresee and prevent such outages is vital. Nothing stimulates research better than new ideas. Therefore, the writer suggests that one entire issue of SPECTRUM be devoted to new ideas on reliability and stability of electric energy supply systems. Obviously, one staff-written article and the letters following it cannot constitute the entire response of the IEEE to an event that rocked the electrical world. The proposed special issue of SPECTRUM could contain not only papers, but also original suggestions on reliability and stability, briefly expressed by IEEE members. In view of the great need for new ideas, their submission should be encouraged, even at the risk that some of them may be found to be impractical.

Henry Greber (SM)
New York, N.Y.

Responsibilities of management and the individual engineer

The Northeast power failure may have had the salutary effect of providing the reason for giving us a cover and feature article in IEEE SPECTRUM on the subject of electric power, but it might also have another effect. Since it is known that not all of the 60-c/s problems have been solved there should be a renewal of interest, on the part of teachers and students of electrical engineering, in the low-frequency end of the art.

Mr. Friedlander's article in the February issue is commendable in its narration of events and its penetration into the technical problems involved. This is not to say that the author has made any discoveries. Electric circuit performance is not a scientific mystery, and the questions raised have been asked before, but Mr. Friedlander has brought them into professional focus, just as that amazingly perceptive and expeditious report of the Federal Power Commission has brought them into national focus.

Both reports deal essentially with the physical and technical aspects of the problem, and the questions they ask are generally directed toward a correction of these. Certainly, this sort of thing ought to be done; but if, in resolving the technical problems, we are to be plagued by some underlying situation that bequeathed us the problems in the first place, then we need to look deeper than the revealed errors of the past and find out what permitted them to occur. Only then will we discover the cause of the failure and preclude the same kind of failure from occurring elsewhere in our industrial society

Mr. Friedlander's comment, in explaining the delays that were evident in the restoration of service, "that during the night of November 9–10, the cause of the massive power failure was not known," is pertinent only within narrow limits. To each of the electric power systems involved, the cause of its failure was different, but it should not have been one for which some of them seemed to be so inadequately prepared.

What does Mr. Friedlander mean by "cause of the massive failure"? Certainly the initiating circumstances were (or should have been) known to someone at the Sir Adam Beck station within moments after its occurrence. Within minutes—or less—it should have been apparent to load dispatchers throughout the vast area involved that they had finally come to grips with a case of electric system instability.

This "phenomenon" of instability, if it may be called one, has been treated academically—if at all—in recent years. In the earlier years of interconnected operation it was a real problem, but it was well defined. As transmission capacity increased, relative to the capacity of generating stations, the problem was relegated to textbooks and an occasional technical paper. Now, as the mass and concentration of generating capacity has overtaken and surpassed existing transmission capacity, the immutable laws of physics have once more asserted themselves. The writer can well recall, on this very subject, being told by a utility executive to "throw away the slide rule and forget you are an engineer." In other meetings, he can still feel the silence with which the mention of the word "instability" was greeted, as if it were something archaic and ethereal.

If a specific event must be singled out as the cause of the failure of November 9, then it should not be sought in the initiating circumstance that triggered the whole operating disaster, but in those events that occurred immediately afterwards. It should be noted, however, that a correction of these events does not ensure that the problem has been solved; it may only have been transferred elsewhere.

One question that should be added to those enumerated by Mr. Friedlander is: What is the ethical responsibility of the engineer in industry to the public, as compared with his responsibility to the management of his company? When the writer graduated from an engineering school some 40 years ago, he gained the impression from one of his teachers, Dr. Alexander G. Christie, that these were separate and distinct responsibilities. The underlying ethic of the engineer's creed was his dedication to the principle that "engineering is the art of applying science to the efficient conversion of natural resources for the benefit of mankind."

For many years, the writer had the privilege of working under the direction of a prominent engineer whose credo was "give me the answer you believe to be right, not the answer you think I want." Having the answer, he would make a decision, based entirely on the authority of his own executive position and without any need or attempt to subvert his subordinate's judgment if it were different from his own.

Today, the engineer's responsibilities are too often tangled with those of management. To get the answers it wants, management will sometimes resort to the strategy of "divide and conquer," with its attendant toll on the professional integrity of its engineering organization. There are those who will be quick to point out that management usually includes engineers who have reached executive positions, and that they will be the outstanding engineers of the industry. Some undoubtedly are, but there are not enough to exercise any dominant role in a subject so complex, so technical, and so far-reaching as the one under consideration. Furthermore, the selection of engineers for executive positions is not always based on their dedication to the public interest, nor are they always the professional intellectuals of the industry.

The tenor of engineering education and development today is toward what is termed "management." The curriculum includes an early elimination of the slide rule, a practice some engineers follow to their advantage.

The questions that surround the dynamic stability of a nearly infinite electromechanical system cannot be answered by this kind of engineering. They cannot be answered by the computer alone, nor can they be answered without it. The computer is only a tool by which a component engineer can test the performance of a given system under the conditions he believes to be valid. It is not a means whereby management can test the effect of changing the variables in order to pick an answer that statisfies some predetermined set of conditions.

There is such a thing as the "wrong" amount of interconnection capacity. The evidence indicates that the wrong amount was being used to intertie the essentially hydroelectric generation around and beyond Niagara Falls and the predominantly thermal capacity of downstate New York and New England. The most accurate computer available, using the most carefully selected input data and the most realistic assumptions, would have predicted the inevitable answer of November 9.

The wrong amount of interconnection capacity is not necessarily a case of being too little; sometimes even less would be better. Nor does this mean that more interconnection capacity would be worse, as some are now contending. The wrong amount of interconnection capacity lies somewhere between the amounts needed to connect two systems loosely or to integrate them completely. It is a dangerous amount, and it possesses the merits of neither philosophy. It is a management compromise that carries with it the aura of engineering approval.

In its haste to meet the challenge of a federal transmission grid, which was proposed a few years ago, the Edison Electric Institute published a map of existing and proposed interconnections submitted by the managements of its member companies. The objectives of this publication were political, and time was not afforded to the engineers of all the companies to examine the proposals that concerned their own systems. Publicity releases such as this, unless successfully challenged, become a part of the fabric of management. It is this kind of thinking which, in my opinion, represents the real cause of the failure of November 9.

It is unfortunate that the subject of Cornwall has been injected into the discussion of causes and remedies for failures like that in the Northeast. That it was done so soon after the blackout by a spokesman for the Consolidated Edision Company can be written off as a bit of transparent propaganda. Cornwall should or should not be built on its own merits. To pretend that it is essential to the improvement of service is ridiculous. Cornwall could have contributed to the situation on November 9–10 only to the degree that its operation was predicated on the need to meet such a contingency. The performance of other stations on the metropolitan system, in the light of reality, leads to the conclusion that only by chance would Cornwall have made any difference.

As a technical society, IEEE has carefully avoided the politics of power supply in this country, as well as abroad. But the first and last sentences of Mr. Friedlander's opening paragraph, concerning the role and the responsibility of IEEE in this matter, touch upon something that must not be avoided if the existence of the Institute is to have meaning. By default, we have permitted to go almost unchallenged a plan, which was politically motivated, for interconnecting the electric systems of North America. Whether or not it was soundly conceived is not the question; the interconnection of electric systems is now one of international scope and national importance. The answers must not depend upon who supplies the power, or under what political philosophy. It is time for an appraisal by a competent and independent engineering authority. Truly, quoting Mr. Friedlander, "The IEEE has a unique role in the wake of the November 9 power failure."

Charles Morrison (SM)
Hagerstown, Md.

Weaknesses cited in power industry operation

The blackout of November 9, 1965, revealed a number of fundamental weaknesses in the electric utility industry. These were, broadly, of two general kinds:

1. Those that permitted a disturbance to be initiated.
2. Those that permitted the disturbance to evolve into a major power blackout.

These weaknesses arose because there was no universally accepted, fundamentally sound philosophy of power system protection. The two major technical functions in the operation of a power system are production and protection.

Electricity can be produced without protection, but to provide high-quality electric service, the protection function must be superimposed on production. Protection is a discipline of wide scope, incorporating everything of a technical nature—apart from production—necessary to provide high-quality service. It includes voltage regulators, speed governors, devices for recording the performance of the power system, protective relaying, system layout, and a sound design philosophy.

The specific weakness responsible for the initiation of the blackout was the acceptance of a form of backup protection that was unsound in its basic conception. It consisted, fundamentally, of impedance relays to provide remote backup protection on five 3-terminal lines, furnished with information from the local line terminal only.

Recognition of the vital relationship existing between system layout and protection relaying would never have permitted this form of protection in this situation. Either the 3-terminal lines should have been rejected, or a more sophisticated form of backup protection should have been provided. In either case, the disturbance that ended in the blackout would then never have occurred.

The additional weaknesses, which permitted the initial disturbance to evolve into a blackout, all involved the protection discipline in its widest connotations. In ascending order of importance there may be noted such factors as: limited energy storage in governor–oil pressure tanks and in air pressure tanks for pneumatic circuit breakers; direct extraction from the power system of power for station auxiliaries, without local energy storage to serve as a buffer in emergencies; the inherent limitations of the steam cycle; a power system that consisted of steam-electric vs. hydroelectric generation in the ratio of 3 to 1, whereas the reverse would have been more realistic; and lack of local energy storage for individual consumers in emergencies.

In the list of weaknesses, lack of hydroelectric generation and local energy storage loom large.

In the entire world there is probably no area so ideally endowed to solve the problems of reliable electricity supply as northeastern North America, including the "blackout" area. Here alone, there is available a potential of 150 000 MW of hydraulic power—about equally divided between run-of-the-river hydro power and tidal power in the Bay of Fundy. The necessary resources are thus available to solve the energy supply problems of the area for the near future. This provides a breathing spell to find alternatives to the inherently insoluble problems posed by the steam cycle, which tend today to make both fossil and nuclear fuels less than ideal primary sources of energy.

The wise course of action for the electric utility industry in northeastern North America is, therefore, to concentrate its major effort on a cooperative attack on the remaining undeveloped hydraulic power resources of the area, deferring the installation of more steam generation in the expectation that a technological breakthrough may make the steam cycle obsolete. Adding to the feasibility of this approach are the facts that hydro power does not pollute the environment, nor does it deplete irreplaceable natural resources.

Obviously, there is an energy storage problem involved in the development of hydro power. Water impounded by a dam is stored energy. In the case of streams that possess no storage basins, the storage of energy must be approached in a different manner. One such approach is pumped storage. . .

B. C. Hicks (SM)
La Compagnie d'Electricité Shawinigan
Montreal, Que., Canada

Protective devices should guard system rather than its components

The Northeast power failure of November 1965 was caused by a system design policy in which protective devices and relays were designed to protect components instead of protecting the system. Relaying, load shedding, and generator tripping should have been under the control of a system-state computer.

For example, a number of generators were shut down by overfrequency protective relaying. Consider a simplified case when four generators, operating in parallel, are supplying full load: If 25 percent of the load is lost, there will be an increase in speed and frequency of all four generators. With overspeed or overfrequency protection, the four generators will continue to increase in speed until the protective relays trip, shutting down all four machines and causing the loss of the remaining 75 percent of load.

The proper system protection that should have been used is a system-state evaluator. At the instant that 25 percent of the load is lost, there is an immediate positive rate of increase of frequency and a positive rate of increase of shaft speeds. Sensors, operating from each dN/dt, can deliver instantly the effect of load lost on each acceleration. From the known loads, inertias, and known mechanical inputs to the generators, a decision function would yield the command that one—and only one—generator should be shut down, if this would compensate for the load loss.

The overspeed or overfrequency relays would not be relied upon, because their information comes too late to save the system, and overfrequency is an average effect over many machines and a period of time. One generator only can be selected and disconnected to save the rest of the system by a decision based upon the system-state vector.

After line Q29BD was tripped out (the first Beck-to-Toronto line to be disconnected), the remaining four lines became overloaded. Overload relays are devices designed to protect a line even if it results in the loss of a system. No longer should individual line relays be used for primary decisions. A system-state evaluator would have recognized that the generation exceeded the load, and would have tried to reclose line Q29BD, reduce generation at Beck, and either increase generation or shed load north of Toronto. These actions could all be taken simultaneously, or preferentially, in sequence. An intelligent decision requires sensing and telemetering the state of the system, and action must be taken within a few cycles—too short a time to depend upon human comprehension of the situation.

The large power surges, following the loss of the five 230-kV lines from Beck No. 2 to the north, resulted in different amounts of acceleration of different generators and different parts of the system. One objective of transient control should be to adjust the relative accelerations so that all of the machine shafts have the same rate of increase of per unit speed. No longer can one depend upon synchronizing power flowing through transmission lines between machines to accomplish this result. Instead, those machines that have a low ratio of inertia to power should be instantaneously overexcited to minimize the phase advance, and those machines that have a high ratio of inertia to power should be instantaneously underexcited to maximize the phase advance. The excitation controls can be actuated from a comparison between: the shaft speed and dN/dt, and telemetered average speed and average dN/dt for the system as a whole. With such controls, the system can survive large power surges.

When the transients that appear at the terminals of a machine are such as to excite a shaft-angle oscillation, with energy oscillating between kinetic and magnetic, the oscillations can be completely removed by optimal decision functions based on machine-state variables and system average state variables.[1] The exciter should be controlled from speed and torque angle during transients. In some cases, this control does a better job of voltage regulation than if the voltage regulator were operative, because of delays in the voltage regulator–excitation dynamic system.

The shaft-state variables are "close" to the sources of energy flows and the causes of the voltage variations, and, therefore, during transients, the shaft-state variables are the superior control variable to use because of their phase lead ahead of other dependent, dynamically related effects such as voltage.

Where load shedding or generator tripping is called for by a system-state controller, the operation of the circuit breakers can either excite or quench serious transient oscillations. When the circuit breaker opens, a transient phasor is introduced into the low-frequency oscillatory systems, consisting mainly of inertias, magnetic fields, and the transmission line reactances. When a low-frequency oscillation already exists, the circuit breaker opening should be timed with respect to the phase of the low frequency so that the new contribution to the transient is as nearly as possible equal and opposite to the previously existing transient, thereby resulting in a sum that is almost zero—meaning that the oscillation has been quenched, and the oscillatory energies have been reduced to zero.

If two parts of a power system, connected by a line with reactance, are oscillating with respect to each other, the oscillation can be removed by a computer opening and closing the circuit breaker in the transmission line at properly chosen times with respect to the phase of the oscillation. Inserting a series capacitor can also produce desirable results.[2] The Northeast blackout has illustrated the necessity for computer control using the state variable or state vector techniques that have been so successful in satellite and rocket controls.

The important components of the system-state vector are the measurables most closely related to the stored energies and to the rates of change of energy. The shaft speed of each machine is the most important measurable, since it is proportional to the stored kinetic energy. The shaft phase angle is another important measurable, being related to the stored magnetic energy. Power flows and telemetered phase-angle differences on tramsission lines are also important state variables.

For system protection, a centralized computer can compare the existing state vector with a desired, realizable state, including the restrictions of components and apparatus that are nonoperative, and issue the sequence of decisions or commands necessary to achieve the desired realizable state. Automatic load shedding would, of course, be one aspect of the computer operation, but it also can control the timing of the opening and closing of circuit breakers to minimize the transient oscillations in phase angle and shaft speed of each machine. Generator shutdown, overspeed protection, fault clearing, and line protection should be under the supervisory control of this computer, which would be cognizant of the system implications of each decision and each action. Each action for component protection would automatically call for a simultaneous second compensatory action selected to maximize the system's usefulness or minimize the transient disturbance.

Otto J. M. Smith (F)
University of California
Berkeley, Calif.

REFERENCES

1. Smith, Otto J. M., "Optimal transient removal in a power system," *IEEE Trans. on Power Apparatus and Systems*, vol. PAS-84, no. 5, pp. 361–374, May 1965.
2. Kimbark, Edward W., "Improvement of system stability by switched series capacitors," *IEEE Trans. on Power Apparatus and Systems*, vol. PAS-85, no. 2, pp. 180–188, Feb. 1966.

Economy vs. service reliability in Sweden

The FPC report and the review in IEEE SPECTRUM are being studied with great interest in Sweden. Any operation engineer may easily imagine the despair of the operator on duty, feeling like a 20th century Hamlet:

"The grid is out of joint—O cursed spite:
That ever I was born to set it right."

The lesson must not be lost by any power pool utility. I am sure there are many things to learn, and a qualified professional discussion within IEEE is to be expected. Permit me to comment on one point:

The FPC report states that in a conflict between economic factors and service reliability, the latter should be given heavy weighting. This conflict is very common and the recommendation is not very definitive. Within the Swedish State Power Administration, we have for several years tried to make an "overall economy" calculation of this conflict by using our equipment failure and service outage statistics and an estimated service outage cost.

This cost, after some investigation, is set to 1 Swedish crown per kW and 2 Swedish crowns per kWh interrupted load. With the rate of exchange of 1 dollar = 5 Sw. cr., the cost of the Northeast blackout seems to be about $80 million. This is not so far from the figure mentioned by the FPC. Isn't that coincidence interesting?

Bertil Mattsson
Swedish State Power Administration
Vallingby, Sweden

A supplement to "Information for IEEE authors"

Under the title "Information for IEEE authors" (IEEE SPECTRUM, vol. 2, pp. 111–115, August 1965), suggestions and requirements for publication in any IEEE journal are described. The need for a good title and abstract is stressed.

After further study of information handling and retrieval, the Publications Board expanded and clarified the requirements for title, abstract, and references. The elements prepared according to the procedures outlined below serve two functions: they are valuable to the present reader and they provide an excellent basis for eventual machine retrieval. The cost of such preparation is reasonable, both in effort and in dollars.

The usefulness of technical articles in IEEE periodicals is greatly enhanced if their indexing is thorough and effective, and if their contents are summarized by a good abstract. In fact, the primary ingredients of good indexing and retrievability are (1) a title, (2) an abstract, and (3) references. From these primary ingredients are derived the needed indexing components—the subject headings or descriptive words or phrases that enable searchers to find the documents in which they are most interested. As our technology improves, we expect this indexing step to become more and more automatic through the use of computer-based techniques.

It will now be IEEE procedure to have the author of each technical article for publication in IEEE periodicals provide:

(1) a title—brief, clear and descriptive, typically less than ten words. (Alternatively, if a brief "clever" title is preferred, a straightforward descriptive subtitle shall be added. For example, "Fingers or Fists?"—"Binary vs. Decimal in Digital Computer Decimal Representation.")

(2) an abstract—succinct and indicative of subject areas concerned. When appropriate, specific informative content shall be included. Suggestions for the preparation of such abstracts are given below.

(3) references—a set of citations chosen appropriately to point to the prior literature that relates closely to the present article. (A bibliography, often having no direct reference to points in the article, is often useful.)

It is also a part of the new procedure for each IEEE editor to apply standards of quality to the title, abstract, and references, just as they are applied to the texts of the articles that he publishes.

Preparing abstracts

An abstract or précis of an article is a collection of statements that comprises the essential qualities of the larger article. A good abstract enhances the value of a technical article for everyone concerned—reader, author, and indexer. It contains the significant words that identify the important ideas, and meets the following criteria:

1. The abstract tells the prospective reader what the article is about, in language appropriate to the field.
2. The abstract states the principal results and conclusions of the article when they are of a nature that permits brief statement.
3. The abstract is no longer than necessary, typically 50 to 200 words; longer only if the complexity of the article demands it.
4. The abstract, together with the title, contains the information needed for indexing the article.
5. The abstract does not contain concepts or conclusions beyond those discussed or arrived at in the article itself.

As an aid in writing good abstracts, the following techniques are suggested:

1. Read over the article and list, in order, the particular topics that are discussed.
2. Review each topic in turn to see if a direct informative statement of the result achieved can be written. If this can be done, use an informative statement instead of the purely indicative topic listing.
3. Now compose a first sentence that will clearly establish the context and scope of the article.
4. The completed abstract must now be edited for style. Use definite statements instead of generalities. Use the shortest clear expression for each thought. Use language appropriate to the potential reader of the article. Eliminate obscure abbreviations or acronyms. Make style and grammar consistent throughout. Check to make sure that the abstract truly meets the five main criteria listed above.

Examples

Title: Laser Radar Tracks Airborne Targets

Abstract: An instrumentation tracker quickly shows position of high-speed targets with great precision. Range is determined by FM-CW. Azimuth and elevation angles are recorded on magnetic tape or read out into a computer.

This abstract is merely indicative of the work done and provides few technical details.

Title: A Precise Optical Instrumentation Radar

Abstract: The instrumentation tracker described provides real-time positional data on high-speed cooperative targets with precisions of ±1 m at ranges between 300 m and 10 km. Unambiguous range is determined by a precise digital FM-CW ranging technique at a rate of 15 per second. A target-mounted beacon and a narrow laser ranging beam permit measurement of target position to values much less than target dimension. Azimuth and elevation angles are read out by precision shaft angle encoders and recorded in binary form, along with range and time, on magnetic tape or directly into a real-time computer.

This abstract is both indicative of the work done and informative as to the engineering techniques and the limits of precision.

Special Conference Report

1966 IEEE Winter Power Meeting

The Third Winter Power Meeting of the IEEE, and the second to be conducted under the auspices of the Power Group of the Institute, was held in New York City from January 30 through February 4, 1966. All sessions and committee meetings took place at the headquarters hotel, the Statler Hilton.

More than 40 technical sessions were held during the week, encompassing a wide range of subjects of interest to engineers in the power industry and related fields. H. C. Barnes, American Electric Power Service Corporation, served as Technical Program Chairman. This article will be devoted to highlights of several of these sessions, as submitted by their respective presiding officers.

Application of probability methods

R. Billinton, University of Saskatchewan, described some of the important, basic aspects involved in the problems of distributed representation of the forced outage probability value in generating capacity reliability studies. He stressed the need for accumulating generating unit outage statistics by means of effective reporting procedures. Another paper submitted by Prof. Billinton provides a comprehensive bibliography of 96 sources dealing with the subject of the application of probability methods in the evaluation of generating capacity requirements.

"The Effects of Availability Upon Installed Reserve Requirements," by A. K. Falk, presented by B. H. Schneider, points out, through the use of probability theory, the value of utilities entering into pooling agreements, installing interconnections, and generally becoming a part of a large system. Because the utility engineer now has the computer as a calculating tool, the equations in this paper are practical not only for determining generating reserve requirements but also for determining unit sizes. However, the paper mathematically proves that as long as the trend to larger generating unit sizes continues, the practical lower limit for reserve or forced outage protection appears to be 8 or 9 percent of load.

As reported by H. Frank and S. L. Hakimi, a power system can be assumed to be a communication network where interconnecting lines are information channels and current flow is equivalent to information flow. Given a set of power stations and cities with current demands, the interconnecting (lossless) links may be unreliable. Techniques are available to compute the probability of satisfying the power demands and methods of increasing this probability. In reality, there is a cost associated with sending a flow through a given line. A synthesis procedure, based on linear programming, can be used in the design of a minimum cost distribution system which satisfies the given power demands.

L. L. Garver, General Electric Company, described a concept by means of which loss-of-load probability results may be interpreted in understandable megawatt terms. His paper also presents a graphical method for estimating the megawatt value of any generator at any outage rate. The information given should be useful in the planning of generating unit additions that maintain constant system reliability.

J. M. Arthur
Duquesne Light Company

Insulation measurements for HV stators

The use of insulation power-factor measurements for evaluation of the condition and quality of stator windings has been the subject of study for several years. E. H. Povey presented the recommendations of an IEEE working group for standardizing the method of measuring insulation power factor and interpreting the results. The report affirms that the change in insulation power factor as a function of test voltage may reflect the compactness of the insulation, the degree of cure of impregnating material, the extent of voids in the insulation structure, and the degree of deterioration of old coils.

R. G. A. Brearley, Aluminium Company of Canada Ltd., reported successful practical use of insulation power-factor measurements as a guide to the rewinding and repair of 22 large hydrogenerators. The theoretical problems and pitfalls associated with insulation power-factor measurements were described by E. M. Fort, Westinghouse Electric Corporation. The numerous discussions disclosed considerable differences of opinion regarding the merits of this method of evaluating rotating-machine windings.

B. L. McHenry, Ontario Hydro, explained the elaborate refinements of equipment and procedure developed by his company for the determination of winding serviceability by testing with high direct voltage. Automatic stepless linear increases of the output voltage from an extremely stable high-voltage electrostatic generator is applied to the thoroughly guarded winding while the current-voltage function is recorded on an *XY* recorder.

C. L. Sidway
Southern California Edison Co.

Power system economics

Five papers were presented at a session on economics sponsored by the Power System Engineering Committee. "A Parametric Equations Approach to the Economic Dispatch Problem," by R. W. Long and J. R. Barrios, introduced a new mathematical approach to the economic dispatch problem. Among its advantages, the method obtains a solution for plants having cost curves with horizontal segments, a situation that has led to convergence problems in the past. E. B. Dahlin of the IBM Corporation presented a paper entitled "Optimum Solution to the Hydro-Steam Dispatch Problem for Certain Practical Systems," coauthored by D. W. C. Shen of the University of Pennsylvania. This paper introduced a mathematical solution of the hydro-steam optimization problem through use of the "maximum principle."

C. W. Watchorn of Pennsylvania Power & Light Company presented his paper, "Inside Hydrothermal Coordination," in which he stresses the physical relationships underlying the various mathematical approaches. A widespread interest in the subject was evidenced by the large number of prepared discussions of the paper presented at the session.

R. Billinton and M. P. Musick are the authors of "Spinning Reserve Criteria in the Combined Manitoba System by the Application of Probability Mathematics," which introduces an analytical consideration of risk levels. A paper by H. Greber introduces a new concept by considering the reliability of electric service at the consumers' point of load in the system. The system's reliability might then be considered as the summation of the service reliability throughout the system. Mr. Greber also employed this approach for economic evaluation of optimum service reliability.

Einar Greve
Chas. T. Main, Inc.

Control response of universal pressure or once-through units

Universal pressure or "once-through" (sometimes called "drumless") boilers are appearing in increasing numbers, not only because of the accelerated trend toward 3500-psi steam pressure, which is above the "critical" pressure, and therefore requires units without drums, but also because of problems associated with long drums and circulation in large 2400-pound drum-type units. Many problems introduced by "drumless" units remain unsolved. Therefore, the progress developed through the three papers presented, and their discussions, is of timely interest to those in the the power generation field. These papers include "Considerations in the Regulation of Interconnected Areas" by Nathan Cohn, "Simulation of Bull Run Supercritical Generation Unit" by B. Littman and T. S. Chen, and "Control Response of Once-Through Steam Generation" by R. R. Walker and W. H. Baker.

Once-through units now in operation have been restricted in rate of response because of control instability, so that the contribution of such units to normal frequency-bias load control is quite small compared with drum-type units. This means that other units must bear a greater share of the regulation of generation. Certain emergency situations were discussed in which the performance of once-through units is dubious, the first involving the sudden loss of a large block of generation on a system, which would load interconnection ties instantaneously to something above the continuous current carrying capacity of the ties. Rapid load-control action is usually required so as to restore tie-line loading to something under its thermal rating within about five minutes. If this cannot be done, then interconnection ties might have to be opened, with rather serious results for the deficit area. The second emergency situation discussed involves the splitting off of a portion of a power system or area with an imbalance of load and generation, causing frequency to drop by one or two hertz when all the steam units would be expected to move to maximum output (almost instantaneously) to halt the decline in frequency and permit restoration of synchronized operation. Discussion developed that types of control available to date with once-through units probably would not permit this type of emergency assistance without danger of control instability and resultant loss of a unit. The third situation involves a rare catastrophe when an area could become bogged down and when it would be highly desirable for a unit operator to disconnect his turbogenerator from the high-tension transmission system by circuit breaker, leaving this unit to supply its own station power from the normal generator connections and thus be able to pick up load and supply station power to start another unit or plant. Although some feel that this can be accomplished, none of the users present had any units that are currently able to perform this emergency function. On the other hand, it can be accomplished readily with drum-type units.

It was the consensus that much more development work is needed, not only on controls but on basic elements of boiler-furnace design, so that drumless-type units can make their own proper contribution to reliability of bulk power supply.

W. R. Brownlee
Southern Services, Inc.

Utility communications techniques and systems

H. M. Matthies of Siemens & Halske, Germany, opened the session with a presentation of a paper by G. R. Bergmann and H. K. Pidszeck of the same company. The types of systems available, channel specifications, and methods of multiplexing various services on the company's latest design of power-line carrier equipment were outlined. The authors feel that power-line carrier has an important role in the control and operation of large interconnected power networks.

"Communication Circuits on 500-kV Lines," by T. M. Swingle and H. I. Dobson, describes results of extensive testing on the line and on the insulated shield wires under dead-line and energized conditions. Special testing also was done to learn more about the modal transmission of carrier signals. Data on radio-interference measurements of a preliminary nature were given and it was reported that no radio or television interference complaints have been received to date.

A. H. Ballard, Bernard Electronics Company, discussed a new multiplex technique for communication systems, in which orthogonal pulse waveforms are transmitted as simultaneous subcarriers. Advantages include high spectrum efficiency and rejection of crosstalk and noise.

"The Tennessee Valley Authority's Direct Distance Dialing Telephone System," by H. H. Howell, D. R. Jernigan, W. M. Foster, and L. H. Luallen, describes TVA's overall communication requirements and gives considerable details on the methods used to provide a suitable total telephone system. Special test facilities were also discussed, as were the plans for future growth.

"A Guide for the Protection of Wire Line Communications Facilities Serving Electric Power Stations" is the result of investigations conducted by the Wire Line Subcommittee of the IEEE Power System Communications Committee. The paper outlines the overvoltage problems that arise on the communication wire-line circuits during lightning and short circuits on the power system. The protective devices and practices recommended should assist the utility engineer in planning his communication system.

T. A. Cramer
General Electric Company

DC machinery

"The Design of Brush Gear for High Current Pulses and High Rubbing Velocities," by R. A. Marshall of the Australian National University, demonstrates the method used to design and develop brush gear for a machine that will produce a pulse of 1.6 million amperes. Results show that metal graphite brushes can be operated at peripheral velocities four to five times conventional values and at apparent current densities of more than 50 to 100 times conventional densities. In addition it was found that the brush can share load adequately without external balancing resistances, provided sufficient brush pressure is applied.

As reported in a paper by R. P. Pardee, Sandia Corporation, tests were conducted on silver-graphite brushes sliding against copper slip rings to determine moisture levels below which the brushes would dust. Brushes performed satisfactorily in CO_2 at the lowest attainable moisture concentration (9–10 parts per million), whereas critical moisture levels were found in air, nitrogen, and helium. The order of increasing moisture dependence to nondusting sliding was CO_2, air, helium, and nitrogen. Brush contact resistance varied inversely with coefficient of friction during nondusting operation in all four gases considered as a group. Theoretical equations relating resistance and friction were derived from fundamental concepts of contact resistance and solid and fluid friction.

The role of the carbon brush and its relationship to current collection were discussed in two papers by K. P. P. Pillai of the Engineering College, Kerala State, India. The first of these papers, which was coauthored by S. P. Roy Choudhuri, deals with analysis of the resistance of sliding electrical contacts. The second paper concentrates on the factors affecting static resistance of electrical contacts.

R. M. Dunaiski
General Electric Company

Induction machinery

Four papers were presented at a session devoted to unusual techniques or applications of induction motors. Two of the papers, by S. V. Ahamed, University of Colorado, cover experimental investigations of brushless single-unit frequency converters using unique designs of rotors. This technique permits the use of induction motors as frequency changers for frequencies both above and below line frequency. Although the use of induction motors for this purpose is not new, these papers present a fresh approach to the problem in the method of construction and method of increasing the output per unit size.

The other two papers relate to the use of induction motors in systems utilizing solid-state controls. One paper, by B. V. Jayawant and G. Williams of the University of Sussex, England, explains the analog and analysis of induction motors when powered by a solid-state static inverter. Although this paper raised considerable question and comment as to the assumptions made in establishing the analysis, moving pictures of the test data revealed supporting agreement with the analog results.

A paper by M. S. Erlicki and E. Livnant, Israel Institute of Technology, outlines the use of power transistors to control the secondary winding of a wound-rotor induction motor. This technique gives effective speed control, but it is limited by the rating of power transistors to only the smallest wound-rotor motors.

The trend of these papers indicates the necessity of considering not only the induction motor, but the system by which it will be powered or controlled, as a complete package.

J. C. Andreas
Century Electric Company

Excitation and energy development

Developments in a new and promising "fluid metal" magnetohydrodynamics approach in direct energy conversion were revealed by L. L. Prem, Atomics International. A small experimental MHD ac induction generator using liquid metal (NaK) accelerated through the generator channel has produced a three-phase ac power of 1840 watts at a line voltage of 215 volts. The induction-generator stator windings are in a flatbed arrangement on opposite sides of the channel. An external three-phase system energizes these windings, which in turn create a