VOLUME 6 □ NUMBER 4 IEEE spectrum APRIL 1969

IEEE spectrum

features

25 Spectral lines: Journals of the IEEE
The most profound change in the Institute's publication program has been the elevation of the Transactions of the various Groups from second-class status to prime journals in which specialists in the field communicate with their fellow specialists

26 Electricity and weather modification
I. A survey of the scientific relationships
Seymour Tilson
Rain, hail, snow, fog (and clouds of every other kind, including the pollution palls created by man), violent thunderstorms, and tornado vortices — even the crisp clear air — are characterized by distinctive electrical attributes that could provide the key to their control

46 Pumped storage — the handmaiden of nuclear power
Sherman R. Knapp
In addition to providing energy during peak periods, a pumped-hydro plant can be designed so that it can be brought on the line rapidly enough to be classed as a reliable source of spinning or assured reserve capacity at all times

53 The crystal resonator — a digital transducer
Donald L. Hammond, Albert Benjaminson
Though not digital devices in the purest sense, temperature and pressure transducers using quartz-crystal sensors convert information from analog physical parameters into the analog parameter of frequency, which in turn is measured by the techniques of digital counters

59 Art and technology I. Steps toward a new synergism
Nilo Lindgren
"After Picasso," quipped an art historian, "women and owls never looked the same." If artists and engineers who want to collaborate on artistic projects have their way, our whole environment may never look the same

69 The early history of electronics
III. Prehistory of radiotelegraphy
Charles Süsskind
Even before the turn of the century much of what is modern microwave engineering had already been anticipated in the devices and techniques developed by Hertz and his successors, and Crookes had proposed using electromagnetic waves for telegraphy

79 How to extend sampling-oscilloscope versatility
H. Allen Zimmerman
Many measurement problems can now be solved much more readily by means of a recently developed sampling-oscilloscope system based on the availability of some 20 different kinds of units, designed to operate in conjunction with each other

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

the cover

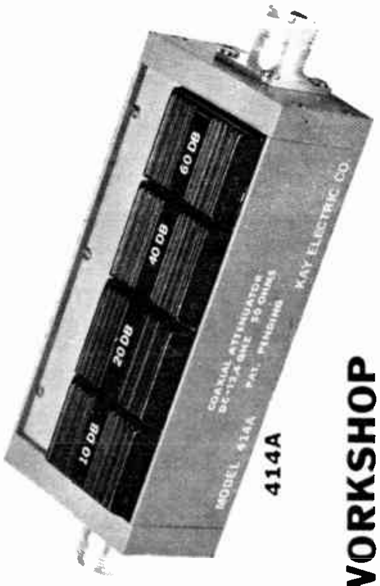
"Homage to New York," the machine that destroyed itself on March 17, 1960 at the Museum of Modern Art, New York, was the work of sculptor Jean Tinguely and engineer Billy Klüver. Not only did this work assault the traditional idea that all art must be lasting, but it also reasserted on older idea that artists and engineers might work together. See page 59.

departments

NEW

DC-12.4GHz COAXIAL STEP ATTENUATOR

from KAY'S ATTENUATOR WORKSHOP



414A

SPECIFICATIONS

ATTENUATION: 0 to 130 db in 10 db steps.
ATTENUATION STEPS: 10, 20, 40, and 60 db.
IMPEDANCE: 50
FREQUENCY RANGE: DC — 12.4 GHz.
CONTROLS: In-line Slide Switches, self locking.
ACCURACY: ± 1.0 db or ± 5% whichever is greater.
MAX. VSWR: 1.5:1
INSERTION LOSS: 1.0 db Max.
CONNECTORS: Type N to Mil C39012.
PRICE: $295.00

Kay Electric Company,
Maple Avenue, Pine Brook, N.J. 07058

Circle No. 3 on Reader Card

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

BOARD OF DIRECTORS, 1969

F. K. Willenbrock, *President*
J. V. N. Granger, *Vice President*
J. H. Mulligan, Jr., *Vice President Technical Activities*
M. E. Van Valkenburg, *Vice President, Publication Activities*
R. W. Sears, *Secretary*
Harold Chestnut, *Treasurer*
S. W. Herwald, *Junior Past President*
W. K. MacAdam, *Senior Past President*

G. J. Andrews
H. P. Bruncke
Werner Buchholz
E. E. David, Jr.
R. G. Elliott
F. A. Hawley
L. C. Hedrick
Hubert Heffner
D. M. Hodgin
D. G. Lampard
H. R. Mimno
D. C. Ports
C. F. Savage
George Sinclair
W. T. Sumerlin
R. H. Tanner
J. G. Truxal
R. P. Wellinger
J. R. Whinnery

A. N. Goldsmith, *Editor Emeritus and Director Emeritus*
Haraden Pratt, *Director Emeritus*
E. B. Robertson, *Director Emeritus*

HEADQUARTERS STAFF

Donald G. Fink, *General Manager*

John L. Callahan, *Staff Consultant*
Richard M. Emberson, *Director, Technical Services*
Ralph H. Flynn, *Director, Publishing Services*
Elwood K. Gannett, *Director, Editorial Services*
William J. Keyes, *Director, Administrative Services*
John M. Kinn, *Director, Educational Services*
Leon Podolsky, *Staff Consultant*
Betty J. Stillman, *Administrative Assistant to the General Manager*
Howard E. Tompkins, *Director, Information Services*

Committees and Staff Secretaries

Awards Board: Una B. Lennon
Educational Activities Board: J. M. Kinn
Fellow: Emily Sirjane
Finance: W. J. Keyes
History: W. R. Crone
Information Services: H. E. Tompkins
Information Systems Advisory: H. E. Tompkins
Internal Communications: Audrey L. van Dort
Intersociety Relations: J. M. Kinn
Life Member Fund: W. J. Keyes
Long Range Planning: W. J. Keyes
Membership and Transfers: Emily Sirjane
Nomination and Appointments: Emily Sirjane
Professional Relations: J. M. Kinn
Publications Board: E. K. Gannett
Sections: Emily Sirjane
Standards: J. J. Anderson
Student Branches: Emma White
Technical Activities Board: R. M. Emberson
Translated Journals: A. A. McKenzie

IEEE SPECTRUM is published monthly by The Institute of Electrical and Electronics Engineers, Inc. Headquarters address: 345 East 47 Street, New York, N.Y. Cable address: ITRIPLEE. Telephone: 212—752-6800. Published at 20th and Northampton Sts., Easton, Pa. 18042 and additional offices. **Change of address** can be made effective with any particular issue if received by IEEE Headquarters by the 10th of the preceding month. **Annual subscription:** IEEE members, first subscription $3.00 included in dues. Single copies $1.00. Prices for nonmember subscriptions available on request. **Editorial correspondence** should be addressed to IEEE SPECTRUM at IEEE Headquarters. **Advertising correspondence** should be addressed to IEEE Advertising Department, at IEEE Headquarters. Telephone: 212—752-6800.

Responsibility for the contents of papers published rests upon the authors, and not the IEEE or its members. All republication rights, including translations, are reserved by the IEEE. Abstracting is permitted with mention of source.

Second-class postage paid at Easton, Pa., and additional offices. Printed in U.S.A. Copyright © 1969 by The Institute of Electrical and Electronics Engineers, Inc. **IEEE spectrum** is a registered trademark owned by The Institute of Electrical and Electronics Engineers, Inc.

BPA

OTHER IEEE PUBLICATIONS: IEEE also publishes the PROCEEDINGS OF THE IEEE, the IEEE STUDENT JOURNAL, and more than 30 Transactions for IEEE Groups with specialized interests within the electrical and electronics field. Manuscripts for any IEEE publication should be sent to the editor of that publication whose name and address are shown on page 10A of the January 1969 issue. When in doubt, send the manuscript to E. K. Gannett, Editorial Services, at IEEE Headquarters, for forwarding to the correct party.

World Radio History

Arithmetic probability paper based on the Gaussian probability distribution is a standard tool of the statistician. This month Warren White discusses the use of a similar paper based on the Rayleigh distribution for the analysis of video data samples.

VIDEO PROBABILITY DISTRIBUTIONS

In many experimental programs it is desirable to plot the probability distribution functions of observed data samples. These functions often convey vital information on the process being observed.

Arithmetic probability paper based on the normal or Gaussian probability distribution is often used for plotting these distribution functions. Since Gaussian distributions plot as straight lines on such paper we have a ready visual test of the hypothesis that the sample observed does in fact come from a Gaussian population. Furthermore, even when the distribution is clearly non-Gaussian it often happens that the tails tend asymptotically toward straight lines. In any event, the scale expansion at both ends of the scale permits the tails of the distribution to be examined in detail.

The utility of Gaussian probability paper suggests that perhaps other graph paper designs might be useful when the underlying process is such that one expects another type of distribution for the standard or normal situation. A case in point is the case where the observations are of detected video voltage samples. If the input to a linear detector consists only of IF Gaussian noise, then the detector output will follow the envelope of the input noise and its distribution is given by the Rayleigh distribution. The probability density is

$$P(v) = \frac{v}{\sigma^2} \exp(-v^2/2\sigma^2) \qquad (1)$$

where σ is the rms value of the IF noise. This function integrates readily and the cumulative distribution is

$$P(v > v_0) = \int_{v_0}^{\infty} P(v)\,dv = \exp(-v_0^2/2\sigma^2) \qquad (2)$$

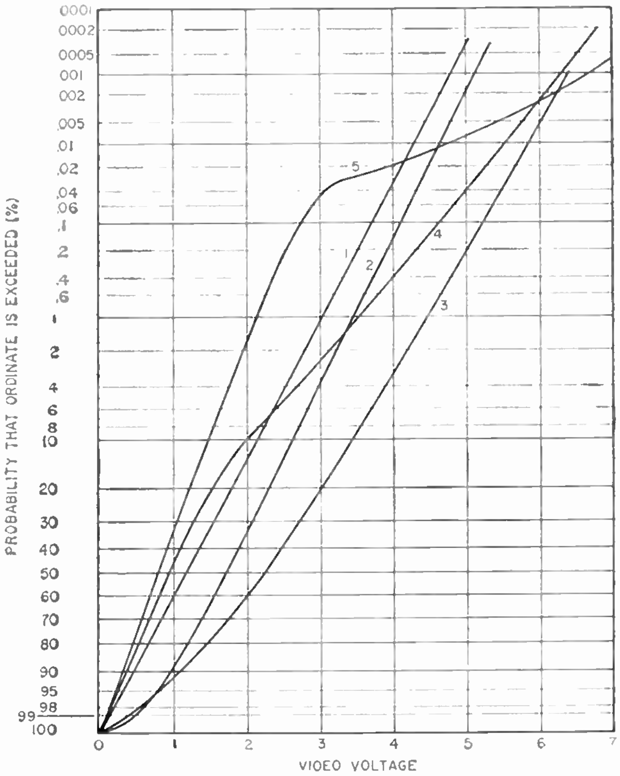
Now let us assume we wish to design a graph paper such that a distribution following equation (2) will plot as a straight line starting at the origin. That is to say we would like the plot to be of the form

$$Y = F(P) + mx = m(v_0/\sigma) \qquad (3)$$

We let $v_0 = \sigma x$ and substitute in (2) and then solve for x substituting the result in (3). We thus retain the required form of the vertical scale

$$Y = F(P) = m\sqrt{-2\ln(P)}$$

where m is an arbitrary constant that may be adjusted to give the desired size sheet for the required range of P.



The illustration shows several theoretical distributions plotted on this type of Rayleigh paper. Curve 1, which is straight, is the type of curve we should get if the input consists only of Gaussian noise and the circuits are behaving as expected. Curve 2, which is displaced somewhat to the right and has a definite curvature at the bottom, is the type of curve we get when the video bandwidth is not adequate to pass all the noise, although the DC component is passed unattenuated. Curve 3 shows the situation when the IF input includes a constant amplitude sine wave in addition to the noise. The signal-to-noise ratio of the case plotted is 3 dB. Higher signal-to-noise ratios produce greater displacements and more pronounced curvature at the bottom.

Curves 4 and 5, which have a reverse curvature, illustrate what can happen when the observed samples come from a mixed population. In curve 4 three quarters of the samples came from one Rayleigh population and the remainder from a population 6 dB stronger. Curve 5 is illustrative of the situation where a little impulse noise is mixed with Gaussian noise. In this case 0.1 percent of the samples came from a population 10 dB stronger than the rest.

The plots can also be used to detect equipment nonlinearities. A curve displaced to the left and then curving down to the origin might indicate that the bottom of the detector characteristic is square law rather than linear, while if all the curves break sharply upward at a particular voltage, saturation effects might be suspected.

Recently AIL had a project which involved the analysis of large quantities of recorded video data. We found the Rayleigh paper to be a definite aid in this analysis. A master sheet was prepared on a computer-controlled plotter and then printed on vellum stock. We are not in the graph paper business but we do have a limited supply of these sheets left over. They are suitable for reproduction and if you would like one or two for your own experiments drop a line to Harold Hechtman of our Public Relations Department.

A complete bound set of our twelfth series of articles is available on request. Write to Harold Hechtman at AIL for your set.

Technological Advances Reflected

FORTRAN with Engineering Applications

By DANIEL D. McCRACKEN, McCracken Associates, Inc.

"... the author skilfully weaves three volumes into one. The first, and least important, is the discussion of FORTRAN IV features; the second is a course in engineering applications.... Third, the book shows how one goes about solving a problem. In this sense, it's an introduction to programming. The book includes many exercises and solutions. Actual programs and results of programs are shown. The text is profusely illustrated..."
—*Computing Reviews*

This book presents the essentials of FORTRAN in a form that can be mastered in a few hours of careful reading and practice. It was written basically for the person who wants to get a quick grasp of the use of the computer in the solution of problems in engineering and related disciplines.

Partial Contents: CASE STUDIES. Resistance of a Wire: Program Basics. Resistance of a Circular Wire: Values of Variables; Integer and Real Constants. Tension in a Rope on a Pulley; Parentheses; Hierarchy of Arithmetic Operators. Radioactive Decay: FORTRAN Functions. A Venturi Flowmeter; The GO TO Statement; Program Execution. Column Design: The Logical IF Statement. The Mass-Spring System or the Series RLC Circuit; Flowcharts; The Arithmetic IF Statement. **1967 237 pages $6.95**

To order, circle #5 on Reader Service Card.

Basic Nuclear Electronics

By HAI HUNG CHIANG, Argonne National Laboratory.

Contents: Basic Electronic Devices and Circuits. Pulse Amplifiers and Single Channel Analyzers. Multichannel Pulse-Height Analyzers. Typical Nuclear and Electronic Instruments. **1969 In press**

To order, circle #6 on Reader Service Card

Music by Computers

Edited by HEINZ VON FOERSTER and JAMES W. BEAUCHAMP, both of the University of Illinois.

Contents: Sounds and Music. SYSTEMS AND PROGRAMS. Online Generation of Sound. A Computer System for Time-Variant Harmonic Analysis and Synthesis of Musical Tones. Some New Developments in Computer Generated Music. ALGORITHMS IN COMPOSITION. Some Compositional Techniques Involving the Use of Computers. Graphical Language for the Scores of Computer Generated Sounds. AESTHETICS. Infraudibles. Operations on Waveforms. Control of Consonance and Dissonance with Nonharmonic Overtones. The Problem of Imperfection in Computer Music.

LP records of computer music are included in the book.

1969 139 pages $14.95

To order, circle #7 on Reader Service Card.

Computerized Approximation and Synthesis of Networks

By JIRI VLACH, University of Illinois.

Contents: Signals and Spectra. Idealized Filters. Linear Networks and Their Mathematical Descriptions. Realizability Conditions for Linear One-Ports and Two-Ports. Characteristics of Linear Networks. Transformations. Mathematical Methods of Approximation and Mathematical Algorithms. Approximation of Amplitude Characteristics. Modified Chebychev Polynomials and Their Use for Approximation or Correction of Characteristics. Approximation of a Constant Group Delay. Approximation of Low-Frequency Phase, Phase-Delay, or Group-Delay Characteristics. Approximation of Time Responses. Some Synthesis Techniques. Synthesis of Networks from Given Approximations.

1969 Approx. 488 pages $14.95

To order, circle #8 on Reader Service Card.

Analytical Functions and Distributions in Physics and Engineering

By BERNARD W. ROOS, John Jay Hopkins Laboratorium for Pure and Applied Science, Gulf General Atomic, Inc.

Contents: Analytic Functions. Fourier Transforms. Causality, and Dispersion Relations. The Wiener-Hopf Technique. Boundary Value Problems for Sectionally Analytic Functions. Distributions. Applications in Neutron Transport Theory. Applications in Plasma Physics. Appendices.

1969 Approx. 528 pages $19.95

To order, circle #9 on Reader Service Card.

Solenoid Magnet Design

The Magnetic and Mechanical Aspects of Resistive and Superconducting Systems

By D. BRUCE MONTGOMERY, Massachusetts Institute of Technology.

Contents: Uniform Current Density Solenoids. Non-Uniform Current Density Solenoids. General Cooling Considerations and the Cooling of Uniform Current Density Coils. The Cooling of High Performance Magnets and Cooling Considerations for Non-Uniform Current Distributions. Magnetic Stresses. Superconducting Magnets. Pulse Magnets. Field Analysis. **1969 In press**

To order, circle #10 on Reader Service Card.

WILEY-INTERSCIENCE a division of JOHN WILEY & SONS, Inc.

In Canada: John Wiley & Sons Canada Ltd.

in New Wiley-Interscience Books

Introduction to Optimization Practice

By LUCAS PUN, Laboratoire D'Automatique et des ses Applications Spatiales, Toulouse, France.

Contents: Implications of Optimization Problems. Static Optimization Techniques. Extremum Seeking Methods. Dynamical Optimization Techniques. Dynamical Suboptimization Techniques. General References. Index.

1969 336 pages $14.95

To order, circle #11 on Reader Service Card.

Principles of Holography

By HOWARD M. SMITH, Eastman Kodak Company, Rochester, New York.

Contents: Historical Introduction. Basic Arrangements for Holography. General Theory of Plane Holograms. General Theory of Volume Holograms. Factors Affecting Image Resolution. Practical Considerations. Color Holography. Applications for Holography. Appendices.

1969 Approx. 256 pages $9.95

To order, circle #12 on Reader Service Card.

Applied Numerical Methods

By BRICE CARNAHAN, The University of Michigan; H. A. LUTHER, Texas A & M University; and JAMES O. WILKES, The University of Michigan.

Contents: Interpolation and Approximation. Numerical Integration. Solution of Equations. Matrices and Related Topics. Systems of Equations. The Approximation of the Solution of Ordinary Differential Equations. Approximation of the Solution of Partial Differential Equations. Statistical Methods.

1969 In press

To order, circle #13 on Reader Service Card.

Physics and Technology of Semiconductor Devices

By A. S. GROVE, University of California, Berkeley.

Contents: Introduction: The Planar Technology. SOLID-STATE TECHNOLOGY. Vapor-Phase Growth. Thermal Oxidation. Solid-State Diffusion. SEMICONDUCTORS AND SEMICONDUCTOR DEVICES. Elements of Semiconductor Physics. Semiconductors under Non-Equilibrium Conditions. *p-n* Junctions. Junction Transistors. Junction Field-Effect Transistors. SURFACE EFFECTS AND SURFACE-CONTROLLED DEVICES. Theory of Semiconductor Surfaces. Surface Effects on *p-n* Junctions. Surface Field-Effect Transistors. Properties of the Silicon-Silicon Dioxide System. Index.

1967 366 pages $12.95

To order, circle #14 on Reader Service Card.

Quantum Electronics

By AMNON YARIV, California Institute of Technology.

Partial Contents: Basic Theorems and Postulates of Quantum Mechanics. Some Solutions of the Time-Independent Schrodinger Equation. Matrix Formulation of Quantum Mechanics. Lattice Vibrations and Their Quantization. The Quantization of Electromagnetic Radiation. Polarization and Energy in Insulators. Dia- and Paramagnetism. Paramagnetic Resonance. Paramagnetism of Ions in Crystals. The Paramagnetic Maser.

1967 478 pages $15.95

To order, circle #15 on Reader Service Card.

Amplifying Devices and Low-Pass Amplifier Design

By E. M. CHERRY, Monash University; and D. E. HOOPER, Plessey Pacific, Pty. Ltd.

Partial Contents: Introduction to Amplifiers. Characteristics of Amplifying Devices I: General Introduction. Characteristics of Amplifying Devices II: Vacuum Tubes. Characteristics of Amplifying Devices III: Bipolar and Field-Effect Transistors. The Field-Effect Transistor. Elementary Amplifier Circuit Theory. Biasing Circuits. Simple Practical Amplifiers I: Small-Signal Dynamic Response. Simple Practical Amplifiers II: Noise. Simple Practical Amplifiers III: Distortion.

1968 1036 pages $29.95

To order, circle #16 on Reader Service Card.

Analog-to-Digital/Digital-to-Analog Conversion Techniques

By DAVID F. HOESCHELE, JR., General Electric Co.

Contents: Conversion Systems. Error Allocation Analysis. Conversion Logic. Switching of Analog Voltages. Digital-to-Analog Decoder Operation. Digital-to-Analog Decoder Design. Reference Voltage, V_R. Analog Voltage Comparators. Analog Multiplexing (Commutation, Subcommutations). Shaft Position to Digital Encoding. Analog-to-Digital Converter Types (For Electric Analog Signals). Analog-to-Digital Conversion of Electric Analog Signals.

1968 455 pages $15.95

To order, circle #17 on Reader Service Card.

605 Third Avenue, New York, N.Y. 10016

22 Worcester Road, Rexdale, Ontario

Forum

Concerned scientists

The statement that follows was received from the Union of Concerned Scientists, a faculty group formed at Massachusetts Institute of Technology because of disagreement with recent trends in the United States regarding the use or misuse of science and technology. The evening program on March 4 included a panel on employment opportunities for scientists and engineers outside the defense industry.

Misuse of scientific and technical knowledge presents a major threat to the existence of mankind. Through its actions in Vietnam our government has shaken our confidence in its ability to make wise and humane decisions. There is also disquieting evidence of an intention to enlarge further our immense destructive capability.

The response of the scientific community to these developments has been hopelessly fragmented. There is a small group that helps to conceive these policies, and a handful of eminent men who have tried but largely failed to stem the tide from within the government. The concerned majority has been on the sidelines and ineffective. We feel that it is no longer possible to remain uninvolved.

We therefore call on scientists and engineers at M.I.T., and throughout the country, to unite for concerted action and leadership: Action against dangers already unleashed and leadership towards a more responsible exploitation of scientific knowledge. With these ends in mind we propose:

1. To initiate a critical and continuing examination of governmental policy in areas where science and technology are of actual or potential significance.

2. To devise means for turning research applications away from the present emphasis on military technology towards the solution of pressing environmental and social problems.

3. To convey to our students the hope that they will devote themselves to bringing the benefits of science and technology to mankind, and to ask them to scrutinize the issues raised here before participating in the construction of destructive weapons systems.

4. To express our determined opposition to ill-advised and hazardous projects such as the ABM system, the enlargement of our nuclear arsenal, and the development of chemical and biological weapons.

5. To explore the feasibility of organizing scientists and engineers so that their desire for a more humane and civilized world be translated into effective political action.

As a first step towards reaching these objectives we ask our colleagues—faculty and students—to stop their research activity at M.I.T. on March 4 and to join us for a day devoted to examination of the present situation and its alternatives. On that day, we propose to engage in intensive public discussions and planning for future actions along the lines suggested above.

If you share our profound apprehension, and are seeking a mode of expression which is at once practical and symbolic, join us on March 4.

Why Miami Beach?

The 1968 International Symposium on Circuit Theory, sponsored by IEEE and the Circuit Theory Group and supported by the Office of Naval Research, was held December 4–6, 1968, at the Hilton Plaza Hotel in Miami Beach. It is my view that this symposium should not have been held there.

The U.S. is faced with tremendous social problems of poverty, racism, the decay of urban areas, and the deterioration of the quality of life. Amid these, there exist areas of great affluence and extravagance, of which Miami Beach is one of the pompous symbols. When Federal programs aimed at reducing poverty, at ameliorating the conditions of life for the disadvantaged, at combatting pollution to make life livable for all of us, and many other programs, including National Science Foundation support for basic research and curriculum development, are being drastically reduced for lack of funds, it is socially irresponsible to hold such an event as the Symposium on Circuit Theory amid the ostentation, extravagance, and excess of a Miami Beach. This is particularly true since the expenses of a great majority of participants were undoubtedly supported by Federal grants or contracts.

Nothing of scientific and engineering importance was achieved at Miami Beach that could not have been achieved at a fraction of the expense and in fuller measure (in view of the Miami Beach distractions) elsewhere, possibly on a university campus.

It is time for circuit theorists, in particular, and electrical engineers, in general, to become more sensitive to their social responsibilities, to examine critically their professional activities, and to ask themselves searching questions concerning the social value and social costs of all that they undertake.

Norman Balabanian
Syracuse University, Syracuse, N.Y.

The decision to hold the symposium in Miami Beach was my own. It was based on the desire of a number of members of the Circuit Theory Group to have the symposium in the East at a place where the climate is warm in December. Two universities were approached for possible conference sites. When it became apparent to me that one would not cooperate and the other did not have adequate housing facilities, I decided to hold the symposium in a hotel. The choice of hotel was influenced by the attractive offer that the Hilton Plaza made to the symposium. We had the use of any number of meeting rooms without charge and a special room rate was given to those who chose to stay at the hotel.

I agree fully with Professor Balabanian that we have social responsibility as well as professional responsibility. I hold, however, that social responsibility can be expressed and implemented in many different ways. The Circuit Theory Group, as a professional body, should urge its members to take part in alleviating the social ills. But as to the precise action one might take and as to the degree of participation, I think these are matters of individual conscience. Many of us, over the years, without much fanfare, have engaged or are engaging in extracurricular activities to help the disadvantaged. I am sure more could and should be done and the Groups of IEEE could play an important role in solving some of the social problems. In fact, the Symposium on Circuit Theory featured a panel discussion on long-range research goals and several speakers, including Professor Balabanian, directed our attention to our social responsibility.

Returning to the point of the letter, I am not sure that holding the symposium at Miami Beach had lessened our sensitivity to our social responsibility and I doubt that holding the symposium at a university would increase it. As to the contention that it would cost less to hold the symposium at a university, this is simply not true. Except for the social hour, which was self-supporting, the symposium paid the hotel nothing. On the other hand, many universities would charge the symposium for the use of meeting rooms.

Omar Wing
Chairman, 1968 IEEE International Symposium on Circuit Theory

Urban communications

The article by Eckert and Kelly in the January issue should be of interest to many groups throughout the U.S. working on problems of communication systems in the urban environment. It is unfortunate, however, that the view of the authors is so narrow that the article is misleading in one significant respect. In the early part basic assumptions of the study are outlined, which, to me, appear to be justified and reasonable. Having been exposed to studies of this type during a recent "tour of duty" with the U.S. Department of Housing and Urban Development (HUD), I believe the general constraints placed on the study are appropriate. However, I feel that an alternative view of the overall problem, which also satisfies the stated constraints, should be used. This is the basic question of what type of communication is required for each task.

It has been observed in many studies conducted by HUD, the Department of Justice, the Department of Transportation, and others, that digital communication is very often not only a possible alternative to voice communication but an efficient, and in many ways, desirable one. In such cases, the spectrum conservation advantages of digital communication may be used to relieve congestion on the hand mobile communications bands.

Of course, going to digital communication is not without its problems. Standard mobile radio equipment would have to be modified. An educational program would be needed to "retrain" policemen, firemen, taxi drivers, and others in the use of the new equipment. However, digital systems, used appropriately, are ultimately easier to use and provide more efficient frequency utilization for many applications. It is worth the trouble to investigate this alternative.

Stephen J. Kahne
University of Minnesota
Minneapolis, Minn.

Engineering education

I agree with the usefulness of an honors program in the undergraduate engineering curriculum as expressed by J. O. Kopplin in "Stimulating Change in Engineering Education," in the January issue of IEEE SPECTRUM. I cannot agree, however, with Mr. Kopplin as to the need for increased liberalization in the engineering education. Mr. Kopplin says that engineers need to be "humanized," that their education should be based on the "needs of mankind today." I believe that the current engineering student is already quite "humanized," and that further liberalization of his studies is not only unnecessary, but a threat to the student in that the time could be more usefully expended toward scientific study.

Mr. Kopplin feels that in order to cultivate the student's responsibility to society, he should be barraged with nonscientific or "humanities" courses. In my mind no more than one sixth of the undergraduate's credit hours need be spent on nontechnical courses. (Such is the case at The Cooper Union, at which I am a student. Each semester of the four-year curriculum includes a three-credit humanities course, which, by the way, I have so far found to be quite stimulating and useful, as well as sufficient.) If care is taken to insure the worth of these courses to the engineer, the time will be well spent. It can be seen that by design the course should be general rather than specific: it should tie in *all* human experience to the mainstream of life, so that the student may gain a perspective not only of his usefulness to society, but also of everything's proper place in the complete makeup of a healthy society. The specific course may very well fail to make its usefulness known to the student, who may, therefore, make no effort to study the "useless" principles. Indeed, Mr. Kopplin's statement is valid, that "an engineer without a liberal education is a mere technician unaware of the cultural and social forces of his age." The engineer, however, need not plunge into a multitude of nonscientific courses in order to learn how to work creatively and thoughtfully with a mind toward his responsibility to the society.

Howard R. Krauss, Bronx, N.Y.

I wish to state quite clearly that it is not my belief nor did I imply that ". . . in order to cultivate the student's responsibility to society, he should be barraged with nonscientific or 'humanities' courses." I did state that a pluralism of educational patterns is needed, that new methods must be sought, that exploration is required, and that a pilot operation can be an effective way of achieving those objectives. Another point apparently missed by Mr. Krauss is that the Honors Program never encourages students. It does encourage individuals.

Mr. Krauss should have noted the examples I used. The Honors Program introduced human problems to engineering courses and sponsored conferences that considered engineering and social issues. The program is currently offering a course entitled "Technology and Urban Problems" open to students in engineering, urban planning, political science, and other disciplines. One of the objectives of the courses is to consider the effects of transportation systems upon people in cities. In some cases more social science and humanities courses may be beneficial, but they are not *the* answer to the problem. I fully agree with Mr. Krauss's concluding statement.

J. O. Kopplin, Professor and Head
The University of Texas at El Passo

Focal points

Experiments indicate that learned behavior controls responses of the nervous system

Experimental demonstrations have shown that many of life's vital processes can be speeded up or slowed down by rewarding changes in them, thus upsetting the widely held belief that the autonomic nervous system, which controls these processes, can learn only through classical conditioning.

In the January 31 issue of *Science*, journal of the American Association for the Advancement of Science, Dr. Neal E. Miller of The Rockefeller University reports research in which he, Dr. L. V. DiCara, and other colleagues, have trained animals to change visceral responses such as salivation, heart rate, intestinal contractions, kidney function, stomach activity, peripheral blood flow, and blood pressure.

For example, in a typical experiment an automatic record was made of tiny drops of saliva secreted by a dog. Whenever the time between two drops was a little shorter, the automatic apparatus immediately rewarded the thirsty dog with water. By rewarding first the small, spontaneous increases in salivation, and then requiring the dog to make larger ones before earning his reward, one group of dogs learned to salivate a great deal. By rewarding decreases in salivation, another group of dogs was taught to stop salivation. The same type of training but with different rewards was used to train the animals to change other visceral responses.

Pavlov, the Russian physiologist, used a training procedure called classical conditioning. If a bell is always sounded before meat is given to a hungry dog, the salivation elicited by the meat is conditioned by the bell. In Pavlovian conditioning, one must use a reinforcement that already elicits the response to be learned.

Prior to the Miller experiments, most scientists believed that classical conditioning was a primitive and automatic process different from the trial-and-error learning responsible for so-called voluntary behavior. They also believed that the vital life processes could be changed only by primitive classical conditioning.

Dr. Miller's theory is that classical conditioning and trial-and-error learning involve the same mechanism. The experiments help to demolish the wall between classical conditioning and ordinary learning and between the control of the internal organs and the skeletal muscles. Dr. Miller believes that his findings should affect four major areas: current theories of learning; individual differences in autonomic responses (those governed by the "involuntary centers" of the brain); the cause and cure of abnormal psychosomatic symptoms; and a better understanding of homeostasis—the state of equilibrium in the living body with respect to various functions and to chemical compositions of the fluids of tissues.

Thermoelectric generator features use of heat pipe

A thermoelectric power generation system, developed for the California Institute of Technology's Jet Propulsion Laboratory by the Westinghouse Astronuclear Laboratory, is said to have the highest efficiency of any such device built in the past and the lowest development cost in terms of watts per dollar.

Output of the unit is about 120 watts, 3.5 volts, at 34 amperes. Its efficiency is nearly 6 percent. The generator consists of a heat source, heat pipe, thermoelectric module, and planar radiator. For testing purposes the unit is operated with electric heaters instead of a radioisotope as the heat source.

Thermal energy produced by the heat source is transferred to the interior of a tubular thermoelectric generator by the heat pipe. The interior of the generator is the hot side of the thermoelectric couples; the radiator is the cold side. The heat pipe is highly efficient; it transfers almost 100 percent of the heat from its evaporator section at one end to the condenser section at the other.

Design of the tubular thermoelectric module is the result of eight years of development and testing. Total testing time on the developmental modules, which can operate in air, inert gas, or submerged in liquids, exceeds 180 000 hours. The monolithic modules can resist impacts as great as 10 000 *g* and are not sensitive to thermal cycling or nominal overheating.

Papers are needed for wire and cable symposium

Technical papers in the field of communication and electronic wire and cable are being solicited for presentation at the 18th International Wire and Cable Symposium, to be held in Atlantic City, N.J., December 3–5. Papers are requested in the following general categories: cable design and application; cable materials (conductors and insulations); process, manufacture, and quality control; testing and reliability; cable interconnection and connective devices; special cable requirements for advance equipment; and performance and field experience.

In its written form a technical paper may be of any length desired by the author; however, a time of 30 minutes is allotted to each paper on the program, of which 15–20 minutes should be allowed for the oral presentation and the balance for a question-and-answer period. Prospective authors must submit ten copies of a summary consisting of at least 500 words, in sufficient detail for evaluation of the paper, no later than May 10. Notification of acceptance will be made by June 15, and either the final manuscript or 1500 copies of the printed paper will be required no later than October 15, 1969.

Copies of summaries and papers, as well as requests for additional information, should be sent to Jack Spergel, Co-Chairman, Wire and Cable Symposium, U.S. Army Electronics Command, Attn: AMSEL-KL-EE, Fort Monmouth, N.J. 07703.

Changes announced in temperature scale, volt

A new International Practical Temperature Scale extending down to much lower temperatures than before has been approved by the International Committee of Weights and Measures. Among other actions taken at its October 1968 meeting in Sevres, France, the committee also recommended a change in the reference base of the volt.

The new International Practical Temperature Scale of 1968 (IPTS-68) is defined in terms of a series of fixed points and interpolation formulas for specified instruments used to realize the scale. It differs from the scale of 1948 in the following ways: (1) The new scale extends down to 13.8°K, whereas IPTS-48 terminated at 90.18°K. (2) New values have been assigned to most of the defining fixed points and to the second radiation constant, C_2, in order to approximate more closely the thermodynamic temperature scale. It is hoped that the National Bureau of Standards will be able both to calibrate thermometers on the new scale and to make adjustments for prior calibrations soon after publication of the text in the April issue of *Metrologia*.

The committee's action on the volt authorized the International Bureau of Weights and Measures to decrease the volt, as maintained, by 11 parts per million. This brings the "practical" volt into much better agreement with the theoretically defined volt, which is derived from the basic mechanical units of length, mass, and time. The change, effective January 1, is to be accompanied by appropriate changes at the various national laboratories throughout the world. As a result, the measured value of a truly constant voltage will increase, as the unit is being made smaller.

The change in the U.S. legal volt, as maintained by a group of standard cells at NBS, will be 8.4 parts per million, because the NBS volt and the BIPM volt have drifted apart by 2.6 parts per million since the last adjustment. In calibration reports issued by NBS during 1969, values will be given on the new basis. Values on the old basis will be included on a supplemental form so that the user will be advised directly of the change.

Leakage current limits for appliances specified

Underwriters' Laboratories, Inc., acting for the United States of America Standards Institute, has released its findings on maximum limits for household appliance leakage current to which users may be safely exposed in normal use. The recommendation is that such leakage current shall not exceed 0.5 milliampere, and applies to appliances having two-wire portable cord connection, rated for use on 120-volt circuits.

To determine the maximum level of leakage current, a program of measurements and tests was devised which included having a subject make a series of motions with a full cup of rice while touching sources of varying leakage current. If the current was sufficient to create an involuntary reaction by the subject, and cause rice to be spilled, this was a factor in determining if that current was excessive. Unless leakage current is controlled, the user of an electric appliance is exposed to potential hazard when touching the appliance and a grounded surface.

The U.S. Public Health Service and some 18 technical and trade associations participated in the study, which was conducted by UL during 1968 on assignment from USASI Committee C-101.

Electronic system may duplicate human brain

Three University of Wisconsin electrical engineers hope to develop an electronic system that can function much like the gray matter of the human brain. The basis of the proposal by A. C. Scott, R. D. Parmentier, and J. E. Nordman is a device called a superconductive tunnel junction neuristor (a neuristor is an electronic device that propagates electric impulses much as a nerve cell does).

The Wisconsin neuristor is a long, narrow sandwich of insulating material between two thin layers of metal. When the device is cooled to the temperature of liquid helium, the metal layers offer almost no resistance to electric impulses. If an electric impulse is applied at one end of the sandwich, it is propagated to the other end much like an impulse in a nerve cell. With this device, it may be possible to produce a mass of neuristors with a packing density that approaches that of the human brain.

The inventors feel that their system has the possibilities of overcoming three major problems: cost, power consumption, and plan of connection. The cost of ten billion nerve cells, the number in the human brain, at as little as one cent per neuron, would be $100 million. By using techniques employed in production of microcircuits, however, the engineers hope to produce the tiny electronic nerves in great quantities at tremendously lowered cost.

An electronic system the size of the human brain, with even small power requirements per nerve, would require as much power as is needed to keep an electric oven going full blast—10 000 watts. The power used in the resting state by the Wisconsin cryogenic neuristor is zero; only when the circuits are actually in use is power required.

The problem of connecting ten billion neurons would be solved by constructing the mass of neurons with bits of ironlike material between each neuristor. As information is fed into the system, certain preferential paths through the neuristors are established. Each time a path is used, the bits of magnetic material along the route become more and more magnetized, thereby allowing current to pass with less resistance. Thus the mass of neuristors can be "taught," much as a brain is taught to prefer certain paths over others.

However, as the Wisconsin scientists point out, the leap from a single neuristor to a mass of interconnected neuristors, and then to an artificial brain, is a tremendous one. It is felt that a more immediate use for the electronic gray matter is as a pattern-recognition system for computers. Patterns of light focused on a mass of light-sensitive

VOLUME 6 □ NUMBER 4 IEEE spectrum APRIL 1969

Spectral lines

Journals of the IEEE. In 1968, journal publication by the Institute totalled 23 759 pages, not counting 9979 pages of Russian, Japanese, and Chinese journals republished in English translation, and not counting advertising. By contrast, the American Chemical Society published 40 225 journal pages, and the American Institute of Physics published 66 100. Pages per member reckon out respectively to 0.15, 0.34, and 1.47. These figures are not cause for surprise, because it is known that scientists lean more heavily on archival publication than engineers do.

Now that the transients caused by the AIEE–IRE merger have pretty well died down, it seems worth while to look over the journal situation in its present quasi-steady state. With some exceptions, the TRANSACTIONS of the Groups, of which there are 31, used to be the outlet for technical articles that were not quite good enough for the PROCEEDINGS. The most profound change in the Institute's publication program has been the elevation of the TRANSACTIONS from their second-class status. They are now seen as the prime journals in which specialists communicate with their fellow specialists.

Maturing of the TRANSACTIONS has altered the character of the PROCEEDINGS, and has checked the growth of that journal, in spite of the large Special Issues. Its slimness in some months might be taken as a sign that the PROCEEDINGS is shrinking. In 1968, however, the PROCEEDINGS totaled 2300 pages of articles and letters, as against 1537 pages ten years earlier. In overall size, there has indeed been some decrease, because the news pages and much of the advertising have been transferred to SPECTRUM.

The Publications Board has repeatedly questioned whether there is still a real role for the PROCEEDINGS to fill. Every time, the answer has been "Yes." The nonspecial issues are the vehicle for three kinds of publication: papers that have interest beyond the area covered by any one TRANSACTIONS, the brief and quickly published Letters, and tutorial or review papers. These last are perhaps the kind of journal material most urgently needed by the membership. Unfortunately, they are the hardest to procure. A visitor from another planet might expect them to flow copiously from teachers in our universities. However, in the past 30 years the groves of Academia have turned into jungles, whose denizens have learned that in the struggle for existence, the writing of review articles has little survival value. It does have value for the development and consolidating of expertise, so we are doing pretty well with review articles by authors in industrial and nonprofit laboratories.

Of even greater tutorial worth than a single article is a well-coordinated Special Issue. Whether judged by their technical quality or by the value placed on them by the subscribers, these issues rank as one of the most successful of the Institute's publication activities.

The upgrading of the TRANSACTIONS has not been the only cause of metamorphosis of the PROCEEDINGS. Another factor has been the birth and growth of IEEE SPECTRUM. The purposes of SPECTRUM, numerous but not vague, were laid down at its inception: they have lately been reexamined by the Publications Board, which concluded that they needed no marked change. One attempt to improve their implementation is the new policy of soliciting and publishing how-to-apply-it articles, in the hope of providing useful, idea-generating material for the practicing engineer.

It seems to be widely believed that all articles published in IEEE SPECTRUM have been written by invitation. That is by no means true. We do solicit specific articles from specific authors, but some manuscripts are just sent in by authors who feel that they have something to say. A considerable fraction of these are accepted, and we emphatically encourage the submission of more manuscripts in this category.

A more bothersome and equally fallacious belief is that articles in SPECTRUM bear the ideological or political cachet of the Institute, and can be taken as indicating "the policy of the IEEE." That notion is quite without foundation. In this column, over his name, the Editor may properly make comment that is colored by his political or ideological bias, but he is expressing his own opinion, not that of the Institute, and if he permits his bias to affect his selection of material to be published, he is not doing a good job.

Actually, except for the governance of internal affairs, "the policy of the IEEE" does not exist. Decisions or recommendations are those of specific boards or committees, and are to be ascribed to those bodies rather than to the Institute as a whole. On this subject, President Willenbrock recently wrote, "I hope I never see a statement which says 'The IEEE believes . . . .' " and he has said that in his view, even the articles prepared for SPECTRUM by the staff writers are not to be construed as carrying any endorsement by the officers of the Institute.

Far from being a pipeline for IEEE policy, SPECTRUM should be a forum for responsible debate on the host of social, political, and economic problems associated with our profession. The editorial willingness is here. Where are the manuscripts? *J. J. G. McCue*

Electricity and weather modification

I. A survey of scientific relationships

Manipulation of the electrical properties of clouds may someday provide the long-sought key to modification of the weather by man. The scientific problems aren't simple, as this first installment makes clear, but they are easy compared with the social ones to be discussed in future installments

Seymour Tilson *Staff Writer*

Ethereal, evanescent, elusive—these are adjectives a poet might use to describe the fragile stuff of clouds. They are accurate, and evocative, but still another adjective—electrified—may prove to be more important in the years immediately ahead. Workers in the fields of atmospheric electricity and cloud physics have accumulated sufficient evidence to suggest that electric fields, forces, and charges in the earth's lower atmosphere play a critical role—perhaps *the* critical role—in the development and behavior of clouds that produce precipitation. This in turn suggests that manipulation of the electrical properties of clouds may someday provide the long-sought key to modification of the weather by man.

J. Doyne Sartor of the National Center for Atmospheric Research at Boulder, Colo., is a leading investigator of electricity's role in the particle collision and coalescence processes within clouds (Fig. 1), whereby droplets too small to fall are welded together into particles large enough to reach the ground. He states categorically that if we could produce and control the same charges and electric fields in the atmosphere that we can in the laboratory, we could drastically influence precipitation. Precipitation might be inhibited by injecting some ionizing agent into growing convective clouds to decrease the electric field strength and the number of free charges within them; precipitation might be increased, on the other hand, by introducing some unipolarly charged aerosol in such a way as to increase electric fields drastically. Essentially this latter possibility is being explored by a group at the New Mexico Institute of Mining and Technology at Socorro. Their effort is in turn inspired by a convective theory for thunderstorm electrification and precipitation proposed several years ago by Bernard Vonnegut of the State University of New York at Albany.

Vonnegut suggested that convectively rising columns of warm air, which support the growth of the cloud, might at the same time carry positive ions—derived from point discharge at the earth's surface—into the base of the growing cloud. A rising column of air might be expected to carry charge into the cloud in much the same way that the belt on a vertical Van de Graaff generator carries charge to a storage sphere or hemisphere at its upper end. Such ions might build up the cloud's field and also help form rain.

In order to investigate this hypothesis and other aspects of cloud electrification, the New Mexico group has built a large vertical wind tunnel, in which they have placed a high-voltage ionizer capable of generating a high concentration of positive ions. The group hopes to inject these ions into the base of a growing convective cloud through a 1000-foot (305-meter)-long polyethylene duct attached to the apparatus (Fig. 2). The idea is to supplement the natural updraft beneath the cloud with a forced convection of artificially charged air, to test the possibility that the electric field generated by these ions will cause cloud droplets to coalesce.

Fantastic? Not at all. The New Mexico experiment is only one of several electrical weather-modification experiments that have been carried out or are now in progress under federal sponsorship.[1-5] But neither this nor any other electrical weather-modification possibility is soon likely to become an operational reality. Not only are the engineering problems formidable, as both Sartor and Vonnegut are quick to acknowledge, but we still have much to learn about how nature produces both clouds and precipitation.

Nature first must set the stage

The first and most important requirement for precipitation is a favorable large-scale environment. Only natural perturbations in the general circulation of the atmosphere can provide this. Convection leading to the development of clouds and precipitation usually occurs under either of two sets of circumstances: (1) when a warmer air mass is lifted by another, colder air mass along the frontal surface that separates the two; or (2) when warm, moist air rises to clear a mountain barrier. Condensation of water vapor and cloud formation are the usual results as the rising air cools, provided that its normal decrease of temperature with height is not reversed by strong temperature inversions at low or at intermediate

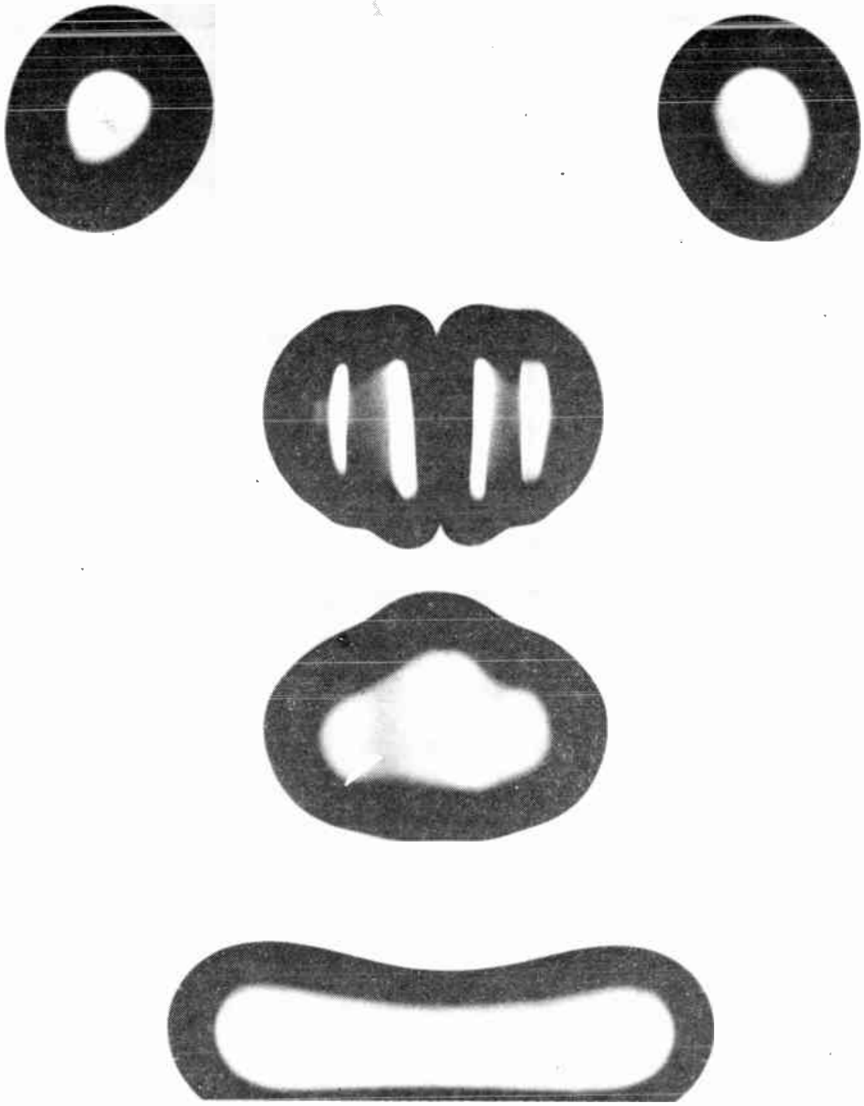


FIGURE 1. Water droplets that first condense from the vapor state to form clouds in nature are only a few micrometers in diameter, far too small to fall as rain. Collisions like the one you see taking place in this laboratory experiment help them to coalesce into larger drops that can reach the ground. In this experiment, one of a continuing series being carried out under the direction of J. Doyne Sartor at the National Center for Atmospheric Research, two streams of water droplets, each 780 μm in diameter and separated from the next by a time interval of 50–100 μs, were projected toward each other from the upper left and upper right. Only the leading droplets in each stream are visible in this photograph. Successive droplet pairs collided and coalesced with monotonous regularity—as you can see—because droplet trajectories, inertias, and interaction times could be manipulated mechanically to achieve this result. To achieve comparable results in clouds, however, and thereby to produce or prevent rain, it may be necessary to manipulate the charges carried by cloud droplets, or the electric field within the cloud, because electrostatic attractive forces contribute significantly to most collisions that lead to coalescence in the early stages of precipitation particle growth. This alternative to conventional cloud seeding has been explored by several groups in the past few years.

heights, where the temperature is not yet low enough for condensation or freezing to take place. Convective clouds like the cottony fair-weather cumulus, however, whose familiar shapes bring welcome shade on warm summer afternoons, often develop within a single air mass rather than along the frontal boundary between two air masses or along a mountain belt. The rising air that supports their growth to the cumulonimbus stage is provided by nonuniform local heating of the earth's surface, though conditions aloft must be favorable as well, if growth is to continue.

However it is initiated, once launched into a favorable environment by any of these large-scale processes, a cloud is on its own. And, in fact, most clouds sustain themselves quite nicely. They improve on the dynamics of their environment by releasing latent heat to the air at strategically distributed places and times, as the water vapor that feeds their growth condenses, or as the water droplets that comprise them freeze. Clouds also abstract heat from the air, of course, when water changes phase in the opposite direction. Clouds are often so successful at carrying out this bootstrap operation that in advanced stages of cloud development—especially in those massive storms whose clouds yield significant amounts of precipitation—the clouds actually develop a quasi-steady, local circulation that contributes to their longevity.

All clouds, quiescent ones as well as stormy ones, die; however, they are reborn without end. Differential heating of the earth's surface by the sun creates air masses. These, driven ceaselessly across the face of the globe by thermally induced pressure gradients, carry water vapor, gained mostly from the sea, and condensation and freezing nuclei, gained from both land and sea. As the temperature changes, the water in the air utilizes these nuclei as the substrate upon which to change, repeatedly—from vapor to liquid to solid, and back again—to form and dissipate clouds whose variety is exceeded only by their beauty. Embedded, somehow, in these apparently elementary phase changes, and in the convective processes that carry air to altitudes where temperatures become low

FIGURE 2. Most modern attempts to induce precipitation from clouds are based on injecting them with ice-nucleating substances such as silver iodide or dry ice. The purpose is to convert the relatively large number of small, supercooled water droplets they contain into fewer, larger, and heavier ice crystals or snowflakes; these will then fall and grow still larger as they sweep up other droplets on their way to the ground. The experiment shown here, however, is exploring an alternative possibility, that altering the electric field in clouds will cause droplets to coalesce into drops large enough to fall as rain. The business end of this device, being developed at the New Mexico Institute of Mining and Technology, under the direction of Stirling Colgate, is the structure at ground level. It houses a vertical wind tunnel within which a high-voltage ion generator has been placed. The idea is to inject the large ions produced by the ionizer into the base of growing convective clouds, through a 305-meter-long, balloon-supported polyethylene duct, only a short test section of which was mounted at the time this photograph was taken.

enough for the contained water vapor to condense or freeze, are the generating mechanisms most immediately responsible for thunderstorm and tornado electricity. The same processes may be responsible for the formation and fall of rain, hail, and snow—all of which usually carry a weak electric charge.

Electrical control of coalescence

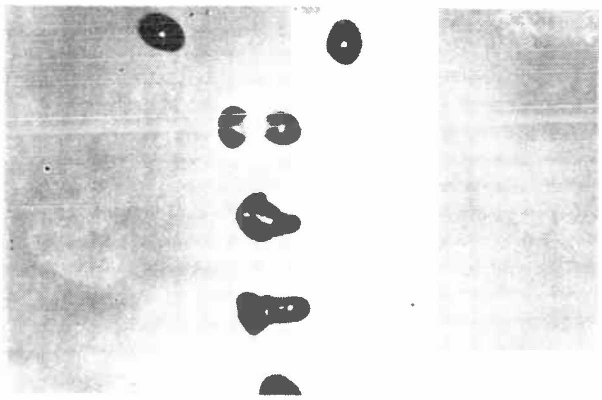
When water vapor condenses on aerosol particles in the atmosphere to form the first droplets that characterize clouds at above-freezing temperatures, the droplets are typically only a few micrometers in diameter, much too small to fall to the ground. At first, the droplets grow rapidly by continued condensation of water vapor from the surrounding air. Further growth occurs when winds and gravitational forces cause these droplets to collide and coalesce into larger, heavier drops. These larger drops can then fall, sweeping through smaller droplets, gathering up some, and growing thereby until they are large enough to fall all the way to the ground as rain. This accretion and fallout process is not restricted to warm clouds—those whose temperature is above freezing. It also contributes to the growth of precipitation particles in colder clouds. The snow and ice that fall from subfreezing clouds, or from the subfreezing portions of clouds transected by the 0°C isothermal surface, grow by accretion after freezing occurs. And their crystals form from supercooled water droplets whose growth is subject to the same collision and coalescence events that control droplet growth in warm clouds.

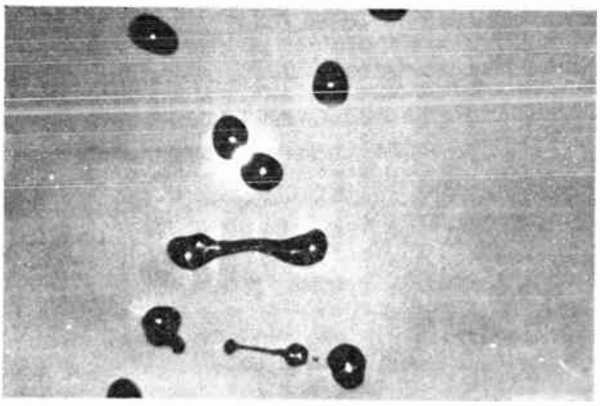
Chance collision and coalescence of water droplets, ice crystals, or snowflakes, followed by their fallout under the influence of the same gravitational and aerodynamic forces that promote collisions, is not the whole story. The best available theories of droplet aerodynamics and coalescence cannot account for the very rapid particle growth deduced from radar observations of well-developed rainstorms. What other factors could control or speed up the rate at which raindrops and other precipitation particles form?

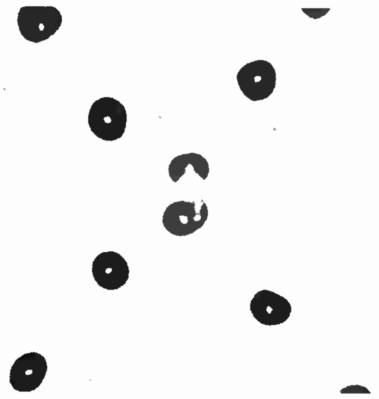
In 1879 the British physicist Lord Rayleigh came up with a possible answer: electricity. His laboratory experiments showed that large uncharged water drops that collided with each other generally bounced apart without coalescing, but that charged water droplets tended to coalesce, as did uncharged droplets colliding in an electric field. Rayleigh was aware of the meteorological implications of his observations. He recognized that electric forces in clouds may in some way be important in the growth of precipitation.

"It is obvious," he wrote, "that the formation of rain must depend very materially upon the consequences of encounters between cloud particles. If encounters do not lead to contacts, or if contacts result in rebounds, the particles remain of the same size as before; but, if the issue be coalescence, the bigger drops must rapidly increase in size and be precipitated as rain. Now . . . we have every reason to suppose that the results of an encounter will be different according to the electrical condition of the particles, and we may thus anticipate an explanation of the remarkable but hitherto mysterious connexion between rain and electrical manifestations."

Although the role electricity plays in rain formation is still far from clear, recent studies by Sartor and his colleagues at the National Center for Atmospheric Research, and by other groups, have revealed many details of electricity's effects on colliding drops. Sartor and his (NCAR) colleagues began their studies in an attempt to pinpoint the mechanism responsible for radio emissions that had been observed from warm clouds. In these laboratory studies, Sartor's group noted that a faint glow appeared when oppositely charged water drops met, whether they coalesced or not. The glow, so weak that it was visible only to dark-adapted eyes, was eventually identified as the sum of many tiny, extremely brief sparks during which electric charge was transferred as the water drops approached each other. The group turned its attention to the problem of recording these sparks photographically, and eventually developed a technique for producing drops that would meet at a pre-







FIGURES 3, 4, and 5. The very rapid, near-concurrent growth of electric fields and precipitation particles in thunderclouds may be explained, in part, by the charge-transfer processes revealed by these laboratory experiments carried out under the direction of J. Doyne Sartor at the National Center for Atmospheric Research. Oppositely charged streams of water droplets, each 800 μm in diameter and carrying a charge of about 20 picocoulombs, were directed toward each other on direct collision (Fig. 3, above), glancing (Fig. 4, above right), and near-miss (Fig. 5, bottom right) trajectories. In each case, the discharge you see is the cumulative, visible result of large numbers of invisible, smaller discharges, each lasting only about one ns. But the electrical consequences of each encounter are quite different. In Fig. 3, where droplets coalesce, the opposing charges carried by each droplet are neutralized; although coalescence—aided, perhaps, by electrostatic forces—results in droplet growth, it does not help increase the cloud's electric field. In Fig. 4, however, where droplets merge and then reseparate, not all of their charge may be neutralized; and in Fig. 5 they may retain most of their charge. If the droplets of Fig. 4 and 5 were droplets in clouds, and if they were of unequal size but like-sized droplets carried charges of like sign, then gravitational separation of these droplets could intensify the cloud electric field—sufficiently, perhaps, so that subsequent droplet encounters that might have resulted in a near miss, or in a collision followed by reseparation, would instead result in coalescence. The same processes might occur even if the droplets were initially uncharged; the induction-charging process illustrated and explained in Figs. 6 and 7 will help you to see how this might take place in clouds.

dictable point, so that time exposures could be made of the light produced by many successive small sparks. Figures 3, 4, and 5 testify to their success. Each spark apparently occurs just before the drops actually touch, at the same time that charge transfer takes place, and lasts only about one nanosecond. Other photographs reveal that the oppositely charged droplets literally reach out toward each other as they approach, increasing the likelihood that their approach will lead to collision and perhaps to coalescence.

Uncharged drops behave differently. Photographs, such as Fig. 1, show that they often flatten slightly as they meet; then they may either bounce apart like rubber balls or coalesce. What they do seems to depend mostly on their size and on trajectory considerations. Larger drops often fail to coalesce because of the barrier set up by an air film between the drop surfaces. Uncharged pairs of smaller drops, which, because of surface tension effects, behave as though they are harder, are more likely to break through this air film and unite.

Droplets that are initially uncharged in the atmosphere, however, may not remain uncharged for long, at least not once the turbulent circulation induced by convection begins. Turbulent aerodynamic forces bring about many more collisions and near-collisions between cloud droplets than downward-acting gravitational forces can. In the presence of an electric field, and an electric field always exists in the atmosphere—even when clouds are absent, the noncoalescing encounters between cloud particles can result in the induction charging of the particles, thus increasing the likelihood that subsequent encounters between the particles will lead to their coalescence. M. H. Davis did the theoretical work that suggested this important possibility in 1964 when he was at the RAND Corporation (he is now at NCAR). Davis showed (Fig. 6) that as two electrically conductive but uncharged spheres of any relative size approached each other, the electric field between them would intensify exponentially as their separation decreased. Sartor's group subsequently showed that this locally intense field can in turn induce opposing charges in the approaching drops (Fig. 7), and, if spark discharges between drops do not completely neutralize these induced charges—as they would if coalescence occurred—the separating drops remain oppositely charged to at least some degree.

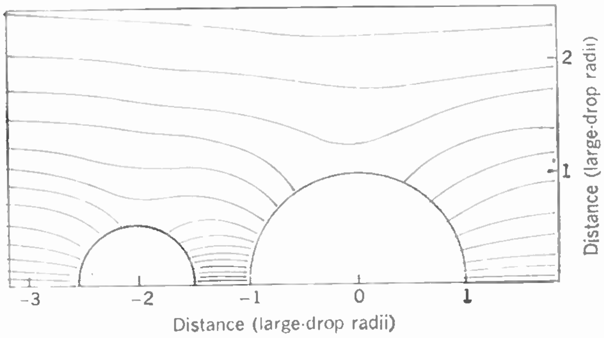
Experiments by other groups have also shown that the coalescence of pairs of drops depends considerably on the strength of the ambient electric field. In the absence of an external electric field, for example, only about 30 percent of the collisions occurring between drops 600 and 100 μm in diameter lead to their coalescence. But in a field of approximately 15 V/cm, a value quickly reached and then greatly exceeded in the growth of a run-of-the-mill thunderstorm, the rate of coalescence of such drops triples. Collision and coalescence of the much smaller droplets that populate nonprecipitating clouds in early stages of their development probably depends to an even greater extent on the intensity of the ambient electric field. Since the momenta and collision cross sections of smaller droplets are much less than those of larger drops, their coalescence is much less likely to take place through impact alone.

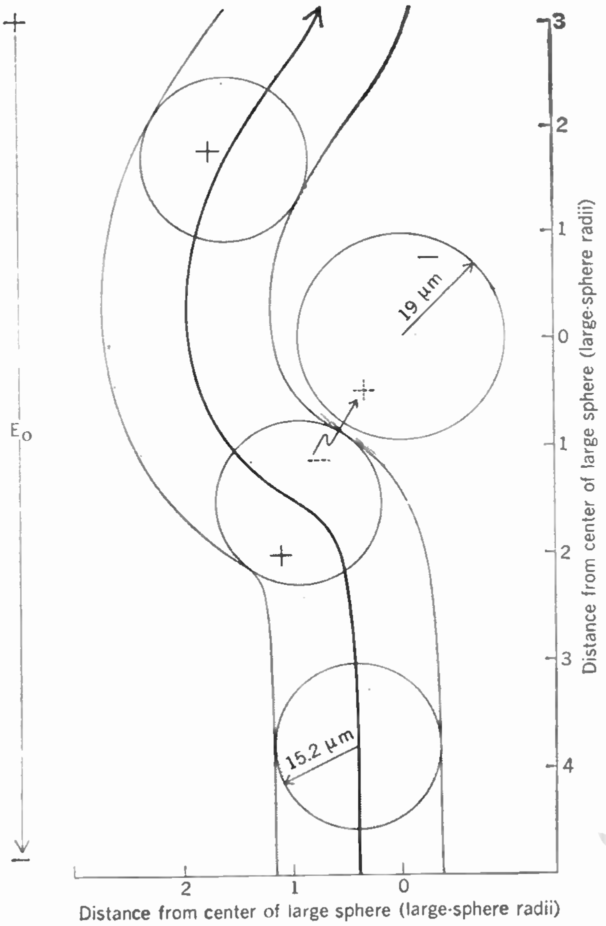
These processes observed in various laboratories suggest the existence of perhaps universal principles of cloud-drop charging, discharging, combination, and bounceoff. They also raise many questions. Do motions of drops in the laboratory reproduce those of real cloud drops? How do drop-size distributions and the ambient temperature and pressure affect bounce and coalescence processes, and what role is played by any secondary drops produced by these collisions? In Sartor's experiments, for example, the discharge between a pair of drops sometimes triggered a second area of luminescence, a positive corona near one of the drops in the next pair streaming through the apparatus. The most important questions, however, are how are collision and coalescence processes modified quantitatively by electric charges and fields in clouds themselves, and how do the electrical properties of clouds respond to electrical events in the surrounding atmosphere?

Perhaps the answers to these and similar questions will lead to the answer to closely related questions that have preoccupied workers in atmospheric and cloud electricity ever since Ben Franklin flew his famous kite and drew electricity down from the sky: What basic mechanisms generate the initial electric fields and charge distributions in the atmosphere and in clouds? And how do these rather feeble initial fields intensify as rapidly and spectacularly as they do in thunderclouds?

A preliminary peek at the thundercloud

In attempting to assess the origin of thunderstorm electricity, most students of these matters point to the fact that the electric field intensity around a thundercloud increases very rapidly soon after frozen particles appear in the cloud. This observation has led most





FIGURES 6 and 7. Two initially uncharged but electrically conductive spheres of any relative size will, upon approaching each other in an electric field, greatly intensify the field between them. Figure 6 (top right) shows this by the convergence and closer spacing of lines representing a field, which, in the absence of the spheres, would be uniform in strength and oriented parallel to the line connecting the spheres' centers. The field between spheres intensifies because the weak dipolar charges initially induced on each sphere by the polarizing ambient field are imaged from one sphere to the other as the two approach each other. The intensifying field in turn induces still stronger dipolar charges on the approaching spheres, and so on. Permanent and opposite, unipolar charging of the originally uncharged spheres will result when induced charges of one sign are transferred from one sphere to the other, via spark discharges like those shown in Figs. 4 and 5. This induction-charging and charge-transfer process is illustrated in Fig. 7, (bottom right), which shows the computed trajectory of a smaller sphere relative to a freely falling larger sphere. When the small sphere is in the lowest position, both spheres are uncharged. As the smaller sphere approaches the larger, both are polarized in the electric field E_0. At the point of closest approach or actual collision, charge transfer takes place, as shown by the dash-line charge symbols and the curled arrow connecting them. As the spheres separate (small sphere at top), each remains charged as shown by solid-line symbols.

atmospheric scientists to conjecture that the freezing process itself is an inherent and necessary part of the cloud electrification process. The charging of small ice splinters ejected from the frozen skin of a supercooled water droplet, which would result from proton (H^+) migration along the thermal gradient set up when the droplet freezes from its outer surface inward, has met with considerable favor as a thunderstorm electrification mechanism. Laboratory experiments indicate that some such process can indeed lead to the creation and separation of large and small ice particles with opposing electric charge; and it is generally believed that gravitational and/or aerodynamic separation of these particles within the cloud can lead to the buildup of the cloud electric field. Other experiments suggest that similar charge separation can occur when dendritic ice crystals lose their fragile end branches, so long as a temperature gradient exists between the center of the ice crystal and the branch ends. Still other experimentally based theories generate electric charge, and separate it more or less in accordance with the charge distribution inferred to exist in thunderstorms, by invoking the frictional charging that would result from collisions between rain, hail, ice, and/or snow particles within the cloud. And still other theories draw sufficiently abundant amounts of electricity out of the effects created by the incorporation of impurities such as alkali halides and/or ammonia into the lattice structure or water-filled voids of growing ice crystals or hail pellets. (References 6–15 can provide more detailed discussions of these conflicting theories by their partisans, proponents, and others, and lead you to the rest of this rich literature, almost none of which appears in primarily electrical journals.)

The charge-separation process Sartor has proposed to explain the growth of the intense electric field of a thundercloud draws heavily on laboratory results such as those shown in Figs. 4 and 5, in which sparks—indicating that charge transfer between drops was taking place—were produced even when coalescence did not result from the encounter. Sartor thinks that when similar noncoalescing encounters take place between unequally sized droplets in clouds, the opposite charges induced on and retained by the drops may be separated as larger drops fall away from smaller ones. This separation, as in other theories, rapidly increases the strength of the initially perhaps quite weak electric field. The droplets themselves, he stresses, need not be charged initially. Uncharged cloud particles of dissimilar size that approach each other in an electric field will become oppositely charged by induction, according to the theoretical principles suggested in Figs. 6 and 7, and gravitational separation will again do the rest.

All thunderstorm electrification theories—and we have not nearly exhausted the list—invoke gravitational (or aerodynamic plus gravitational) separation of charged particles, but not all stress the collision and induction-charging process as strongly as Sartor's theory does. Sartor's numerical integration of the charge separated by his process appears to produce electric fields of appropriate magnitude well within the severe time constraints imposed by observations of real thunderstorms. But, then, so do several other theories, according to their proponents. The important thing to realize is that all thunderstorm electrification theories are in a preliminary stage of development. These preliminary theories provide a nucleus on which future theories can grow. Just as important, they highlight the many kinds of measurements still badly needed to flesh out, and to decide among, competing theories.

More accurate field measurements are critical

Take an example of just one of the fundamental electrical quantities for which more accurate and widely distributed measurements are needed. Published information on the strength of the electric field in and around thunderclouds is so contradictory and confusing that 25 percent might be considered "accurate" here, though measurements of the electric field in the cloud-free atmosphere or in less strongly electrified clouds are considerably more accurate than this. Walter Evans, of the University of Arizona Electrical Engineering Department, discussed this problem in a recent review [16] describing methods of field measurement. Most workers in atmospheric electricity agree with his bleak assessment. Evans became involved with instrumentation problems when he began field studies aimed at assessing electricity's role in droplet coalescence, and at assessing the effects the convection-driven distribution of space charge might have on cloud electrification.

Earlier studies of cloud electrification had shown that the electric field increased only very slightly in strength from the ground to cloud levels. This was contrary to expectation. It had been assumed prior to these measurements that the field strength would increase rather more sharply with height, because point discharge from objects on the ground beneath the clouds was expected to produce a layer of positive space charge adjacent to the ground. Hugging the negatively charged surface of the earth (a consequence of the so-called "electrode" effect), this layer of positive ions and other positively charged particles was expected to reduce significantly the electric field strength at or near the ground. But instead, the field strength appeared, from these early measurements, to be relatively reduced at cloud level and not at ground level, presumably because the positively charged ions and other electrified aerosols released by point discharge were, by one means or another, rapidly reaching the clouds. The question was, how? The mobility of the large ions produced by point discharge appeared to be much too low for such ions to reach the cloud base in time to reduce electric fields there to the surprisingly low values observed by earlier investigators. Evans points out that only two alternatives can explain this apparent anomaly: Either "excess" space charge, of the positive polarity necessary to reduce field strengths at cloud levels to the observed values, is carried rapidly to the cloud base by updrafts—as Bernard Vonnegut of the State University of New York has in fact suggested in the convective cloud-electrification theory (Fig. 8) being tested with the New Mexico group's wind tunnel (Fig. 2)—or the available measurements are wrong.

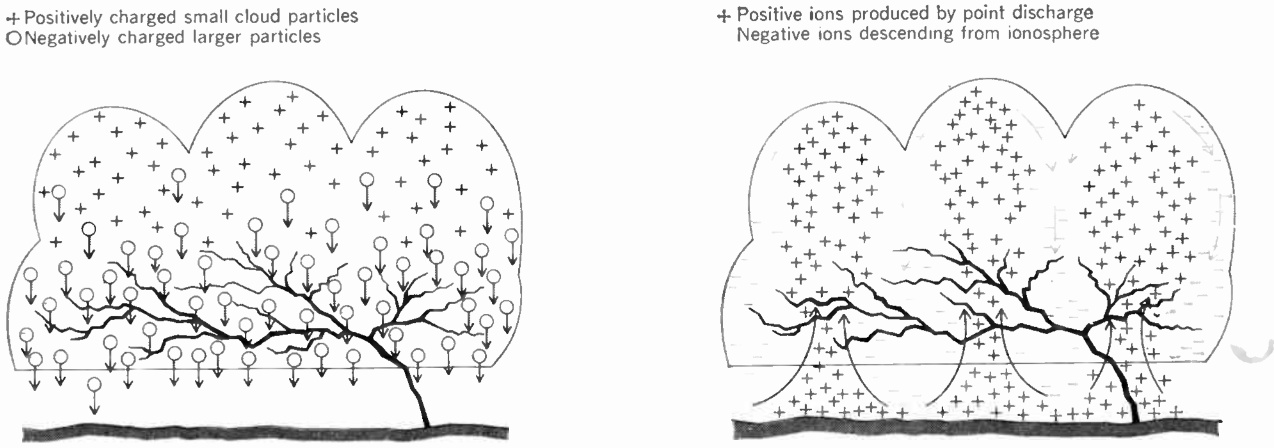
The importance of deciding between these alternatives can be appreciated when one realizes that Vonnegut's convective theory for cloud electrification, one of the leading contenders in a crowded field, is strongly at variance with most others that invoke various permutations of phase-change and collision processes within the clouds themselves in order to explain the charge buildup that leads to lightning. Vonnegut's theory is an attractive one, however, because it is one of the very few that can be directly checked by field experiments like the one being undertaken at New Mexico. Diverse other experiments also designed to change electrical conditions in fair-weather cumulus clouds, not thunderclouds, by utilizing natural convection, have been carried out by Vonnegut and various colleagues[4,17–19] and have met with some success.

In attempting to resolve such fundamental disputes, as well as just to get more data for reconnaissance purposes, most investigators use a simple device known as a field mill to measure the strength of the electric field. The basic principle underlying the operation of the electric field mill is illustrated in Fig. 9. In this arrangement a pulse whose polarity and magnitude depend on the sign and magnitude of the field will appear on the oscilloscope each time the shield is removed, exposing the test plate to the electric field. The pulse is shown as a decaying exponential on the oscilloscope display, but any other type of indicating device may be used in place of the oscilloscope. Some synchronized rectifying device must be used for any dc indicating device that includes the sign of the field, however, since a second pulse occurs when the shield is replaced.

One type of modern field mill is illustrated by Fig. 10. In this version the shielding of the plate is accomplished many times per second by the combination of the cover plate and the rotating, propeller-shaped plate. The signal from the test plate is alternating, therefore, and the sign of the electric field must be recovered by some type of synchronized chopper or commutating signal. According to Evans, this type of mill can be made quite sensitive, while remaining stable, by the use of amplifiers tuned to the frequency of rotation of the rotating plate. Meters that will accurately measure to one tenth of a volt per centimeter are not unduly difficult to design. Mills designed to operate on the ground, however, are usually operated in an inverted position to protect them against rain and dust. Although inverted operation decreases the sensitivity, it is easily recoverable by increasing the amplifier gain, within limits set by the signal-to-noise attributes of the amplifier.

When electric field measurements are made off the ground, in the free atmosphere or in the neighborhood of strongly electrified clouds, single-circuit mills like those shown in Figs. 9 and 10 will not do. Differential measurement methods using dual-circuit field mills are required, since ionization and static charging will make the results obtained with single mills meaningless, unless they are operated very carefully under conditions that are difficult to attain in practice. Evans, under National Science Foundation sponsorship, a few years ago attempted to develop a field-measuring device of reasonable cost that would be at least as reliable as existing mills when used under the difficult conditions encountered in clouds. He has made some progress toward overcoming problems of cost, accuracy, and corona discharge with the dual mill illustrated in Fig. 11. Developed for specific use in clouds—designed, in fact, to be dropped into the cloud from above by parachute, together with radiosonde and telemetry equipment—the inputs of this

FIGURE 8. Most workers in cloud physics agree that the upper part of the typical strongly electrified thundercloud is charged positively and the lower part charged negatively. They further agree that this charge distribution is responsible for the electric field buildup that leads to lightning. Here agreement ends. How are the oppositely charged cloud particles created in the first place and how are they separated within the cloud? These key questions remain unanswered. Two major classes of charge generation-and-separation theories differ, as suggested by these diagrams. Most workers in cloud physics prefer some variant of the field-generating process shown at the left, in which oppositely charged particles created by micro-physical-chemical processes (dozens of possible processes have been suggested) are separated principally by gravitational forces. Bernard Vonnegut of the State University of New York at Albany, however, and some others, believe that charge separation and cloud charging may result as thermal updrafts carry positive ions—derived from point discharge at the ground—to higher levels in the cloud, whereas cooler downdrafts carry negative ions—derived from the ionosphere—to lower cloud levels. This theory, shown at right, is the basis for the weather modification experiment illustrated in Fig. 2.

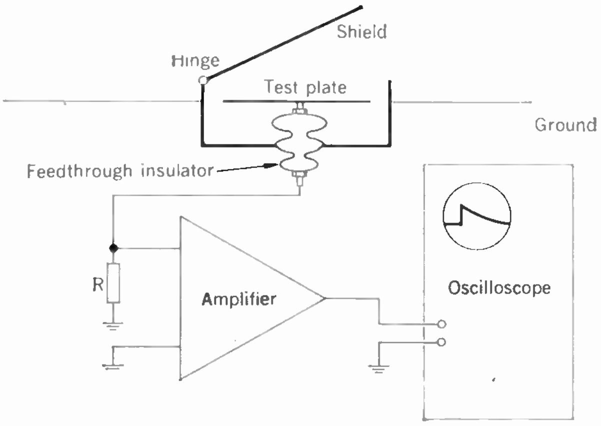


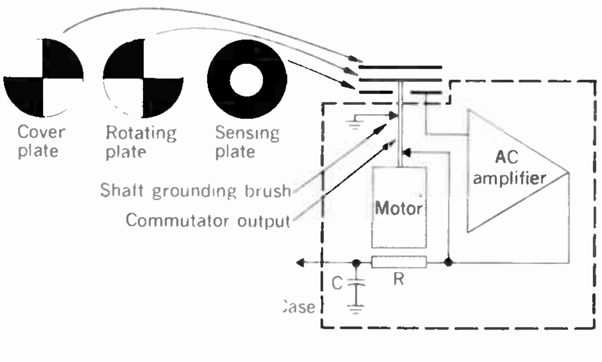
field mill are electrically balanced so that static charging of the case will not cause erroneous results, and mechanical symmetry is also preserved so that equal fields will be sampled as the device descends.

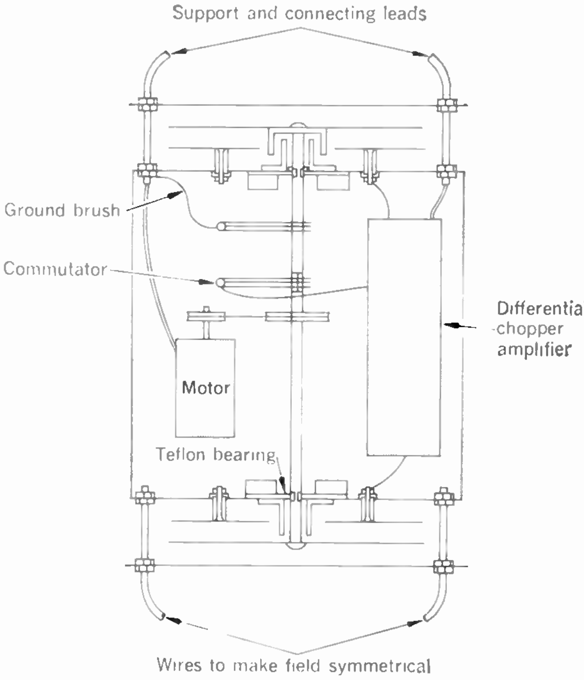
Modern field mills such as those illustrated in Figs. 10 and 11 have a particularly—and desirably—simple underlying theory. Since a fixed fraction of the vertical field flux must terminate on the cover plate or the test plate (with the fraction determined only by the geometry and the voltage of the test plate), the physical problem is accurately represented by the simplified arrangement shown in Fig. 12.

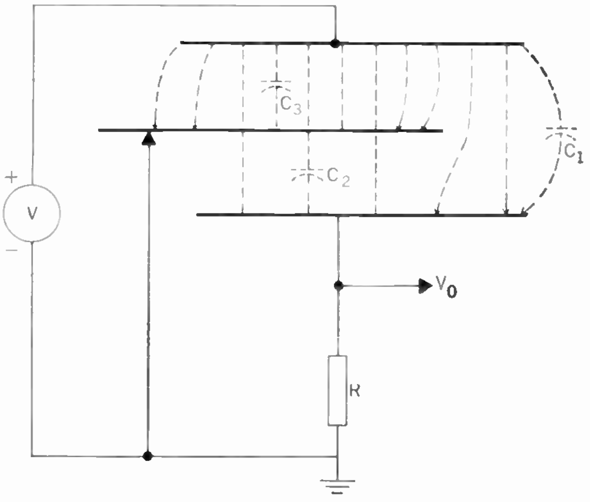
Induction-type field meters that utilize a rotating element have been used since the mid-1920s. They were first used aloft, on dirigibles, in the early 1930s and have been subsequently developed and adapted for use on aircraft, primarily by workers in Germany, the U.S.S.R., and the United States. Meters properly installed on aircraft appear to provide a reliable, fairly trouble-free method of measuring fields of appreciable, though not thunderstorm, magnitude. D. R. Fitzgerald and his colleagues at Air Force Cambridge Research Laboratory[20] have used aircraft-mounted field meters routinely for cloud penetration studies, and have encountered no particularly serious environmental difficulties. Although corona pulses are of course noted at times, they can be easily recognized on recorded traces of the electric field and do not interfere significantly with analysis of the field patterns. But the problem of separating the aircraft charge field from the external field has not been solved. Existing instrumental techniques for maintaining a zero charge on the aircraft during cloud penetration are rather cumbersome, and in most cases cannot cope with the large currents encountered during thunderstorm penetrations.

FIGURES 9, 10, 11, and 12. This series of drawings illustrates the basic principles and some varieties of mills used to measure the sign and strength of electric fields in the atmosphere and within clouds. Figure 9 (below) portrays the basic principle underlying all common variations of the field mill. Figure 10 (below bottom) shows the essential parts of a modern electric field mill. Figure 11 (top right) shows the mechanical arrangement of a differential field mill that was designed by W. H. Evans of the University of Arizona, in an NSF-sponsored attempt he made to overcome certain of the environmentally induced accuracy problems inherent in existing designs. Figure 12 (bottom right) portrays the equivalent circuit of field mills like those shown in the preceding two figures. Text discusses additional data needs in cloud physics.









Things are in worse shape when we turn to the question of charge distribution in thunderstorms. Better techniques are needed for determining the density and distribution of at least three kinds of charge carriers: (1) cloud droplets, and large ions attached to various nuclei and aerosols; (2) small ions of both signs; (3) precipitation particles. More accurate measurements of these charges would reveal how the net charge per unit volume develops in convective clouds and this would, of course, greatly assist in evaluating the various charge-generation theories that have been proposed.

Clouds emit radio noise

Direct measurements of the electric field, charge distribution, and other parameters in and around clouds and in the clear atmosphere with modern instruments represent one necessary approach to observational problems in weather electricity. Other, indirect, approaches are also possible, of course.

In thunderstorms, at any rate, and even in less intensely electrified clouds, charges carried by raindrops and other particles are transferred between particles when they approach each other closely enough. And when sufficiently large numbers of cloud particles do discharge simultaneously, detectable amounts of ultrahigh-frequency radiation are produced. Although it is difficult to measure the number and distribution of free electric charges in thunderstorms directly, some estimates can be made from observations of this radio emission from clouds. In 1963, when Sartor and his colleagues at NCAR began to investigate the radio signals emitted by warm clouds, they were searching for a technique with which clouds could be studied from a safe distance. They hoped to develop a technique that would reveal whether cloud droplets were indeed discharging, and perhaps coalescing to form raindrops as Rayleigh had conjectured nearly 100 years ago, and whether the strong electrification closely associated with the appearance of precipitation particles was somehow connected with the particle collision and/or coalescence process. It now appears probable that such studies of electromagnetic emission may indeed lead to the remote sensing of important meteorological properties of highly electrified clouds. Evans and his colleagues at the University of Arizona, and many other investigators, have, for example, been monitoring the prelightning-stroke electromagnetic radiation from thunderclouds for the past several years. They have been able to detect radio-frequency signals originating from cumulonimbus clouds several miles away from their antennas, at times up to 15 minutes before the first lightning strokes occurred. They offer no theory to explain these signals, most of which are near the limit of detectability in the presence of atmospheric noise, but they assume that they are caused by minor streamer processes in the cloud as fluctuations in charge density take place.

It is quite clear that remote-sensing studies of cloud electricity are important, and that more are needed. But let's not deceive ourselves; before such remote signals can be interpreted adequately, much more information is needed on the consequences of collisions between charged and uncharged particles—liquid and solid—of different sizes and perhaps shapes, both in the presence and in the absence of electric fields; and more information is needed on the collective behavior of charged particles in cloudlike swarms. Nor is this all the information that's needed before electrical approaches to weather modification can be removed from the realm of speculation.

Chicken-versus-egg controversy

There is still basic disagreement, for example, over so apparently simple an observational question as whether the prolific precipitation—the so-called "rain gush"—that often pours from a thundercloud immediately after a lightning stroke, comes before or after the buildup of the locally intense electric field that leads to the lightning. Dr. E. J. Workman of the University of Hawaii, originator of one of the more widely accepted theories for thunderstorm electricity[6,21,22] has observed that we have been collecting evidence on this disputed question since the weather-modification days reported in the *Book of Exodus:*

"And Moses stretched forth his rod toward heaven, and the Lord sent thunder and hail, and the fire ran along upon the ground; and the Lord rained hail upon the land of Egypt. So there was hail, and fire mingled with the hail, very grievous, such as there was none like it in all the land of Egypt since it became a nation."

Moses may have been overenthusiastic in these early weather-modification efforts. He made his point, in any case. But the point here is that even in this account it was not made clear which weather phenomenon took precedence. Which came first—the hail, the thunder, or the "fire" (biblical nomenclature, perhaps, for either lightning or the luminous corona discharge known as St. Elmo's fire)?

Modern attempts to answer such observational questions rely heavily on radar techniques for detecting precipitation particles at early stages of their formation, within the clouds themselves, before they reach the ground. In the few careful studies that have been carried out, radar observations have been supplemented by observations of the electric charge, potential gradient, and conductivity—in the air within and around the clouds and at the ground below. Even such careful observational programs have not resolved the question of precedence to everyone's satisfaction.

Charles B. Moore (of the New Mexico Institute of Mining and Technology), Bernard Vonnegut, and their colleagues are perhaps the leading proponents of theories that accord precedence to the electrical phenomena, or at least to the lightning phenomenon. They have proposed that coalescence triggered by lightning discharges can explain the rain gush from thunderclouds, as schematically outlined in Fig. 13. This basic effect has in fact been demonstrated in a laboratory cloud chamber by Sartor. In separate experiments, H. Plumlee at the University of Illinois found that a horizontal electric field of 3500 V/cm increases the collision efficiency of a 30- and 5-μm pair of particles by an astounding factor of 34.5; of a 40- and 5-μm pair by a lower but still respectable factor of 5.6; and of a 50- and 5-μm pair by a still lower but equally respectable factor of 5. A field of 3500 V/cm

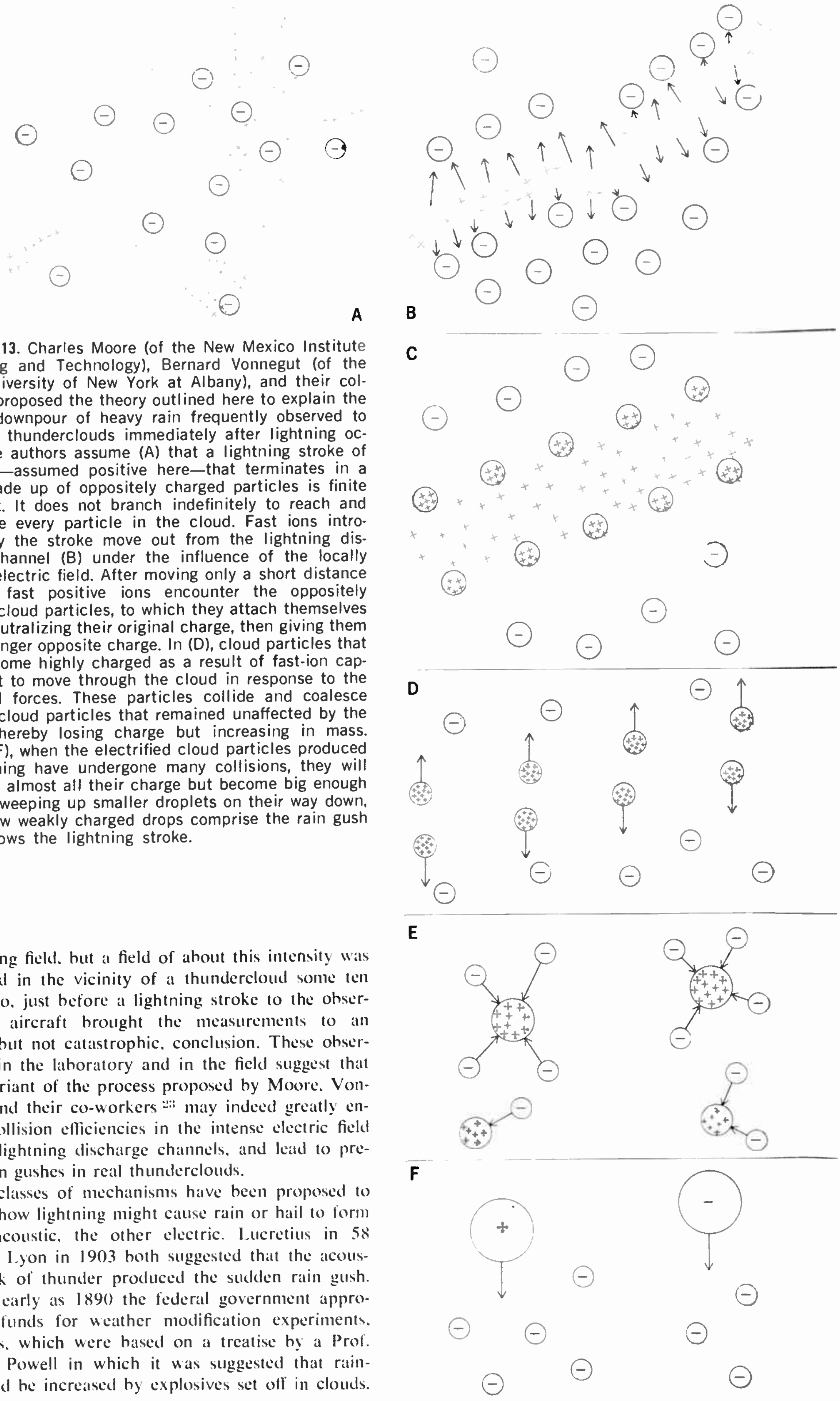


FIGURE 13. Charles Moore (of the New Mexico Institute of Mining and Technology), Bernard Vonnegut (of the State University of New York at Albany), and their colleagues proposed the theory outlined here to explain the sudden downpour of heavy rain frequently observed to fall from thunderclouds immediately after lightning occurs. The authors assume (A) that a lightning stroke of one sign—assumed positive here—that terminates in a cloud made up of oppositely charged particles is finite in extent. It does not branch indefinitely to reach and neutralize every particle in the cloud. Fast ions introduced by the stroke move out from the lightning discharge channel (B) under the influence of the locally intense electric field. After moving only a short distance (C), the fast positive ions encounter the oppositely charged cloud particles, to which they attach themselves—first neutralizing their original charge, then giving them a far stronger opposite charge. In (D), cloud particles that have become highly charged as a result of fast-ion capture start to move through the cloud in response to the electrical forces. These particles collide and coalesce (E) with cloud particles that remained unaffected by the stroke, thereby losing charge but increasing in mass. Finally (F), when the electrified cloud particles produced by lightning have undergone many collisions, they will have lost almost all their charge but become big enough to fall. Sweeping up smaller droplets on their way down, these now weakly charged drops comprise the rain gush that follows the lightning stroke.

is a strong field, but a field of about this intensity was measured in the vicinity of a thundercloud some ten years ago, just before a lightning stroke to the observational aircraft brought the measurements to an abrupt, but not catastrophic, conclusion. These observations in the laboratory and in the field suggest that some variant of the process proposed by Moore, Vonnegut, and their co-workers[23] may indeed greatly enhance collision efficiencies in the intense electric field around lightning discharge channels, and lead to precipitation gushes in real thunderclouds.

Two classes of mechanisms have been proposed to explain how lightning might cause rain or hail to form—one acoustic, the other electric. Lucretius in 58 B.C. and Lyon in 1903 both suggested that the acoustic shock of thunder produced the sudden rain gush. And as early as 1890 the federal government appropriated funds for weather modification experiments, in Texas, which were based on a treatise by a Prof. Edward Powell in which it was suggested that rainfall could be increased by explosives set off in clouds.

Congress withdrew the funds, however, when the scientific basis for the experiments began to appear dubious. The acoustic possibility still appears dubious.

G. G. Goyer and his colleagues at the National Center for Atmospheric Research have scrutinized the acoustic effects of thunder and explosives on cloud and precipitation particles with great care in the last several years. Having detonated explosives near the plume of Old Faithful geyser in early experiments, and produced a shower of fine hail and curtains of rain from the plume, they subsequently explored, in the laboratory and with computer models, the possibilities that cavitation, droplet shattering, or adiabatic cooling caused by the acoustic shock waves might be responsible for these results. Although not claiming that they have exhausted all the possibilities, Goyer and his colleagues are ready to discount the importance of shock as a major influence in rain- and hail-forming processes in clouds in the sky. Their investigations show that the various alternatives for shock-induced precipitation falter in the light of careful scrutiny. The striking results they achieved at Old Faithful remain an unexplained anomaly, but they do not necessarily violate this general conclusion, since neither the high chemical and gas content of the geyser water, nor the great concentration of droplets in the geyser plume, reproduce conditions in thunderclouds.

The alternative explanation—that the electrical effects of lightning cause the precipitation gush from real thunderclouds—is also of long standing. Moore and his collaborators favor an electrical mechanism over an acoustic one because they think the events set in motion by lightning could operate for periods of many seconds after the stroke. Acoustical forces, on the contrary, would have to operate over a much shorter period—the length of a thunderclap—and all subsequent droplet growth could proceed at rates determined only by the gravitational fall of the droplets. Arguing by analogy, the Moore group supposes that lightning strokes hidden from view in the cloud are similar to those that are visible outside of the cloud, or similar to the electric sparks that can be produced artificially. These familiar high-voltage sparks are like a river-and-tributary system, or like the branches of trees—in which the branching process, while extensive, is finite; the branches do not develop to the point where they occupy the entire available volume. This argument assumes that the invisible lightning sparks within a cloud also do not subdivide almost indefinitely, until a tiny discharge reaches to, and neutralizes, each individual charged particle within the cloud, but that the branching process stops short of such extreme development because sufficient electric energy probably is not available for the creation of an ionized path to each and every charged cloud particle. The lightning does not bring opposite charges—those in the cloud and those carried into it by lightning—into electrical contact; it only brings them much closer to each other. This argument is supported by the well-known electric discharge pattern that B. Gross created over ten years ago when he irradiated a Plexiglas plastic block with megavolt electrons. The spark pattern "frozen" into the plastic block is finite; it does not subdivide indefinitely. A lightning discharge thus may not immediately neutralize all of the charged cloud particles but may instead accomplish only a gross neutralization.

If this interpretation of the unseen lightning process within clouds is correct, it follows that even though the typical lightning discharge destroys a dipole moment of the order of 100 coulomb-kilometers within the cloud, it may still, temporarily and locally, create in-cloud regions of very high space-charge density and very high electric field strength. And in these regions the electrically accelerated drop-coalescence process outlined in Fig. 13 may begin.

Needed: one comprehensive theory

Vonnegut, Moore, Sartor, Workman, and many others who have devoted their careers to investigating electrical aspects of weather-making processes, and arguing their diverse viewpoints in the literature and at meetings, can be expected to stress the importance of the relationships connecting weather and electricity. But in recent years such authoritative and presumably more highly objective groups as the National Science Foundation and the National Academy of Sciences–National Research Council have also stressed the still tenuous but possibly critical nature of this relationship.

In its most recent (ninth annual) report on weather modification,[2] issued last summer and covering all federally sponsored activities in the field through mid-1967, the NSF said: "The role that atmospheric electricity plays in weather processes is becoming more fully appreciated, but is still only poorly understood. There seems to be no question, however . . . that electrical forces play a major role in the formation of weather. When these processes can be more clearly understood, it is very probable that major new tools will be made available to the weather modification scientist."

Three years ago, a special Panel on Weather and Climate Modification that surveyed problems and prospects in this field for the National Academy of Sciences–National Research Council,[1] was equally emphatic in its appraisal:

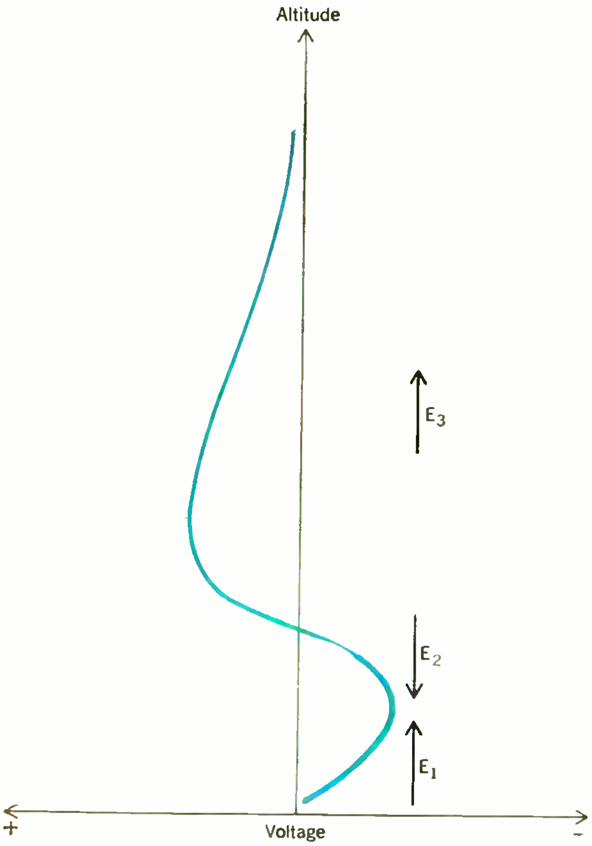
"The natural electrification process in convective clouds remains a mystery. It is possible that electric-charge distribution is a controlling influence not only, obviously, upon lightning, but also subtly upon precipitation and convective cloud dynamics. Vigorous research in cloud electrification has been urged in the past. The need is still urgent, mainly for the development of a more complete and tenable theory for the total cloud electrification process."

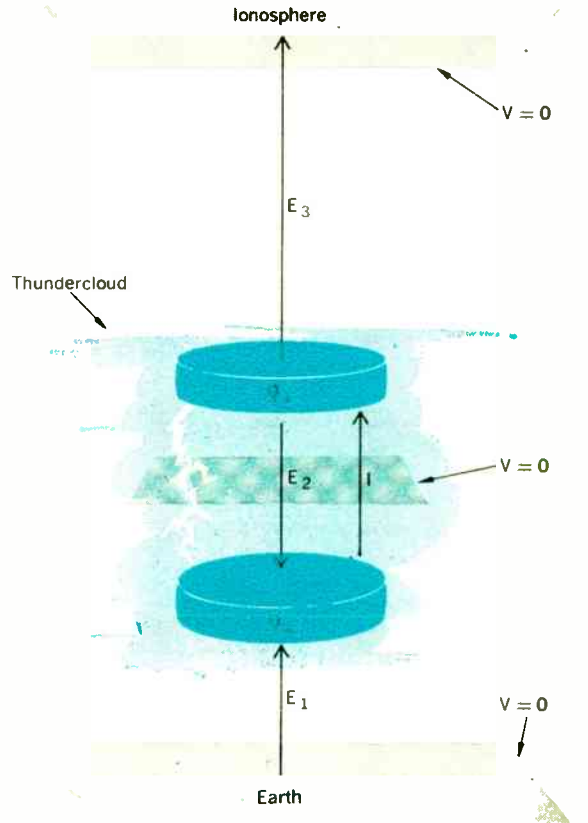
Cloud electrification theories do exist, of course, and they are numerous, as I have mentioned. None has gained clear ascendancy over the others because all are extended extrapolations from fragmentary data. Moreover, most attention, both theoretical and observational, has been directed, understandably, to the most obviously electrified cloud—the cumulonimbus or thunderstorm cloud.

When it really lets go, the typical thunderstorm cloud can spawn a lightning stroke every 20 seconds or so. Each stroke, on the average, destroys an electric dipole moment of about 100 coulomb-kilometers; for the dimensions involved in the average thunderstorm, this corresponds to a negative charge transfer of 20–30 coulombs. During the 10- to 30-minute active life of the typical storm cell lightning discharges can account for a negative charge flow at the rate of 0.3 to 1 ampere to the earth below. During the same period, and beginning, often, somewhat in advance of it, these storms elicit from the atmospheric layer immediately adjacent to the earth a positively charged, and therefore neutralizing, counterflow that approaches the rate of 0.1 to 1 ampere in a typical storm. This upward flow of charged particles takes place silently, however, by the unspectacular—usually invisible—process of corona or point discharge. In this process the discharging tips of every tree and bush, every roof, steeple, house corner, and spire, every blade of grass—and perhaps every hair on your head—cooperate passively in the electrical drama being played out with greater fury in the thundercloud above. Above the cloud silence rules again, a silence in which positive charge flows from the cloud to the highly conductive layers of the upper atmosphere at a rate of perhaps 0.5 to 1 or more amperes.

The extraordinary variation of voltage with altitude in the vicinity of a thunderstorm is shown by the qualitative curve, Fig. 14. George Freier of the University of Minnesota derived this curve using the thunderstorm electrical model shown in Fig. 15. In this thunderstorm model used by most workers, the upper region of the thundercloud carries a dominantly positive charge and the lower region carries a net negative charge. Since these charges are generally not equal to each other, because lightning strokes from cloud to ground drain negative charge from the base of the cloud but leave the upper, positive charge center unaffected, the typical thundercloud is at once dipolar in terms of its internal electrical structure, but effectively unipolar with respect to its surroundings. Between the upper and lower charge centers of the thundercloud an upward-directed current is usually assumed to flow. This current presumably regenerates the charge that the thunderstorm loses by conduction currents within itself and by lightning strokes to the ground. The conductive earth below the cloud, negatively charged with respect to the atmosphere, is effectively a zero-potential surface in the earth–thunderstorm–ionosphere system. Assuming that the ionosphere can be considered to constitute a second plane at zero potential (since its 0.2×10^6-volt potential with respect to the earth, measured for the first time in 1968, is negligible in contrast to the much higher voltages believed to exist in the mature

FIGURES 14 and 15. Professor George Freier of the University of Minnesota derived the curve showing the qualitative variation of voltage with altitude—Fig. 14 (below)—by using the widely accepted thunderstorm model shown in Fig. 15 (below, right). Attempts to verify and quantify curves and models like these will require worldwide surveillance of the atmosphere's electrical parameters at all altitudes in the vicinity of thunderstorms. To develop still more complete and tenable theories for the total cloud-electrification process—the prerequisite, perhaps, to developing reliable, safe, and economic methods of weather control—such surveillance must extend to fair-weather regions and to less strongly electrified clouds as well.

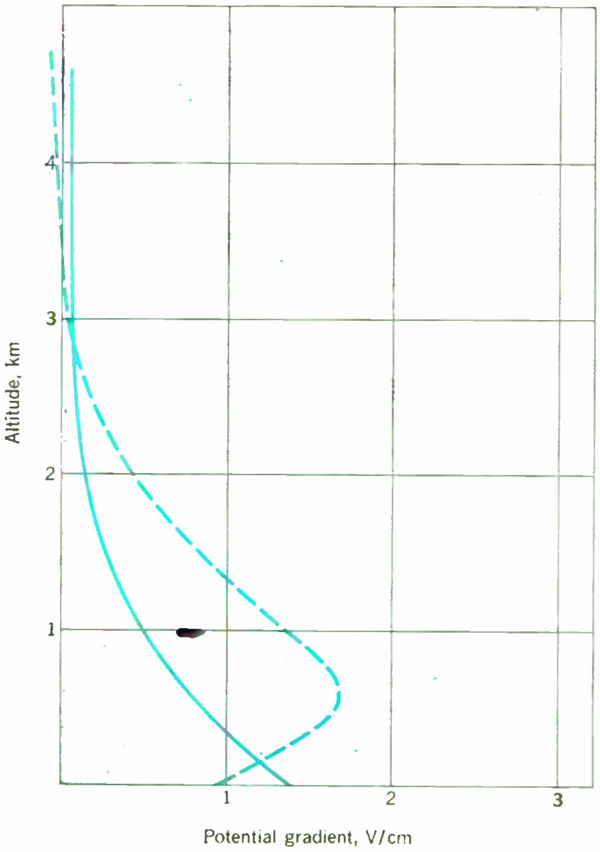




thunderstorm). Prof. Freier infers that a third surface of zero potential lies between the upper (positive) and lower (negative) charge centers of the thundercloud. With this model, the electric field must be directed upward, both above and below each charged region in the thundercloud: that is, the electric field must be directed upward between the cloud and the ionosphere, and between the earth and the cloud. Between the charge centers in the thundercloud, however, the electric field must be directed downward. Using these field directions and assuming that the line integral of the electric field from the surface of the earth, through the thunderstorm, to the ionosphere must be approximately zero, Freier is able to show that the qualitative variation of voltage with altitude in thunderstorm regions must be similar to the curve shown in Fig. 14.

The electrical perturbations created throughout the atmosphere by a mature thunderstorm are impressive. But even when the data needed to explain these perturbations are in hand, we shall not have the data needed to build a "complete and tenable theory for the total cloud electrification process." Many other of the scores of cloud types that the meteorologist distinguishes can also become appreciably electrified, and any theory that pretends to be comprehensive will have to deal with these clouds as well. Indeed, such a theory will have to explain the electric fields and charges that suffuse the skies even on cloudless days (Fig. 16), since these feeble fair-weather fields contribute somehow to the development of cloud electricity and such weather processes as precipitation. The work carried out by the NCAR group under Sartor, for example, on particle collision and separation, shows conclusively that even fields as weak as the average fair-weather field of 1 V/cm must be taken into account in computing the growth of the electric field in clouds. Most clouds are electrified comparably feebly; electric fields within and around them are only slightly more intense than the fair-weather electrical background.

FIGURE 16. These idealized curves portray altitude profiles of the electric field that might be obtained within the lower few kilometers of the atmosphere under fair-weather conditions. The solid curve is typical of profiles obtained in pollution-free atmospheres. The dashed curve typifies profiles obtained under polluted conditions, when additional free space charge is released into the atmosphere. Stormy area gradients are much more intense.



Then there is the tornado.

Electrical generation of tornadoes

The twisting, elephantine appendages (Fig. 17) that sometimes descend from thunderclouds are the most violent and destructive storms known. The lower end of a single tornado funnel may cover a ground area no more than 100 meters in diameter each time it touches down. Yet so diminutive a phenomenon has been estimated to require an energy input from its parent thunderstorm that is a hundred times greater than the combined capacity of all the electric generating plants in the United States. To drive its vortex winds at velocities that have been estimated to range from 100 m/s to the speed of sound, a tornado may require about 10^{11} J/s, or 10^{8} kW, of power. However, a single, run-of-the-mill thunderstorm cell contains some 10^{15} joules of energy, enough to power a tornado for several hours. Bernard Vonnegut is the source of these estimates. He offered them in support of yet another theory—an electrical theory of tornado generation that he proposed some years ago and refined recently. Though incomplete and much debated, this theory, like his convective thunderstorm theory (Fig. 8) serves a useful purpose: It is a "clay pigeon" for others to shoot at, in weather-modification studies and in more basic theoretical undertakings.

Vonnegut thinks tornadoes may be driven not only by the thermal and kinetic energy contained in thunderstorms, but by their electric energy as well. Lightning strokes, striking repeatedly through the same path at many times their normal thunderstorm rate, may heat the air enough to generate a coherent, sustained thermal updraft of appropriately tornadolike intensity; while charged particles of air may be simultaneously accelerated to extreme tornadic velocity in the strong electric field associated with the storm.

Energetic electrification may be necessary, though not sufficient in itself, to generate tornadoes of the particularly intense kind, in which any or all of the following evidences of intense electrical activity have been observed:

1. Luminous discharges in or near the funnel.
2. Strong sensations of heat near the funnel.
3. Scorching or desiccation of vegetation along the path.
4. Strong odors of ozone or nitrogen oxides near the funnel.

5. A buzzing noise, resulting perhaps from point discharge, that sounds like a swarm of millions of angry bees.

6. Unusually intense or frequent sferics impulses—radio-frequency signals generated by lightning—at receivers distant from the storm.

Vonnegut concedes that the mere presence of adequate electric energy cannot account for a tornado of this extremely intense kind. Adequate wind shear and vorticity must also be present in the air before these highly electrified vortices can form.

In a recent discussion of his theoretical approach,[21] Vonnegut has emphasized that in approaching the scientific and practical problems posed by the tornado, it is important to determine whether the electrical activity is an unimportant by-product of the storm or the storm's primary source of energy. Possibly the most direct way to answer this question would be to make temperature measurements in the tornado. If electricity is a significant item in its energy budget, such measurements should reveal local updrafts of air that have been heated 50°C or more over their surroundings by electric discharges. Air temperatures could be measured by several techniques: large numbers of inexpensive, perhaps recoverable, probes that would be sucked up into the tornado could be used, or instrumented drone aircraft or rockets could be directed into the storm. Current-density determinations made either from the ground beneath the tornado or in the air above the tornado-producing cloud would be especially useful. Remote measurements of other electric, magnetic, and electromagnetic parameters might also provide clues to the magnitude and nature of the currents flowing and thus permit the heating effect of these currents to be calculated.

Remote measurements of thunderstorm and tornado sferics have, of course, been carried out for many years. Such work has confirmed that electric discharges are much more intense and frequent in tornadoes than in thunderstorms: high-frequency (150-kHz) sferics are about as abundant as low-frequency (10-kHz) sferics in the average tornado, but in the average thunderstorm they are about a thousand times less abundant. The electromagnetic energy radiated shifts from lower to higher frequencies as the intensity of the storm increases. Tornado sferics, furthermore, appear to have a definite aperiodic burst-rate structure, which in general is lacking except perhaps in the most intense portions of thunderstorms. These differences between tornado sferics and thunderstorm sferics can only be explained by assuming that electrical and mechanical phenomena cooperate in perhaps fundamentally new ways to produce both sferics and tornadoes.

Enter electrohydrodynamics: EHD

John Carstoiu (of the International Consultant Scientists Corporation in Brookline, Mass.) thinks that electric and mechanical energy may in fact be coupled in tornadoes, perhaps in the ways Vonnegut suggests, by a new kind of wave—an "electrohydrodynamic" wave—whose theoretical existence he was led to postulate in attempting to explain certain peculiarities of sferics propagation. Electrohydrodynamic waves, to risk stating a complex matter in a much oversimplified way, might be thought of as the electrical analogs of the magnetohydrodynamic waves that have engaged the attention of many plasma physicists, astrophysicists, and not a few electrical engineers through the past decade or so. Electrohydrodynamics is a new subject, however, and quite distinct from magnetohydrodynamics. In MHD, interactions between electrically conductive fluids (plasmas) and

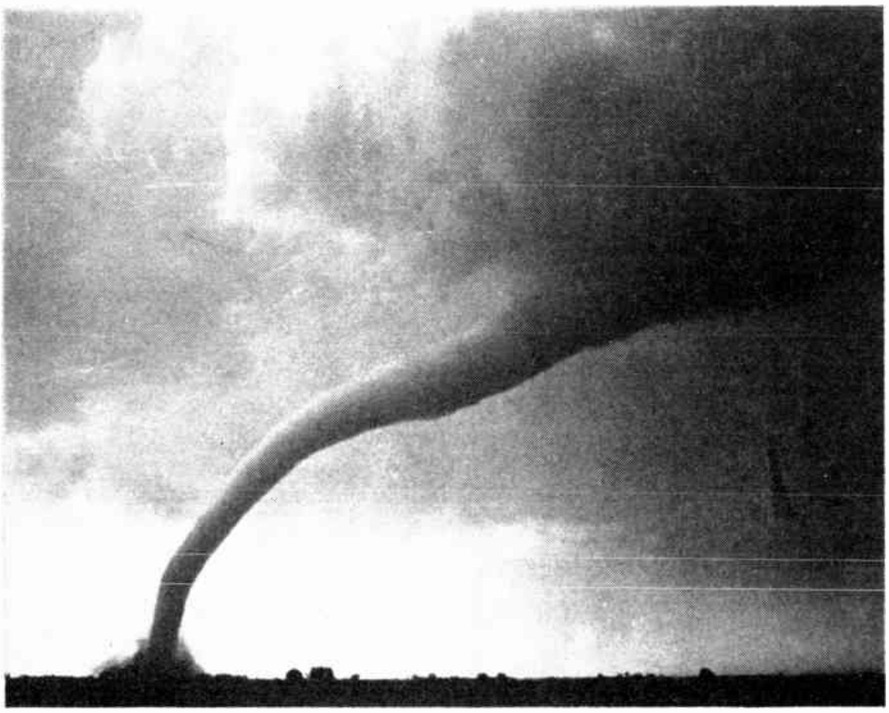


FIGURE 17. Tornadoes like this spectacular specimen may derive part of their energy from the electric energy contained in the thunderclouds that spawn them. Bernard Vonnegut (of the State University of New York at Albany) suggested that coupling of electric and mechanical energy possibly could account for funnel winds that have been estimated to range from 100 m/s to the speed of sound. He thinks that repeated lightning strokes along the same channel may generate intense thermal updrafts in which the air is heated by 50°C or more over the surroundings; it's also possible that the intense electric fields in such lightning discharge regions may help to accelerate charged air particles to tornadic velocities. Recent theoretical work by John Carstoiu tends to confirm Vonnegut's suggestion, as explained in the text.

magnetic fields are the dominant concern; but in EHD attention is directed instead to the reciprocal interaction between fluids in motion and electric fields. Magnetic and electric effects feed upon each other, of course; they are very difficult to separate. But in the basic equations of EHD, electric fields and electrical effects are assumed to be dominant and, as a result, many of the equations and concepts familiar to most electrical engineers may seem to be turned on their heads. As one consequence of this work, which he discussed at the Fourth International Conference on Universal Aspects of Atmospheric Electricity in Tokyo last spring,[25] Carstoiu thinks that electrohydrodynamic shocks (analogous to magnetohydrodynamic shocks) might occur in nature, and that they might manifest themselves—corporeally, so to speak—in both tornadoes and lightning. Thus, he suggests, a relativistic approach may be needed in studying such phenomena.

Although there is, in Carstoiu's view, no question that electrodynamic phenomena in the earth's atmosphere must obey the Maxwellian equations familiar to every electrical engineer, he asserts that Maxwell's system alone cannot explain these phenomena, especially not such phenomena as tornado clouds. These clouds, in his hypothesis, are conceived to be clusters of magnetically—not electrically—polar particles; therefore their explanation lies beyond the range of Maxwell theory. Neither can sferics be explained by Maxwell theory. They might also be profitably approached through further development of his still tentative electrohydrodynamic wave theory, in much the same way that whistlers—the audio-frequency waves that, like sferics, are thought to be generated by lightning—may be approached by magnetohydrodynamic wave theory. Sferics and whistlers are difficult to explain because their velocity of propagation is not constant; it depends strongly on the wave frequency in dispersive propagation media such as the atmosphere.

In his theoretical approach to such phenomena, Carstoiu defines a complex new quantity—the "magnetic diffusity" of the atmosphere (or of clouds)—and takes it into account in deriving the definition of electrohydrodynamic waves. These he defines as the sum of progressive waves traveling with velocities $\pm C_0$, where $C_0 = (\epsilon_0/\rho_0)^{1/2} E_0$, and ϵ_0 is the dielectric constant in a vacuum, ρ_0 is the average density of air, and E_0 is the electric field externally impressed.

Electrohydrodynamic waves travel along the electric lines of force, but they diffuse outward into the medium as they travel. Thus, as the travel time increases, the diffusion phenomenon comes to predominate over the propagation phenomenon. The fascinating result is that these wave disturbances will, in effect, disappear; they will "dissolve" into the environment. This happens when they reach a critical limiting speed that depends on the value of the previously mentioned "magnetic diffusity" of the dispersive propagation medium, be it atmosphere or cloud. It is especially interesting that the velocity expression for electrohydrodynamic waves to which Carstoiu was led in his search for theoretical clues to the sferics phenomenon, was discovered independently by the plasma physicist Stirling Colgate (now president of the New Mexico Institute of Mining and Technology) over ten years ago. But Colgate described this velocity as a limiting condition in tornado rotation!

Acknowledging that the phenomena under consideration are very difficult indeed, and that workers like himself and Colgate have taken only the first steps toward unravelling them theoretically, Carstoiu, like Vonnegut, also appeals for more observational and experimental evidence. He concluded his appeal to the workers in atmospheric electricity assembled in Tokyo with an eloquent quotation: "So far, admitted Lao Ting, it is more in the nature of a vision. There are, of necessity, many trials, and few can reach the ultimate end. Yet even the Yangtze-Kiang has a source."

The source of the mighty Yangtze River turned out to be not all *that* hard to find. Like most rivers, its source could be found in the far mountains and in the clouds that draped their peaks. The source of the electrical events that may control the clouds, however, will not be so easy to discover. Water's ubiquitous phase changes may indeed generate charge in the cloud-filled lower atmosphere, and convective and gravitational separation of these charged particles may indeed account for cloud electrification. But most thunderstorm electrification processes that have been suggested at various times by various authors invoke the prior existence of the earth's fair-weather electric field in order to account for the early stages of the electrification process. If the fair-weather field contributes significantly to cloud electrication, then we must ask: Where does the fair-weather field arise? To answer this question we must raise our sights above the clouds; we must look at the interactions taking place in the plasma envelope that surrounds the earth at ionospheric and higher altitudes. And we must seriously consider, once again, the long-discredited idea that sunspots have something to do with the weather.

Electrical structure of a small planet

Recent observations at mountain-top altitudes[26,27] have shown that variations in the earth's fair-weather field and atmospheric conductivity correspond strikingly with variations in solar-flare activity. And worldwide thunderstorm precipitation has been shown to correlate with some of these same solar emissions, but with a different time lag. Where then, does the "atmosphere" end? Where do the electric phenomena that may influence the weather begin?

J. Alan Chalmers, recently deceased British author of the standard textbook on atmospheric electricity,[28] limited his attention to those phenomena that occur or originate between ground level and an altitude of about 50 km. Above this upper boundary, highly mobile small ions and free electrons are supposed to be abundant enough to permit planet-wide equalizing currents of atmospheric electricity to flow; charges reaching this assumed equipotential surface (or region) presumably spread around the earth in a very short time. Below 50 km the air conductivity is much less, and the more sluggish motion of the large electrically charged particles—chiefly molecular ions and still larger particles—that predominate at these lower levels is determined by a combination of elec-

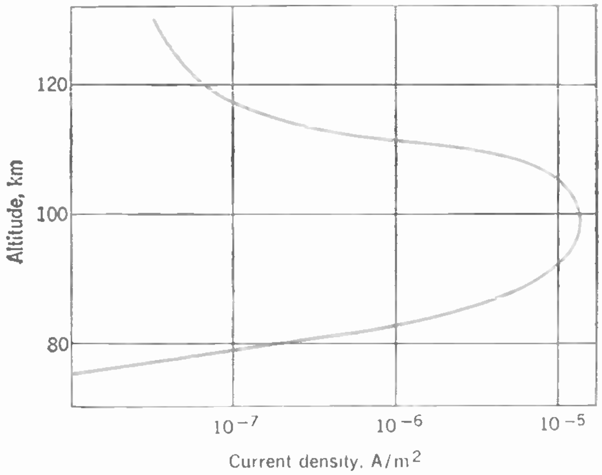


FIGURE 18. This curve shows the current densities that Willis Webb (of the White Sands Missile Range) infers are produced in the ionosphere as the daily passage of the solar heat wave drives charged particles in the ionosphere across the earth's magnetic field lines with a daily, systematic, tidelike rhythm. According to Webb, these Hall-like currents are the primary source of atmospheric electric phenomena. These dynamo currents drive such diverse phenomena as the airglow, aurorae, thunderstorms, fair-weather currents, telluric currents, and magnetic disturbances. Figure 19 shows the paths taken by these ionospheric currents across the earth.

trostatic, gravitational, and aerodynamic forces. This traditional approach to atmospheric electricity assumes that electrical events taking place in the ionosphere, above 50 km, and in the Van Allen belts at still higher levels—out to ten earth radii, in fact, where charge is carried by energetic electrons freed from air molecules by solar and cosmic radiation and by protons and electrons borne to earth by the solar wind—do not influence electrical or weather events in the lower atmosphere. Satellite and rocket studies of the past decade suggest that this assumption may not be justified. Such high-altitude studies indicate, on the contrary, that if we hope someday to successfully manipulate the electrical properties of clouds, and to understand fully the potential environmental consequences of what we are doing, we must view the system earth–atmosphere–space–sun as an electrical entity.

A few years ago I attempted, if not to synthesize, then at least to juxtapose or align, the circuit elements of the planet-wide electrical system in a discussion that appeared in *International Science and Technology*.[29] Just last spring, Willis Webb of the White Sands Missile Range, drawing heavily on data provided by a meteorological rocket-sounding network operated under his direction for the past several years, attempted a comparable but much more quantitative synthesis of global electric structure[30] at the International Conference on Universal Aspects of Atmospheric Electricity in Tokyo. Webb's picture, though quantitative, is still only a first approximation. He acknowledges that it depends heavily on order-of-magnitude estimates, at best, and that it extrapolates far-reaching consequences from rocket soundings obtained from only a few points in the stratosphere.

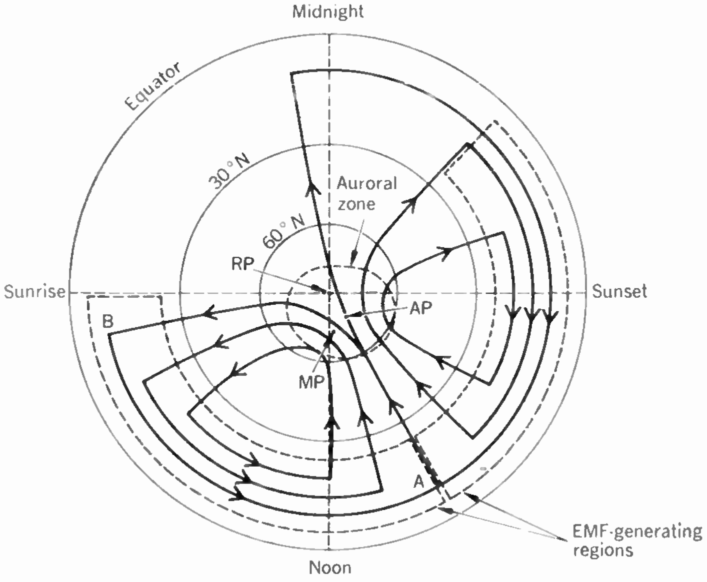


FIGURE 19. This equatorial projection shows the ionospheric current system that drives most atmospheric electric phenomena, according to the scheme proposed by Willis Webb (see Fig. 18 and text). Ionospheric dynamo currents—indicated by lines bearing arrows—are themselves driven by the potential difference set up in the EMF-generating regions as charged particles of differing mobilities are separated by passage of the solar heat wave through the upper atmosphere. Maximum potential develops in early afternoon sector, point A, while lowest potential develops at very early morning, point B; potential difference between A and B is estimated by Webb to be about 400 000 volts. Thunderstorms are a short circuit across the lower atmosphere through which the earth's surface is connected to lowest ionospheric potential, point B. The primary currents in the ionosphere are symmetrical with respect to earth's rotational pole (RP), but asymmetrical with respect to the magnetic pole (MP), auroral pole (AP), and auroral zone. Primary currents crossing magnetic field lines generate a third major electric circuit, according to Webb—the Van Allen radiation belts.

The stratosphere is that portion of the atmosphere extending upward from 10 km, more or less (the exact height of its base depends on latitude, season, and the weather to some extent), to about 40 km. In the stratosphere, the clouds that characterize the lower atmosphere—the so-called troposphere, which has been the traditional focus of meteorological attention—are rare.

Webb's central point, which in his paper is badly obscured by the atrocious syntax for which he has been widely and justly criticized, is quite simple. He believes that meteorologically significant amounts of heat may be delivered to the stratosphere and perhaps ultimately even to the troposphere from much higher atmospheric levels. Conventional meteorological wisdom discounts this as improbable; in it the lower atmosphere is heated exclusively from below, by long-wave infrared radiation and conduction from the earth's surface. Webb thinks that some portion of the heat that may be descending from higher atmospheric levels might be joule heat, produced by intense electric current systems in the ionosphere, and that these may affect thermal and electric energy exchanges throughout the atmosphere.

According to Webb, the rocket sounding data suggest that in addition to the planet-girdling jetstream winds long known to flow at stratospheric altitudes, there also occur daily, systematic, tidelike, vertical air motions. Such vertical motions have long been inferred to occur at higher levels, in the ionosphere, but the stratospheric-level rocket data reveal their possible magnitude for the first time. The vertical motions are perhaps most pronounced at low latitudes in the earth's sunlit hemisphere. They diminish in intensity toward the poles and the nighttime hemisphere. They seem to be powered by solar heating of the upper boundary of the stratosphere. These vertical motions may, if they extend up into the ionosphere, force the charged particles that abound at these levels across the earth's magnetic field lines. There, they may generate ionospheric current systems by a Hall-like effect (Fig. 18), as Chapman and various others first suggested many years ago.[31,32]

Most intense at low latitudes in the E region of the ionosphere, these Hall-like currents may serve as the primary dynamo currents of the earth's atmosphere. They flow east and west toward the early afternoon sector of the globe (Fig. 19) at low latitudes, and return along paths that may be mostly parallel to meridians of longitude, except at higher latitudes. This flow pattern is symmetrical with respect to the earth's rotational poles and equator, but it is asymmetrical with respect to the earth's magnetic poles and magnetic field lines. As a result of this asymmetric flow of charge, secondary currents develop and these flow along magnetic field lines. In Webb's view, these secondary currents flowing parallel to the magnetic field lines constitute the basic particle fields and power sources for the familiar Van Allen "radiation" belts. Currents in the inner Van Allen belt serve to equalize differences in EMF between magnetic hemispheres, whereas currents in the outer Van Allen belt serve instead to equalize differences in potential that develop between the northern and southern auroral zones at subpolar latitudes.

In addition to these two major planetary electric circuits—the rotationally symmetrical primary dynamo circuit in the ionosphere, and the magnetically symmetrical Van Allen circuits—thunderstorms, through their lightning discharges, provide a third circuit. This thunderstorm short circuit connects the earth's surface to the negative potential of the nightside primary dynamo in the ionosphere.

The primary dynamo circuit in the ionosphere thus serves, according to Webb's attempted synthesis of the work of many men, as the source for all of the following features of atmospheric and terrestrial electricity:

1. Airglow, the luminous phenomenon of molecular excitation, which is caused by interactions between charged particles flowing in the ionosphere and radiation belts with the neutral atmosphere.
2. Auroral activity, which is in part caused by charges shuttling between magnetic poles along high-latitude magnetic field lines, currents that are driven by differences in potential created by the E-region dynamo currents in polar regions. Charged particles collected from the solar wind are also sorted and precipitated, as well as accelerated to higher energy, in these regions by the E-region electric fields.
3. Thunderstorm electrification, in which the lightning discharge paths provide an effective short circuit connecting the earth's surface with the primary dynamo circuit in its low-potential region (point B in Fig. 19).
4. Fair-weather electric field, which represents nothing more "normal" or "fundamental" than the difference between the average potential of the atmosphere and the potential established at the earth's surface by the thunderstorm short circuit.
5. Fair-weather electric current, which conducts approximately 1500–2000 amperes to the earth continuously as an auxiliary return path of the primary dynamo circuit.
6. Telluric currents, which flow beneath the earth's surface to complete the global circuit, which includes the fair-weather electric current and the charge transported through thunderstorm lightning.
7. Magnetic disturbances at the earth's surface, which are the result of variations in current intensities in the primary dynamo circuit—variations which are in their turn caused by variations in the thermally induced vertical circulation of the atmosphere, variations in E-region electrical conductivity, and variations in the efficiency with which solar particles are collected in the auroral regions.

Among these diverse but interrelated phenomena, the thunderstorm and fair-weather portions of the global circuits are of most immediate concern to the weather-modification scientist. Let's look at them more carefully.

The spherical condenser theory

Thunderstorms have been assigned diverse roles in electrical schemes proposed by various authors. To some they are the original source of all electrical effects in the lower atmosphere; to others, they play a strictly passive role in the earth's electrical structure. It is true, however, as Webb points out, that whatever the causes and cosmic role of thunderstorm electrification, the physical processes associated with it—and particularly the lightning discharges that are so characteristic a feature of it—do provide a low-impedance path across the lower atmosphere, where the conductivity otherwise is relatively low. Using approximations of continuous lightning-discharge paths across the lower 10 km of the atmosphere, Webb finds that the nominal value of the columnar resistance of the atmosphere may be reduced by at least an order of magnitude. And it may be further reduced at higher altitudes, as the work of others has suggested. Webb's studies and these others indicate, therefore, that thunderstorms can indeed provide a mechanism whereby the earth's surface assumes an electric potential almost, but not quite, equal to that of the ionosphere.

The earth's surface and the ionosphere can thus be thought of as the oppositely charged plates of a huge condenser—earth negative, ionosphere positive (Fig. 20). Estimates of the potential difference across the plates of this condenser have ranged from a few hundred thousand to a half million volts; but recent measurement [33] of the ionosphere's potential indicates that the lower value may be more accurate.

Negative charge is transported downward, ulti-

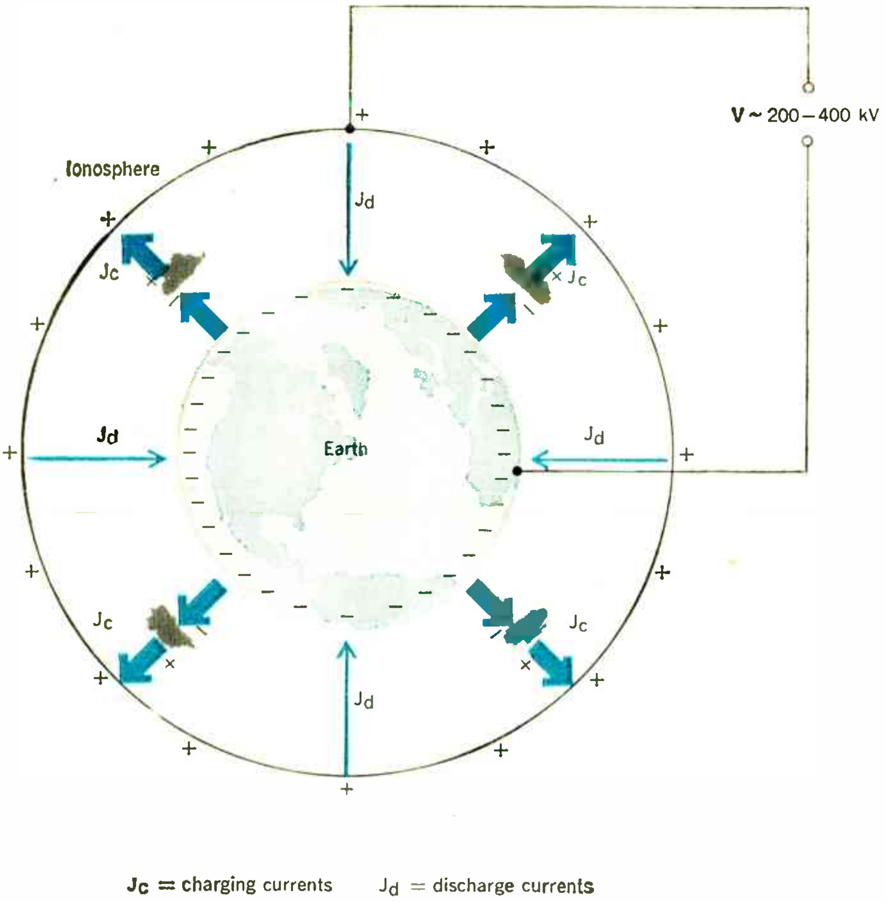


FIGURE 20. The classical "spherical-condenser" theory, illustrated schematically here, seeks to account for the earth's fair-weather electric field and fair-weather electric current. According to the theory, negative charge is concentrated on the earth and positive charge is concentrated in the ionosphere. The potential difference across the intervening, poorly conductive, lower atmosphere accounts for the field; while the current (J_d) that flows in fair-weather regions because the lower atmosphere is not a perfect insulator is the fair-weather current. This current tends to discharge the global condenser. To maintain the global charge difference, currents that can compensate for these diffuse, feeble, but widespread discharging currents are obviously required. It has been generally assumed, on the basis of a fair amount of evidence, which, however, is not definitive, that charging currents (J_c) are generated in all regions occupied by thunderstorms. This theory will be tested in a worldwide, ten-year observational program in atmospheric electricity that is just getting under way.

mately from the ionosphere, by lightning discharges in amounts approximately equal to the amount of positive charge that thunderstorms elicit from the earth via point discharge. Since the negative charge on the earth does not keep increasing indefinitely as a result of this charge transport, but remains sensibly constant at some 1.1×10^{-9} C/m^2, it is generally assumed that the lower-level atmospheric circuit of which the thunderstorm electrostatic generator is a part is closed by the charges being transported in fair-weather areas.[7] In fair weather, when clouds, precipitation, fogs, dust, strong winds, and the like are absent, the electric field near the ground is usually directed as though the earth were negatively charged and the atmosphere positively charged. The mean field intensity at these times is about 130 V/m. During precipitation of any kind, and especially during thunderstorms, the field may reverse direction and reach an intensity of 10 000 V/m or more. According to the spherical-condenser theory, the fair-weather current transfers positive charge to ground and negative charge to the upper atmosphere (Fig. 20). The average density of this fair weather current is 3×10^{-12} A/m^2. Estimates of the average current generated by all the thunderstorms occurring at any moment over the whole earth show that it is about equal in magnitude but opposite in direction to the global fair-weather current, which amounts, over the entire globe, to perhaps 2000 amperes.

This picture of tropospheric electrification attributes worldwide thunderstorm activity to the effects of a global current system in which the positive charge collected by the earth from the ionosphere in fair-weather regions must flow through the crust of the earth to points that lie below thunderstorms. From these locations the positive charge flows back to the ionosphere via the thunderstorm route. Webb points out that measurements of telluric currents in the earth's crust do indeed seem to indicate a systematic flow of charge from the poles toward the equator during the afternoon, and from the equator toward the poles at night. He concludes, therefore, as many other investigators before him have concluded, that the early evening maximum in thunderstorm occurrence around the world is of prime importance in establishing the difference in potential between the E region of the ionosphere and the earth's surface.

This may not be the case, say a pair of Russian investigators.

Watch those stratus clouds

As a result of their investigations, two leading investigators of atmospheric electricity in the Soviet Union have reached a conclusion that seems to challenge the long-assumed supremacy of thunderstorms and ionospheric phenomena in the electrical scheme of things. And they have buttressed their assertion with huge amounts of observational data.

I. M. Imyanitov and E. V. Chubarina base their conclusions on more than 2000 aircraft soundings of meteorological and electrical parameters carried out over three cities in the U.S.S.R. during the International Geophysical Year (1957–1958) and the followup Year of International Geophysical Collaboration. Their report, published in Russian in 1965, was translated into easy-going English for NASA and NSF by the Israel Program for Scientific Translations in 1967.[31]

Imyanitov and Chubarina agree with most other workers that the primary electric generators charging the earth's atmosphere and leading to the appearance of its electric field are located in the troposphere, at altitudes of only a few kilometers. They stress, however, that thunderclouds may not be the only, nor even the most important, generators in this region. The Russian scientists think that stratus clouds, though hundreds of times less active electrically than thunderclouds, may play an at least comparably important part in electrical processes in the atmosphere by virtue of their much greater areal extent. Unlike thunderclouds, which develop most spectacularly in the vertical direction, stratus clouds develop most prominently in the horizontal direction. They cover, at any time, an area 200–300 times larger than the area covered by thunderclouds. Horizontal "decks" of stratus clouds of one kind or another (there are several varieties) constantly cover the sky over about half the globe.

It is especially interesting, from the weather-modification standpoint, that the two stratus varieties that produce the most precipitation along frontal surfaces—nimbostratus and altostratus—also contribute most to the exchange of charge between the earth and the atmosphere. Nimbostratus clouds in particular, the most steadfast and continuous producers of snow and rain of all clouds, although not prolific producers of lightning, contribute most to this charge exchange. Imyanitov and Chubarina calculate that the average amount of charge transferred to earth each year in the northern hemisphere by nimbostratus clouds is about 1.15×10^5 coulombs. The annual average amount of charge transferred to earth by all stratiform clouds in the northern hemisphere could amount to perhaps twice this estimated value—or about 2×10^5 coulombs. Assuming that stratiform clouds in the southern hemisphere (for which there is little or no data since large areas of ocean occupy this hemisphere) carry the same electric charges and have the same distribution as they do in the northern hemisphere, the Soviet scientists estimate that about 4×10^5 coulombs of charge could be transferred annually to earth by stratiform clouds. Comparing this figure with the absolute value of the proper charge of the entire earth, which they estimate does not exceed 3.5×10^5 coulombs, they conclude that the charge of the earth must depend not only on the electrical activity of thunderclouds but also—and quite clearly to a rather overwhelming degree, if their estimates are valid—on the electrical activity of stratified clouds as well.

These studies suggest that the classical spherical-condenser theory, in which the thunderstorm electrostatic generator plays the key role in maintaining the charge on the opposing plates—the earth's surface and the ionosphere—may be a rather gross oversimplification. Because stratus clouds provide such an extensive electrical cover, an electrical structure in which the fair-weather field is maintained according to the two-layer scheme, troposphere-earth and troposphere–ionosphere, may be more accurate. Charge is no doubt exchanged between the troposphere and ionosphere, say the Russians, but the current directed to the upper layers of the atmosphere may be small in contrast to the charge flowing between the earth and stratus-cloud levels of the troposphere. We may be dealing with two somewhat leaky spherical condensers instead of one, and with rather more complex atmospheric circuitry than shown in Fig. 20.

The Russian work thus shifts the focus of electricity generation back to the lower several kilometers of the atmosphere, away from the ionosphere.

Toward a definitive theory

Both regions must be investigated, of course, if the comprehensive theory needed to guide electrical weather modification attempts is to be built.

Like most differences of opinion, these will be settled, someday, by more observations. The fundamental assumptions of the classical spherical-condenser theory, at any rate, are easy enough to check, at least in principle.

Since the charges carried by the earth and the atmosphere are presumably determined by the worldwide relationship between the thunderstorm charging current and the fair-weather discharging current (Fig. 20), any worldwide intensification in thunderstorm activity should act to increase the charges on the global condenser plates, and consequently to increase the potential difference between the earth and ionosphere. This effect should be detectable as an increase in the electric field intensity in the intervening atmosphere. Conversely, any worldwide decrease in thunderstorm activity should show itself as a corresponding decrease in the potential difference and field intensity in the atmosphere. Either such change should be detectable promptly, because both the earth and the ionosphere are highly conductive. Charges delivered to either region, by either the charging or discharging currents, should spread over the surface of that condenser plate practically instantaneously. As a result, variations in the potential difference across the condenser plates should occur over the whole globe essentially simultaneously, and they should be detectable by suitably widely distributed simultaneous measurements of the atmosphere's electrical parameters. Making such measurements, in an attempt to test the spherical-condenser theory of atmospheric electricity definitively, is one major goal of an ambitious Atmospheric Electricity Ten-Year Program that is just getting under way. If this and the other equally ambitious goals of this global program are achieved, we will be a lot closer to the definitive theory needed to explain the total cloud electrification process—that theory which the NAS-NRC Panel on Weather and Climate Modification thought might be central to the search for effective weather-control technologies.

The rest of the story

Future installments will explore possibilities offered by the Atmospheric Electricity Ten-Year Program and view these within the framework provided by a still broader Global Atmospheric Research Program, which the meteorological community is now tooling up to pursue through the decade of the 1970s. We shall also take a closer look at ongoing weather-modification activities, their scientific and social roots, and the directions they may take in the future, paying particular attention to the problems and opportunities these activities may provide for the electrical community. Other—let's for the moment call them "human"—dimensions of the weather-modification equation also have strong implications for the electrical community, in whose hands the future of weather modification may lie. The final installment will also deal with these.

REFERENCES

1. "Weather and climate modification, problems and prospects" (two volumes), Publication No. 1350, National Academy of Sciences–National Research Council, Washington, D.C., 1966.
2. "Weather modification, ninth annual report, 1967," Publication No. 68–21, National Science Foundation, U.S. Government Printing Office, Washington, D.C., 1968.
3. Coroniti, S. C., ed., *Problems of Atmospheric and Space Electricity*. New York: Elsevier, 1965.
4. Vonnegut, B., Smallman, C. R., Harris, C. K., and Sykes, W. G., "Technique for the introduction into the atmosphere of high concentrations of electrically charged aerosol particles," *J. Atmospheric Terrest. Phys.*, vol. 29, pp. 781–792, 1967.
5. Kasemir, H. W., "Lightning suppression by chaff seeding," Final Report to USAECOM under Project Order 67–95906, Amendment 68–95920, Environmental Science Services Administration, Boulder, Colo., June 1968.
6. Workman, E. J., "The production of thunderstorm electricity," *J. Franklin Inst.*, vol. 283, pp. 540–557, June 1967.
7. Malan, D. J., "Physics of the thunderstorm electric circuit," *J. Franklin Inst.*, vol. 283, pp. 526–539, June 1967.
8. Vonnegut, B., "Some facts and speculations concerning the origin and role of thunderstorm electricity," *Meterol. Monographs*, vol. 5, pp. 224–241, Sept. 1963.
9. Sartor, J. D., "The role of particle interactions in the distribution of electricity in thunderstorms," *J. Atmospheric Sci.*, vol. 24, pp. 601–615, Nov. 1967.
10. Mason, B. J., "Charge generation in thunderstorms," in *Problems of Atmospheric and Space Electricity*, S. C. Coroniti, ed. New York: Elsevier, 1965, pp. 239–254.
11. Brook, M., "Thunderstorm electrification," in *Problems of Atmospheric and Space Electricity*, S. C. Coroniti, ed. New York: Elsevier, 1965, pp. 280–283.
12. Kasemir, H. W., "The thundercloud," in *Problems of Atmospheric and Space Electricity*, S. C. Coroniti, ed. New York: Elsevier, 1965, pp. 215–230.
13. *Proc. 1968 Internat'l Conf. on Cloud Physics*, American Meteorological Society, Boston, Mass., pp. 597–677.
14. Pruppacher, H. R., Steinberger, E. H., and Wang, T. L., "On the electrical effects that accompany the spontaneous growth of ice in supercooled aqueous solutions," *J. Geophys. Res.*, vol. 73, pp. 571–584, Jan. 1968.
15. Latham, J., "The influence of electrical forces on cloud physical phenomena," *Proc. Fourth Internat'l Conf. on Universal Aspects of Atmospheric Electricity*, S. C. Coroniti and J. Hughes, eds. New York: Gordon and Breach, in press.
16. Evans, W. H., "The measurement of electric fields in clouds," *Pure and Appl. Geophys.*, vol. 62, pp. 191–197, 1965.
17. Vonnegut, B., Moore, C. B., Stout, G. E., Staggs, D. W., Bullock, J. W., and Bradley W. E., "Artificial modification of atmospheric space charge," *J. Geophys. Res.*, vol. 67, pp. 1073–1083, Mar. 1962.
18. Vonnegut, B., Moore, C.B., Semonin, R. G., Bullock, J. W., Staggs, D. W., and Bradley W. E. "Effect of atmospheric space charge on initial electrification of cumulus clouds," *J. Geophys. Res.*, vol. 67, pp. 3909–3922, Sept. 1962.
19. Vonnegut, B., and Moore, C. B., "Preliminary attempts to influence convective electrification in cumulus clouds by the introduction of space charge into the lower atmosphere," in *Recent Advances in Atmospheric Electricity*. New York: Pergamon Press, 1959, pp. 317–331.
20. Fitzgerald, D. R., "Measurement techniques in clouds," in *Problems of Atmospheric and Space Electricity*, S. C. Coroniti, ed. New York: Elsevier, 1965, pp. 199–211.
21. Workman, E. J., "The possible role of ammonia in thunderstorm electrification," in *Proc. 1968 Internat'l Conf. on Cloud Physics*, American Meteorological Society, Boston, Mass., pp. 653–656.
22. Workman, E. J., "Thunderstorm electricity," in *Problems of Atmospheric and Space Electricity*, S. C. Coroniti, ed. New York: Elsevier, 1965, pp. 296–303.
23. Moore, C. B., Vonnegut, B., Vrablik, E. A., and McCaig, D. A., "Gushes of rain and hail after lightning," *J. Atmospheric Sci.*, vol. 21, pp. 646–665, Nov. 1964.
24. Vonnegut, B., "Luminous phenomena in nocturnal tornadoes," *Science*, vol. 153, pp. 1213–1220, Sept. 9, 1966.
25. Carstoiu, J., "A tentative synthesis of electrohydrodynamic phenomena in the earth atmosphere—theoretical developments," in *Proc. Fourth Internat'l Conf. on Universal Aspects of Atmospheric Electricity*, S. C. Coroniti and J. Hughes, eds. New York: Gordon and Breach, in press.
26. Reiter, R., *Felder, Ströme und Aerosole in der Unteren Troposphäre*. Darmstadt: Dr. Dietrich Steinkopff Verlag, 1964.
27. Cobb, W. E., "Evidence of a solar influence on the atmospheric electrical elements at Mauna Loa Observatory," *Monthly Weather Rev.*, Dec. 1967.
28. Chalmers, J. A., *Atmospheric Electricity*, 2nd ed. New York: Pergamon Press, 1967.
29. Tilson, S., "Environmental electricity," *Internat'l Sci. Technol.*, vol. 3, pp. 26–38, Jan. 1964.
30. Webb, W. L., "The source of atmospheric electrification," in *Proc. Fourth Internat'l Conf. on Universal Aspects of Atmospheric Electricity*, S. C. Coroniti and J. Hughes, eds. New York: Gordon and Breach, in press.
31. Chapman S., and Bartels, J., *Geomagnetism* (two volumes). Oxford: Clarendon Press, 1940.
32. Chapman, S., "The electrical conductivity of the ionosphere: a review," *Nuevo Cimento*, vol. 4, suppl., pp. 1385–1412, 1956.
33. Muhlheisen, R. P., "Recent potential gradient and ion measurements on the Atlantic Ocean in 1965," in *Proc. Fourth Internat'l Conf. on Universal Aspects of Atmospheric Electricity*, S. C. Coroniti and J. Hughes, eds. New York: Gordon and Breach, in press.
34. Imyanitov, I. M., and Chubarina, E. V., *Electricity of the Free Atmosphere*. Jerusalem: Israel Program for Scientific Translations, 1967. (Available from the U.S. Department of Commerce, Clearinghouse for Federal Scientific and Technical Information, Springfield, Va. 22151.)

Seymour Tilson, staff writer for IEEE Spectrum, was born and educated in New York City. Following two years as an aeronautical engineering student at the Polytechnic Institute of Brooklyn, he transferred to Brooklyn College, from which he received the bachelor's degree in geophysics in 1949. He received the master's degree, also in geophysics, from New York University in 1957, and has completed all requirements at Columbia University for the Ph.D. degree in geology, which awaits defense of the dissertation. Mr. Tilson came to IEEE Spectrum from the Travelers Research Center, Inc., Hartford, Conn., a "think tank" specializing in the environmental and management sciences. As director of scientific liaison in the office of the president, he divided his time between writing duties and substantive research in the areas of environmental science, policy, and politics. He joined Travelers in 1967, after six and a half years as an associate editor of the interdisciplinary professional magazine, International Science and Technology. Before entering the broader field of science-related journalism, Mr. Tilson taught geology for eight years at New York University, Hofstra University, and Brooklyn College. Prior to that, in the early 1950s, he was a design engineer for various industrial companies.

Pumped storage—the handmaiden

The power companies of the New England states, where fossil-fuel cost is high, have determined that low-variable-cost nuclear generation is desirable for this region. The construction of pumped-storage hydro plants, favored by the topography of the area, will offset the relatively high fixed costs of nuclear plants by maximizing the capacity-use factor of the nuclear facilities and thereby minimizing the total cost per unit of energy output. This combination provides efficient base load, economical peaking capacity, and effective spinning reserve, along with many recreational and other social benefits.

Rapid growth of load on electric utility systems in the United States is imposing the necessity for adding large blocks of generating capacity that must supply both the long-hour base load and the morning and evening peaks. Historically, utilities have met the base loads with fossil-fired steam generating plants. The peak loads have been met with run-of-river, limited-pondage hydro installations located on the major rivers, as well as older, less-efficient steam units. As utilities pursued this method of expansion, two problems that developed forestalled further expansion of this type. Utilities such as those in the northeastern region of the United States are remote from sources of fossil fuel and are faced with rising rail and water transportation costs. At the same time, development of hydro-generating facilities on the rivers is reaching the point of full utilization of the natural head.

Faced with these problems and the need to reduce the overall cost of power, the systems turned to nuclear-fueled steam generating stations to meet the base load. It is characteristic of today's nuclear power installation that the initial cost is higher than the initial cost of an equivalent size fossil-fuel plant. The operating cost of a nuclear plant, however, is substantially lower in those sections of the United States where transportation expense represents a large proportion of fossil-fuel cost. The result is a total power cost for nuclear generation that compares favorably with fossil-fuel power generating costs. The low incremental energy cost of power from nuclear plants makes the consideration of pumped-storage hydropower very attractive.

Pumped-storage hydroelectric energy is not new in the United States. The first large-scale U.S. hydroelectric pumped-storage plant was constructed in 1928 in New England, where "run-of-river" hydromechanical, and later hydroelectric, energy have been an integral part of the economy for over 300 years. The Rocky River Pumped Storage Plant [1] was built on the Housatonic River, a major river in the northeast section of the United States, as another step in the integration of this source of energy into the power needs of an industrialized territory and in the full utilization of an inexhaustible natural source.

It formed Lake Candlewood in western Connecticut, which has developed into an important recreational area. With a shoreline of some 60 miles (96 km), it is the largest lake in Connecticut. Although the Rocky River Plant's capacity of 32 000 kW can be utilized on a day–night schedule, the plant's principal function has been to provide seasonal reserve in order to firm up the capacity of the more or less run-of-river plants further downstream. For each kilowatthour used in pumping during the spring freshets, when water would otherwise be lost over the dams at downstream plants, 1.14 kWh are recovered by the gradual release of water from Lake Candlewood during the normally drier periods of summer and fall.

Advantages of pumped-storage hydro

Among the numerous advantages attributable to the present-day pumped-hydro plant is the fact that it pro-

# of nuclear power

Where the terrain permits its use, the pumped-storage hydro plant can be valuable for providing an improved plant factor for the large-base-load nuclear units now being installed or planned by many electric utility companies

Sherman R. Knapp *Northeast Utilities Service Company*

vides, by means of its pumping load, a better plant factor for the large-base-load nuclear units that are being installed today by many utilities. In addition, the pumped-hydro plant can be designed in such a manner that it can be brought on the line rapidly enough to be classed as a reliable source of spinning or assured reserve capacity at all times. Further, its ability to accept or reject load almost instantaneously makes it much more flexible than other types of generation, peaking or base load. It can follow the sharp peak-load fluctuations that occur on a minute-to-minute basis on any large system. This same ability to follow the peaks of the system load permits more uniform and efficient loading on the fossil-fueled and nuclear-fueled units that are operating as base-load units in conjunction with the pumped-hydro plant.

Finally, the installation of relatively small pumped-hydro generating units as a means of meeting the system peaks permits an overall reduction in the installed reserve required on the system. This reduction is possible because of the superior reliability of the pumped-hydro unit as related to alternative types of peaking generation.

These many advantages, along with extensive engineering economic studies, resulted in the planned construction of a 1000-MW pumped-storage hydro development at Northfield Mountain in Franklin County, Mass. This plant, when completed in 1971, will be the largest installation of its kind in the world.

The Northfield Mountain Project

The design of the Northfield Mountain Pumped Storage Project is being carried forward on the basis of developing a generating plant capable of high-speed response. The units will be designed in such a manner that they can be brought up to speed from a complete shutdown, synchronized to the system, and fully loaded, all in less than three minutes' time. Thus, they can be fully loaded in a shorter period than a modern large fossil-fueled or nuclear-fueled steam unit can be brought from the minimum load to full load, even though the latter may already be synchronized to the system. If this desired flexibility of operation is realized, these units can be classified as a portion of the New England systems' assured reserve capacity, even when they are at a standstill. During periods of emergency, when a reserve is limited, it is proposed that the Northfield Mountain units will be operated in the "spinning-in-air" mode, rotating in the generate direction. When operated in this manner, the units can be brought to full load in less than one minute. Thus, the pumped-hydro plant, properly designed, can produce much greater flexibility than any other form of peaking or base-load capacity.

To obtain the degree of speed regulation and governor stability necessary to meet the conditions of assured reserve and to be capable of responding to the rapid changes occurring in system demand, the pressure conduit length of the Northfield Mountain plant has been kept to a minimum. As a consequence, the powerhouse has been located underground adjacent to the upper reservoir. The resulting long tailrace tunnel will be subject to surges, which will be controlled within necessary limits by the underground surge chambers located a short distance downstream of the pump turbines.

Method of operation. During the peak-load hours of the day, the Northfield Mountain Project will operate as a generator by releasing the water from its upper reservoir on top of the mountain through the turbines to its lower reservoir, the Turners Falls Pond. The impoundment of the entire usable storage of the project's upper reservoir will be made possible in this lower reservoir by the installation of bascule-type spillway gates on the existing Turners Falls Dam. During off-peak hours, the water previously used for generation and subsequently stored in the lower reservoir will be returned to the upper reservoir by pumping it via the same hydraulic conduits utilized in the generating cycle. The project will depend upon other generation sources to provide the power required to operate the units as pumps.

Inasmuch as the operation will not be dependent upon the flows of the Connecticut River and since the project depends entirely upon nonproject generation sources for its production of power and energy, the project can be classified as a pure pumped-storage type of system.

The project will operate within the confines of its own upper reservoir volume and the volume provided specifically for this operation within the Turners Falls Pond. The usable storage for generation above the minimum drawdown elevation of 938 feet (275 meters) will be 12 750 acre-feet (15.7 million cubic meters) resulting in some 8500 MWh of stored electric energy. For normal operation, it is expected that this reservoir will fluctuate between these elevations, retaining always in storage some 2500 MWh for assurance of reserve and for use during system emergencies. The project will operate generally on a modified weekly cycle. The upper reservoir is large enough to allow it to refill either on a daily basis or on a weekly cycle in order that it may take advantage of the lower weekend pumping costs.

The lower reservoir will not be drawn below elevation 176 feet (53.7 meters), which is the present self-imposed limit of drawdown during the summer months of the Turners Falls Project. The maximum crest gate elevation of 185 feet (56.5 meters) will not be reached during normal operations, inasmuch as sufficient storage volume will always be held ready to accept the full

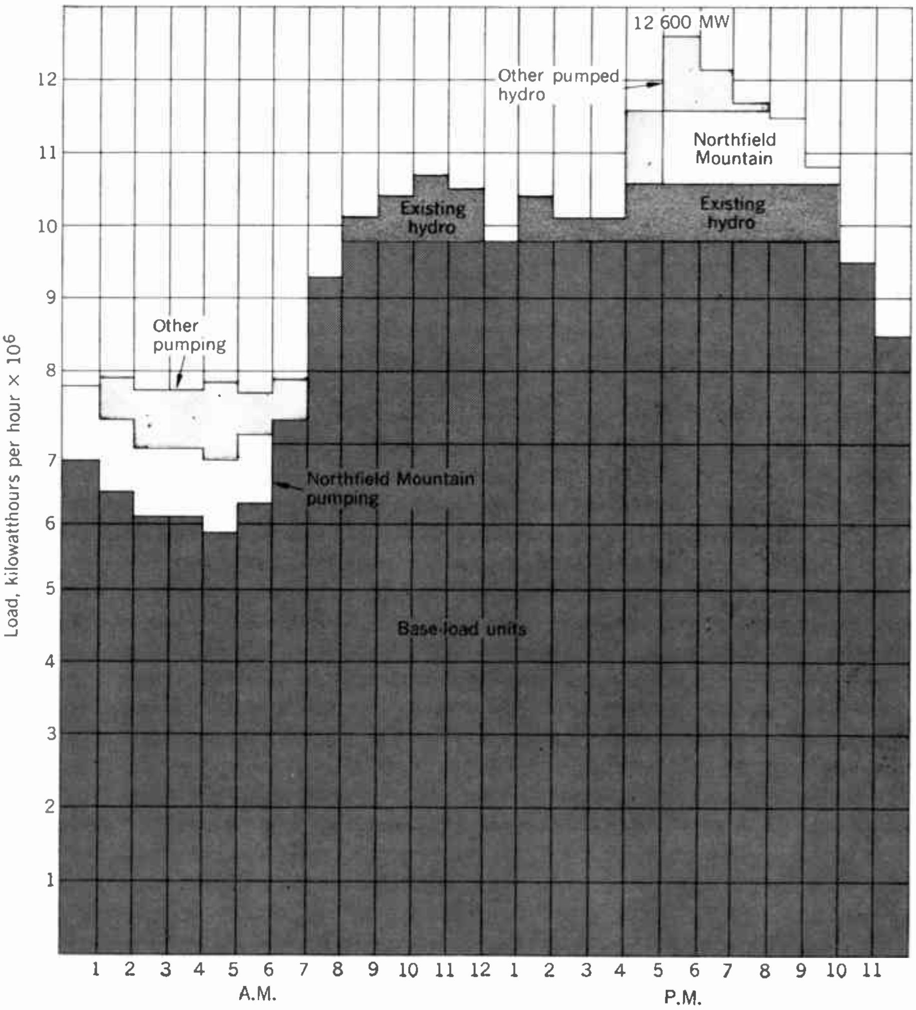


FIGURE 1. New England load and generation for peak day in December 1973.

drawdown of the upper reservoir, should this be required by emergency conditions.

The site is such that the construction of the large upper reservoir is economically feasible and, as a result, the generating capacity (1000 MW) will not be limited by the amount of stored electric energy available. As previously noted, a substantial portion of the upper reservoir is being reserved for operation only during system emergencies. Studies of project operation during a simulated emergency showed that the total available energy (8500 MWh) would be sufficient to maintain a firm capacity of 1000 MW during the first day of an outage, and the system would need to rely only partially on other system reserves during the remaining weekdays. The studies indicated that the system could survive this type of emergency until the weekend, at which time the upper reservoir would be recharged by normal pumping.

The economics of pumped-storage hydro

The economics of high fixed-cost and low variable-cost generation, such as that produced by nuclear plants being installed by the New England utilities, dictate long hours of use in order to maintain a low overall cost of energy. Thus, the number of plants of this type that can be installed on any given system is a function of the minimum load existing on that system at any time and the magnitude of the portion of the on-peak load that can be sustained for an appreciable number of hours during the day. Utilization of pumped-hydro plants, such as the Northfield Mountain plant, to meet the peak-load hours, imposes an appreciable amount of additional pumping load on the system during the early morning hours when the load is at its minimum. This additional minimum load permits an increase in the amount of capacity that can be base-loaded 24 hours of the day and minimizes the cycling that other base-load units must undergo.

This situation is illustrated in Fig. 1, which shows the estimated New England load and generation during the peak day in December 1973. Northfield Mountain and an additional pumped-hydro plant of the same size, together with existing hydro in New England, are assigned to the peaks occurring during the daytime hours. Nuclear and fossil units carry the rest of the load occurring throughout the day. Between the hours of 12 midnight and 7 A.M., the minimum load period of the day, the pumping necessary to refill the upper reservoir of Northfield Mountain and its sister plant occurs. As can be seen in Fig. 1, the minimum load on the New England system has been increased from slightly less than 6000 MW to approximately 7750 MW, permitting an additional 1750 MW of base-load capacity to operate through the full 24 hours of the day.

It can also be seen that the magnitude of the load between the minimum load and the peak load carried by the hydro units has been reduced from 3750 MW to only 2000 MW. This load must be met by base-load units that are capable of being operated at or near minimum load conditions during the early morning hours.

Operation of this type makes maximum use of the highly efficient base-load units and minimizes generation produced from inefficient fossil-fueled steam units. Thus, it tends to optimize the overall use of our natural resources (nuclear and fossil fuels) by the process of storing energy in a reservoir at times when such energy can be produced on the most efficient base-load units.

During the course of the many studies conducted in New England to determine the optimum methods of meeting the growth of load, both in the form of base-load generation and peaking generation, planning engineers have made economic comparisons of the most commonly used types of peaking installations—that is, the gas turbine, the quick-cycling steam unit, and the pumped-hydro plant. The studies have indicated that New England topography is such that pumped-hydro installations can be built at a lower cost per kilowatt than either the gas turbine or the quick-cycling steam unit. They have determined that the pumped-hydro plant, operating in conjunction with the base-load nuclear units, produces the lowest overall production cost for the system. Furthermore, only the pumped-hydro unit is sufficiently flexible to permit it to be classed as assured reserve in New England.

A résumé of the results of the economic comparisons is given in Tables I and II. This form of comparison was selected in order to show clearly the effect of pumped hydro working in conjunction with nuclear generation as a source of pumping energy. These figures show, for the years 1973 and 1977, the comparison of the three types of peaking units in terms of annual costs, together with the annual costs of the nuclear base-load units operating in conjunction with each type of peaking unit. Energy penalties accruing against the gas-turbine and pumped-hydro combinations have been omitted. The applicable energy penalties would have amounted to about $1.2 million for the gas turbines and $1.0 million for the pumped hydro in 1973. The corresponding values would be considerably less in 1977. Omitting the energy penalties does not affect the relative cost relationship between generation combinations, as tabulated in mills per kilowatt-hour.

As can be seen from the basic data in Tables I and II, quick-cycling peaking steam units can be installed in New England for about $99/kW. Gas turbines, in large sizes, can be installed for approximately $85/kW. However, in New England it is not difficult to find sites where large pumped-hydro plants can be developed for costs at about $70 to $75/kW. This lower installed cost, together with the low fixed-charge rate attendant to a plant with a long life, such as the hydro plant, results in appreciably lower annual fixed costs for the pumped-hydro plant than for its alternative choices.

The unit costs estimated for 1973 show that the combined cost per kilowatthour for the nuclear units, operated in conjunction with the pumped-hydro units, is 4.6 mills/kWh. The corresponding combined figure for nuclear and quick-cycling steam units is 4.7 mills/kWh; the combination of nuclear units with gas turbines produces an overall cost of 4.81 mills/kWh. Thus, the use of pumped-hydro generation as a means of meeting the system peaks is more economical than the use of either gas turbines or quick-cycling steam units (see Table I).

A similar analysis of the year 1977 (see Table II)

I. Cost comparison of generation combinations on a large-area load based on 1973 load in New England states

	Combination I			Combination II			Combination III		
	Base-Load Nuclear Units, 1972	Gas Tur-bines	Total	Base-Load Nuclear Units, 1972	Steam Peaking	Total	Base-Load Nuclear Units, 1972	Pumped Hydro	Total
Basic data:									
Installed capacity, MW	2 932	1 000	3 932	2 932	1 020	3 952	2 932	1 000	3 932
Installed cost (× 10^6)	$436.0	$85.0		$436.0	$99.0		$436.0	$72.0	
Fixed-charge rate, %	12.5	12.0		12.5	12.0		12.5	9.5	
Generation, kWh × 10^6	21 491	26	21 517	21 491	3 931	25 422	21 694	140	21 834
Load factor, %	83.7	0.3		83.7	44.0		84.5	1.6	
1973 production and investment costs, millions									
Fuel expense	$30.3	$ 0.3	$ 30.6	$30.3	$13.4	$ 43.7	$30.6	$ —	$ 30.6
Operating expense	8.0	0.2	8.2	8.0	1.4	9.4	8.0	0.5	8.5
Fixed charges	54.5	10.2	64.7	54.5	11.9	66.4	54.5	6.8	61.3
Total annual costs	$92.8	$10.7	$103.5	$92.8	$26.7	$119.5	$93.1	$ 7.3	$100.4
Mills/kWh			4.81			4.70			4.60

II. Cost comparison of generation combinations on a large-area load based on 1977 load in New England states

	Combination I			Combination II			Combination III		
	Base-Load Nuclear Units, 1972	Gas Tur-bines	Total	Base-Load Nuclear Units, 1972	Steam Peaking	Total	Base-Load Nuclear Units, 1972	Pumped Hydro	Total
Basic data:									
Installed capacity, MW	2 932	1 000	3 932	2 932	1 020	3 952	2 932	1 000	3 932
Installed cost (× 10^6)	$436.0	$85.0		$436.0	$99.0		$436.0	$72.0	
Fixed-charge rate, %	12.5	12.0		12.5	12.0		12.5	9.5	
Generation, kWh × 10^6	20 105	26	20 131	20 105	2 234	22 339	22 048	1 340	23 388
Load factor, %	78.3	0.3		78.3	25.0		85.8	15.3	
1977 production and investment costs, millions									
Fuel expense	$25.6	$ 0.3	$25.9	$25.6	$ 7.6	$ 33.2	$27.9	$ —	$27.9
Operating expense	8.0	0.2	8.2	8.0	1.4	9.4	8.0	0.5	8.5
Fixed charges	54.5	10.2	64.7	54.5	11.9	66.4	54.5	6.8	61.3
Total annual costs	$88.1	$10.7	$98.8	$88.1	$20.9	$109.0	$90.4	$ 7.3	$97.7
Mills/kWh			4.91			4.88			4.18

shows that the difference in total cost per kilowatthour for the pumped-hydro plant over the other forms of peaking generation is even more pronounced. This results from the greater utilization of the pumped-hydro plant, and the fact that the load increase between 1973 and 1977 was presumed to have been met by additional nuclear installations. As a consequence, a greater proportion of the pumping energy for the pumped-hydro unit comes from the low-incremental-cost nuclear units in 1977 than in 1973.

Of considerable significance is the effect on load factor of the various types of peaking units as the load grown from the 1973 level to the 1977 level, and as additional base-load nuclear units with their extremely low incremental fuel cost are added to the system. The gas turbines, with a high fuel rate, would operate at a very low load factor (0.3 percent) in 1973 and in exactly the same manner in 1977. The quick-cycling steam units, with a reasonably good heat rate, would operate at a load factor of 44 percent in 1973. However, with additional low-cost nuclear units available to meet the loads by 1977, the load factor of the quick-cycling steam units will be reduced to 25 percent. In the case of pumped hydro, in 1973 the bulk of the pumping power will come from the best of the fossil-fueled steam units; and the pumped-hydro plant is expected to operate at a load factor of only 1.6 percent. By 1977, however, with an appreciable influx of additional low-incremental-cost nuclear units on the system, the pumped-hydro plant will operate at a load factor of 15.3 percent.

The effect on the load factor of the pumped-hydro plant, as additional nuclear generation is added to the system, illustrates quite effectively the "negative obsolescence" of the pumped-hydro plant. As improvements in the art of developing low-variable-cost nuclear plants take place, the cost of pumping energy for a pumped-hydro plant decreases. While other forms of peaking units are depreciating and becoming obsolete, the pumped-hydro plant tends to become more efficient. The result is the exact opposite of normal obsolescence.

Further studies by the planning engineers of the New England utilities have been conducted to establish the optimum amount of peaking capacity that should be available to the New England utilities to meet their load at the lowest overall cost. Information available to date from these continuing-type studies indicates that the New England systems, as a group, can economically utilize an amount equivalent to approximately 18 percent of their total peak load in the form of pumped-hydro generation.

Pumped-storage plants need not be limited to hydro application. Recent developments in the area of pumped-storage compressed-air techniques are being watched with interest. This method of energy storage combines the conventional gas turbine electric generator with a large air-storage cavity. During off-peak periods, air is stored in the cavern by means of the compressor of the gas turbine driven by the electric generator operating as a motor. This is done with base-load energy. During peak periods the air is released, mixed with fuel, and expanded by combustion to drive the turboelectric generator. The efficiency of the cycle derives from operating the turboelectric generator uncoupled from the compressor section but fed air stored at off-peak periods with low-cost energy. One obvious advantage of a compressed-air plant is its possible location directly at a load center. However, the natural topography of the New England area is ideally suited to pumped-hydro plants and several are planned between now and 1990. (See Fig. 2 for locations of future base-load nuclear and pumped-storage plants.)

Social benefits of pumped-storage hydro

The foregoing discussion has dealt with the high efficiency of economic resource allocation that results by combining nuclear steam generation with pumped-storage hydro. Capital resources are best utilized by virtue of the high plant factors available, and limited fuel supplies are effectively conserved by the ability to store the output of the best steam units.

However, the planners of the Northfield Mountain Project, while primarily interested in the most economical mode of generation, are also conscious of the social benefits that can be derived from well-planned hydro projects. Their earlier pumped-storage hydro projects (the Rocky River plant previously discussed) provides one of the major recreational areas in Connecticut. Water from the project also served to bolster water supplies of the surrounding communities during the five-year period of drought that occurred in the early 1960s.

Because of its past experience, and with the possibility of social benefits in mind, Northeast Utilities proposed a comprehensive multipurpose development of the natural resources in the Northfield Mountain area by including augmentation of the public water supply in addition to recreational and conservation uses.

Water supply. Traditionally, the major rivers of New England have been used by industry both as a source of process water and as a natural waste-disposal system. The most recent drought, of five years' duration, caused many communities to critically reexamine their water resources. The projected population growth and resultant increase in the demand for water require careful planning to assure that the best possible use be made of all existing water supplies.

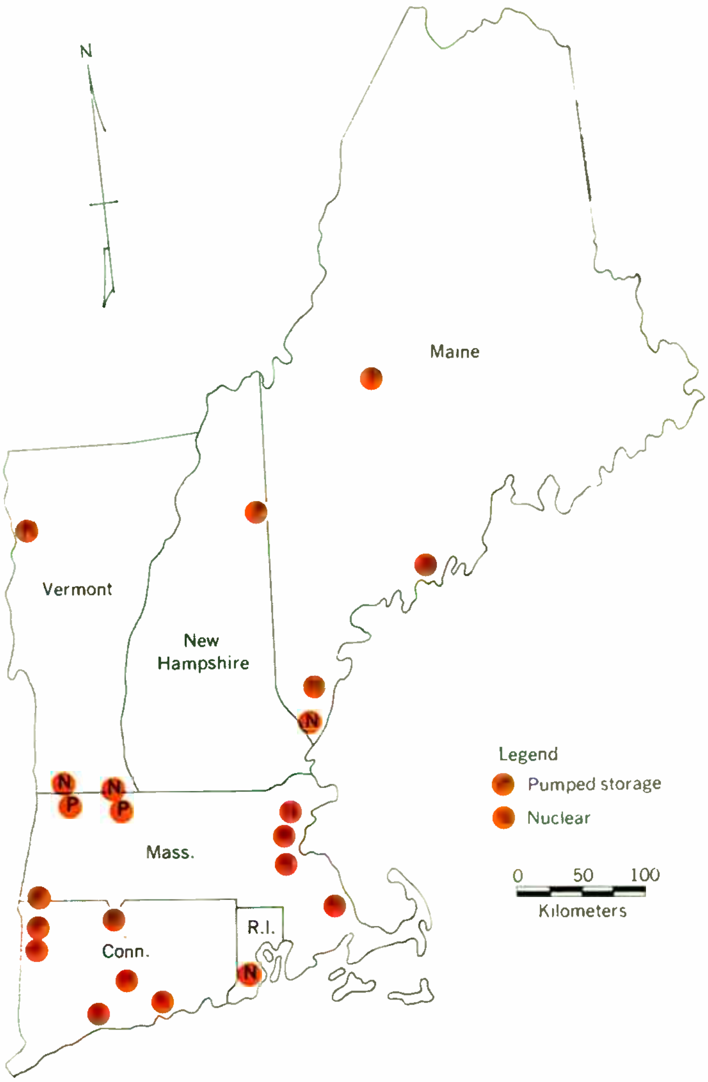


FIGURE 2. Proposed nuclear and pumped-storage plants for the New England states, 1973–1990.

Recognizing this need to conserve water, the participating companies have proposed to utilize the pumping and storage facilities of the Northfield Project to supply Connecticut River water at times of high freshet river flows, primarily in the spring, to the Quabbin Reservoir, which supplies the greater Boston, Mass., area with water (see Fig. 3). The proposal is predicated upon an estimate of an average project pumping rate of 10 000 cubic feet, or about 75 000 gallons, per second (283 m^3/s) for 1.4 hours per day for a total of 50 million cubic feet (1.42 million cubic meters) per day. The pumping will be limited to periods when the Connecticut River flows are high enough (above 15 000 ft^3/s) that water would be spilling over the Turners Falls Dam; that is, it would be excess flow to the Turners Falls Project, to the downstream hydro development at Holyoke, Mass., and to other downstream facilities. A small portion of this excess water, which would otherwise be lost directly to the sea, will be retained by being pumped to the upper reservoir during the limited period of time between the normal generating and pumping cycles or during the normal pumping

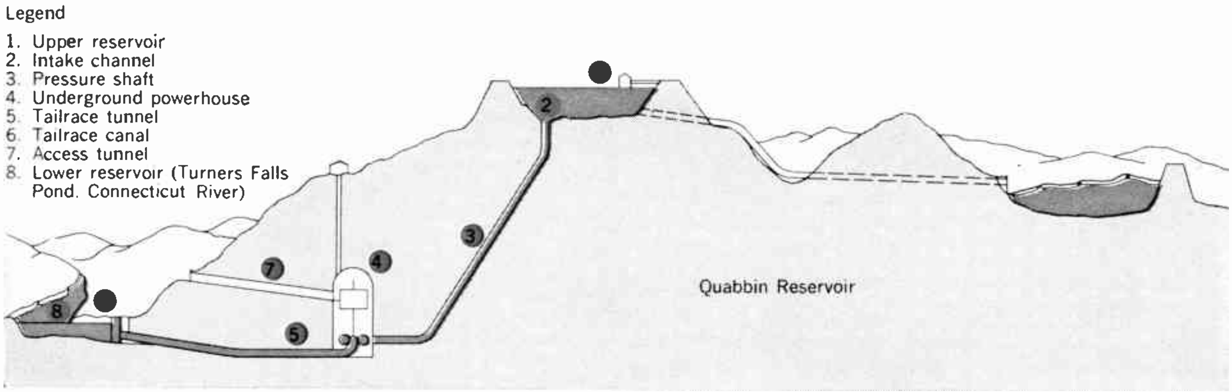


FIGURE 3. Cutaway drawing of Northfield Mountain Pumped Storage Project.

cycle of the Northfield Mountain Project.

To retain this additional volume of water on a daily basis it will be necessary to increase the dam and dike height by 4 feet (1.22 meters); the cost will be paid for by the operators of Quabbin. A separate water supply intake structure will also be required. This excess river water will then flow by the almost 500-foot (154-meter) gravity head available to the Quabbin Reservoir some 10 miles (16 km) to the southeast. Based on an average of about 70 days per year when the river flows are above 15 000 ft^3/s (424 m^3/s), an annual volume of some 26.3 billion gallons (100 million cubic meters) of water can be transferred, with a resultant increase in water level at Quabbin of approximately 3 feet (1 meter). This volume represents about one quarter of the yearly Quabbin usage.

Recreation. More leisure time, higher levels of income, and a constantly increasing population have placed a strain on existing recreational facilities. The Connecticut River and its valley, located reasonably close to large population concentrations, have been called one of the most beautiful areas in the country; however, much of the river's shoreline remains undeveloped.

The Northfield Mountain Pumped Storage Project will include extensive development of a portion of the valley for public recreational purposes. The developers have pledged $1.35 million toward the implementation of a $4 million recreational program; plans are for the balance to be paid by others, including the State of Massachusetts.

The general recreation plan calls for nine areas, with such facilities as swimming pools, camping sites, picnic areas, cabins, bridle paths (usable by snowmobiles in winter), boat launch sites, and a marina.

Other social benefits. Studies indicate that transmission lines play an important part in wildlife conservation, acting both as a refuge and as a feeding site for birds and game.

Moreover, in an age concerned with the real problem of polluted air, the combination of nuclear and hydro generation results in reducing the amount of harmful pollutants that would otherwise result. This contribution to clean air represents another substantial social benefit.

Conclusion

The methods of achieving maximum efficiency in the generation of electricity vary with location and the relative costs of the factors of production.

The power companies of northeastern United States have determined that their stewardship of available natural resources is best served by combining nuclear generation with pumped-storage hydro—a most economically efficient and socially desirable mix.

This article is based on a paper presented at the 1968 World Power Conference, held in Moscow, U.S.S.R., August 20–24. The original paper will appear in the proceedings of the conference, copyright Soviet National Committee.

REFERENCE

1. Amberg, E. J., "Rocky River hydroelectric development," *AIEE Trans.*, vol. 47, pp. 1100–1109, Oct. 1928.

Sherman R. Knapp is chairman, chief executive, and a trustee of Northeast Utilities, a registered holding company whose principal subsidiaries are The Connecticut Light and Power Company (CL&P), The Hartford Electric Light Company, Western Massachusetts Electric Company, and Holyoke Water Power Company. He is also chairman and a director of Northeast Utilities Service Company. He received the bachelor's degree in electrical engineering from Cornell University in 1928. He then joined CL&P, where he served as an engineer in the Operating and Sales Departments until 1937, when he was made manager of the company's New Milford District. In 1941 he became assistant to the sales vice president and in 1948 assistant to the president. The following year he was elected executive vice president. In 1952 he was elected president and a director, in which capacity he served until 1964 when he was elected chairman, a position he still holds.

Mr. Knapp is president and a director of Connecticut Yankee Atomic Power Company; vice president and a director of Yankee Atomic Electric Company and Maine Yankee Atomic Power Company; and a director of several other firms related to the electric power industry in New England. Actively engaged in power industry affairs, he has served as president of the Edison Electric Institute and the Electric Council of New England. He is a vice president and a director of Atomic Industrial Forum, Inc., and a member of the National Industrial Conference Board. In 1956 he received an honorary D.Sc. degree from Hillyer College.

The crystal resonator—a digital transducer

An examination of two types of sensors reveals how the quartz crystal can provide quick and highly sensitive digital information concerning the physical parameters of interest—almost independent of other environmental factors

Donald L. Hammond, Albert Benjaminson

Hewlett-Packard Laboratories, Palo Alto, Calif.

Precision piezoelectric quartz resonators have evolved during a half century of research, development, and engineering. It is an evolution that started with rudimentary concepts and a precision of one part in a thousand, and emerged today with a precision of a few parts in 10^{12}. Although there have been significant steps forward through single contributions, the evolution has been a continuous improvement with an order of magnitude achieved every five to seven years. Not that anyone was particularly aware of these milestones, but, viewed in total, this improvement of a billion to one must stand as one of the truly amazing developments of this electronic age.

In the field of crystal resonators, the major areas of contribution have created a better understanding of the modes of piezoelectric resonator vibrations, better materials, vastly improved high-vacuum techniques, superior electroding methods, and major advances in mounting techniques. With these achievements, it seems appropriate to ask the question, "How can we convert this extreme precision of frequency into sensors for physical phenomena and still preserve at least part of the precision of the frequency stability of the quartz resonator?" As an answer, two sensors will be described here that at least partially effect a compromise—one for measuring temperature, and the other for measuring pressure.

The quartz thermometer

The concept of the quartz thermometer is not totally new. The first quartz-crystal units fabricated were made with a very high temperature coefficient and thus were good thermometers. It was not until 1933, however, that a rotated y-cut with a zero temperature coefficient was invented. Crystal units with high temperature coefficients were extensively utilized to measure the stability of crystal ovens.

Through the work of Wade and Slutsky,[1] Gorini and Sartori,[2] Flynn *et al.*,[3] and Smith and Spencer,[4] the quartz thermometer had evolved as a device with the inherent advantages of high sensitivity, low power dissipation within the sensor, and nondirect digital readout of temperature. The particular efforts of the authors were directed at augmenting these advantages with further improvements directed toward a linear frequency–temperature relationship, small size, short response time, and improved stability with respect to time and temperature cycling.

Over the temperature range from −50°C to +250°C, frequency–temperature behavior of a quartz-crystal unit can be well represented by the third-order polynomial, $F = F_0(1 + AT + BT^2 + CT^3 + \ldots)$. For thermometer applications, the ideal quartz resonator would have a linear dependence of frequency upon temperature. Thus the coefficients B and C should be zero in the ideal case. Through extensive work by Bechmann and others, the coefficients in this third-order expansion of frequency with respect to temperature have been related to the dependence of dimension, density, and the fundamental elastic constants upon temperature. This task, once completed, made possible the analytical calculation of the first three temperature coefficients of frequency for the three thickness modes in a plate having any generalized crystallographic orientation.

Using this approach, and rather generous application of a computer, we explored all possible orientations in crystalline quartz and found that the extensional mode had no regions where a linear dependence of frequency upon temperature existed. For all orientations, the second- and third-order frequency dependence of the extensional mode are negative.

In a similar manner, the two shear modes were explored. The higher-frequency shear mode, commonly referred to as the B mode, had both positive and negative values of the second- and third-order temperature

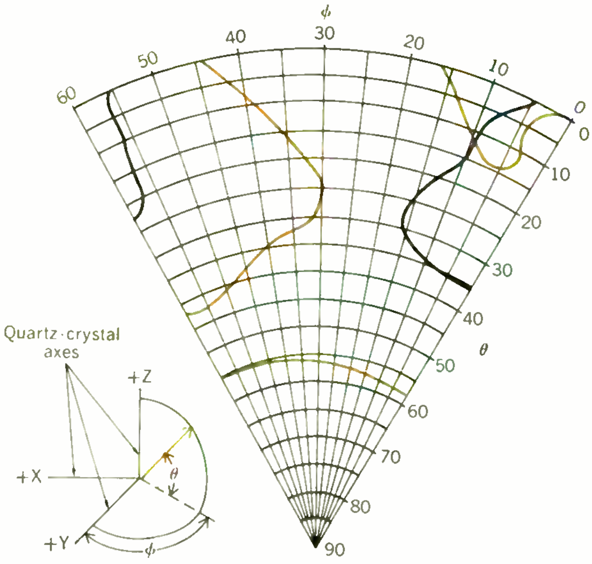


FIGURE 1. Loci of zero second-order (black curves) and third-order (colored curves) temperature coefficients $T_c^{\prime 2}$ and $T_c^{\prime}$ for the C mode.

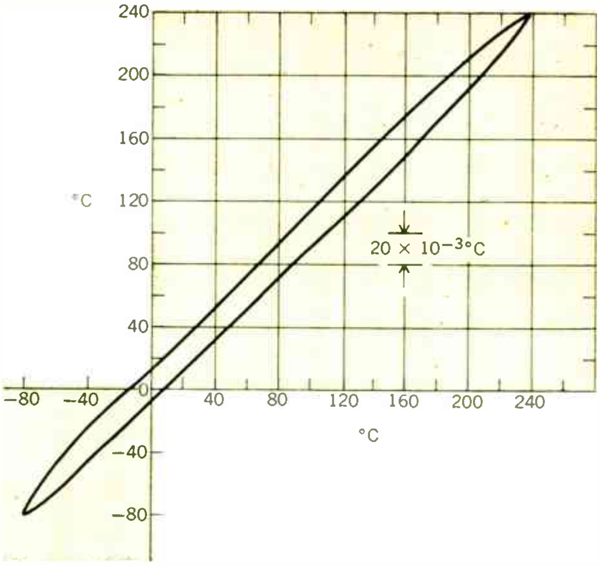
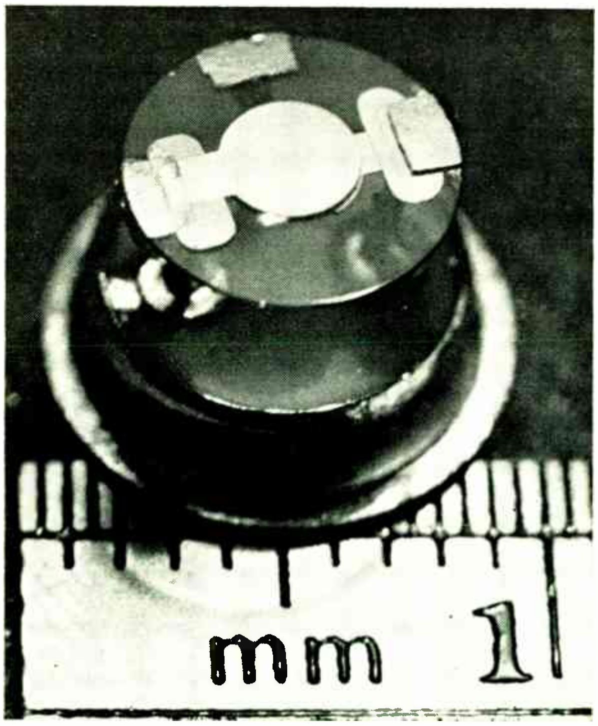


FIGURE 3. Hysteresis effect for a complete cycle from −80 C to +240 C.

FIGURE 2. Photo of LC resonator.



coefficients. However, zero values of these coefficients do not coincide at any orientation. The *C* mode, or lower-frequency shear mode, proved to be the most interesting. Figure 1 depicts the region in crystallographic space for θ between zero and 90 degrees and φ between zero and 60 degrees. This primitive region is repeated by symmetry to provide a description of any possible orientation in crystalline quartz. The black curves represent the locus of zero values for the second-order temperature coefficient of the *C* mode. Similarly, the colored curves represent the contours of the zero third-order temperature coefficient for the *C* mode. At an orientation of $\phi = 8.44$ degrees and $\theta = 13.0$ degrees, these loci cross. At that orientation, the first two nonlinear terms vanish. Experimentally, the actual orientation for linear dependence of frequency upon temperature is $\phi = 11.116$ degrees and $\theta = 9.393$ degrees, which differs from the analytically predicted value owing to small errors in the published values of the second-order and third-order temperature coefficients of the elastic constants. This linearity is accompanied by a first-order temperature coefficient that is about one third the maximum obtainable in quartz.

It is very fortunate that, at the single orientation in quartz exhibiting a linear frequency–temperature behavior, the desired mode (*C* mode) has the strongest piezoelectric excitation of the three thickness modes for an applied field in the thickness direction. All three modes, the two shear and the extensional modes, have been observed experimentally and the analytically predicted frequencies, mode strengths, and temperature coefficients have been verified. Agreement with the computed frequencies is better than one percent. This linear mode has been rather arbitrarily designated as the LC cut to indicate a linear coefficient.[5]

Figure 2 shows the LC-cut resonator mounted on a ceramic header using high-temperature brazed mounting structures and gold electrodes to support and excite the resonator. To avoid excessive aging, provide reliable operation at 250°C, and assure performance under high shock and vibration conditions, the crystal resonator is brazed to its three-ribbon mount. Shock studies of these units indicate retention of calibration at shock levels of 10 000 *g* and vibration levels of 1000 *g* over the frequency range from 10 to 9000 Hz.

A copper cap is cold-welded to the header flange to seal the sensor from contamination. The spacing between the crystal and the cap is held to three thousandths of an inch and the case is backfilled with one atmosphere of helium, providing a thermal time constant for the basic resonator and its case of three tenths of one second and less than two seconds for the complete thermometer probe, which includes the stainless-steel housing. Backfilling with helium also makes possible the use of mass spectrographic leak detection techniques to assure the integrity of the hermetic seal.

In order to normalize the first-order temperature coefficient of 35 parts per million per °C, a frequency of operation on the third overtone of 28.208 MHz at 0°C was chosen. This resulted in an increase of approximately 1000 Hz per °C, with a sensitivity of a millidegree per one-second sampling time. Using longer sampling periods or frequency multiplication results in greater sensitivity limited by the observed short-term stability of 5×10^{-6} °C (rms).

The frequency of 28 MHz is high enough to enable a 0.250-inch (0.64-cm)-diameter resonator to be used without introducing perturbations in frequency caused by the mounting structure or anomalies in the frequency–temperature curve introduced by the existence of activity dips. A plano-convex, third-overtone resonator was utilized to reduce the mounting effects further and to decouple interfering modes of motion from the desired thickness shear mode. With these design parameters, it has been possible to eliminate completely any perturbations of frequency over the temperature range from −80°C to +250°C.

One of the major engineering problems encountered in the development of the quartz thermometer was associated with the phenomenon of hysteresis (Fig. 3). If the ice point is approached from +250°C, the indicated temperature will be 20 millidegrees or approximately one part per million in frequency higher than if the ice point is approached from −80°C. This very elusive effect has been shown to be independent of plating material or type of crystalline quartz used (such as natural, synthetic, doped, or swept), and unrelated to the mounting structure. Recent studies have shown it to be related to the differential expansion between the quartz plate and the thin film, as well as micro contamination within the crystal holder. With the proper processing techniques, this effect can be reduced to a few millidegrees for temperature excursions from −80°C to +240°C. For smaller variations in temperature, the effect is negligible.

The long-term stability of the quartz thermometer after the initial aging period indicates a divergence of a few thousandths of a degree Celsius per month. Using an ice bath, periodic calibration is, of course, straightforward.

Figure 4 illustrates the encapsulated quartz-crystal sensors sealed in probes at the end of 12-foot (3.7-meter) half-wavelength cables that provide thermal isolation between the sensor and the transistorized instrument. The crystal oscillator can be removed from the rest of the instrument by a cable as long as 5000 feet (1500 meters). Figure 5 shows an oceanographic probe designed to withstand 20 000 lbf/in^2. It contains the sensor and oscillator, the reference and its oscillator, as well as a mixer.

Figure 6 displays the block diagram of the quartz thermometer. The complete HP2801A quartz thermometer contains a reference oscillator with a frequency equal to the sensor frequency at zero degrees Celsius, a mixer to provide the difference frequency, a sensing circuit to determine whether the temperature is plus or minus, a divider chain for the counter circuit, control circuits to select the absolute or differential modes of operation for the two independent crystal sensors, control circuits for the sampling time, and a counter for direct digital display of temperature in either degrees Celsius or degrees Fahrenheit. Thus, the quartz resonator provides a thermal sensor of high sensitivity, low power dissipation, direct digital readout of temperature with linear dependence of frequency upon temperature, small size, fast response time, and stability with respect to time and temperature cycling.

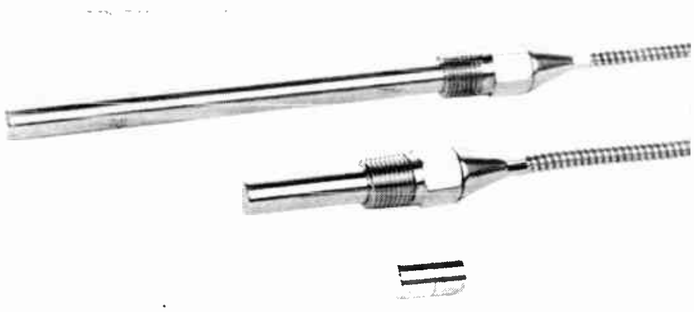


FIGURE 4. Thermometer probes.

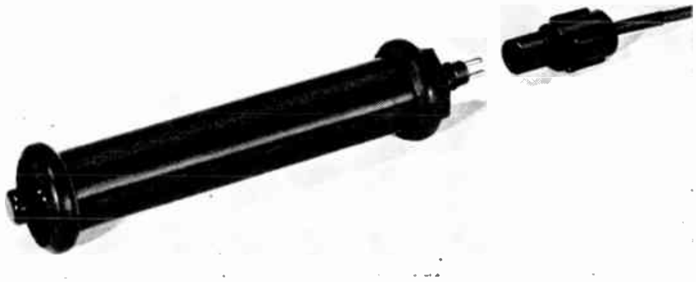


FIGURE 5. Oceanographic sensor.

The quartz pressure transducer

A quartz hydrostatic pressure sensor utilizing a separate phenomenon will now be described. If a quartz resonator is operating in a thickness mode, frequency will be determined by the thickness, density, and an appropriate elastic constant. If a force were to be applied along a diameter of the resonator, a frequency change would be expected due to the change of thickness proportional to Poisson's ratio and the increase in density as shown by $f_r = (1/2t)\sqrt{C/\rho}$.

Experimentally, however, the effect is much larger than predicted on the basis of density and thickness alone. From the observed anisotropy, the large effect appears to be related to nonlinearities in the elastic properties of the material. These observations were first reported 16 years ago by Virgil Bottom *et al.*,[6] and have been important in helping to understand certain mounting and acceleration problems associated with precision resonators. Experimentally, the phenomenon has been investigated by Mingins, Barcus, and Perry.[7]

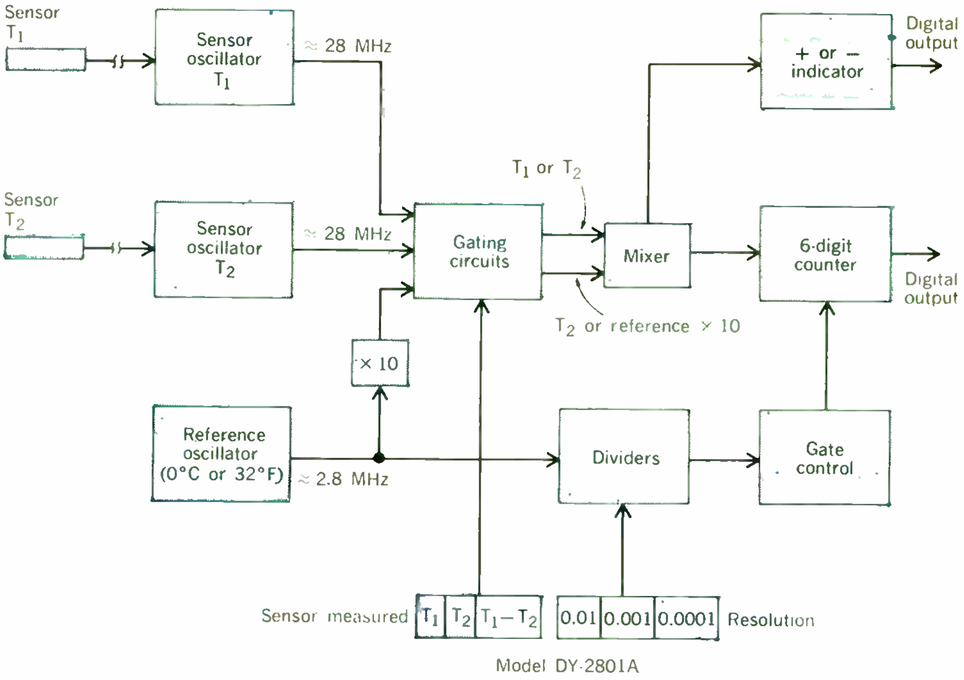


FIGURE 6. Block diagram of the quartz thermometer.

Some attempts have been made to utilize this phenomenon for compensating crystals over wide temperature ranges; other proposals have been made for using them as force sensors. Thurston, McSkimin, and Andreatch[8] have measured ultrasonic plane-wave phase velocities as a function of uniaxial stress and hydrostatic pressure for many orientations in quartz, and from these data have deduced the third-order elastic constants. Hence, it is possible to predict the frequency perturbations caused by force along any diameter of a resonator in any generalized orientation.

In the original work reported by Bottom *et al.*, a linearity of better than one percent was observed. An earlier program at Hewlett-Packard, carried out by Fritz Baur, employed an *AT* quartz resonator with a zero temperature coefficient as the force sensing element in a weighing device. The force was applied in tension along the x-axis. Baur observed high sensitivity and high linearity, however, and repeatability was consistently limited to one tenth of one percent.

After considerable study, it became apparent that this lack of repeatability was related to instabilities in the bond between the quartz and the mounting structure. If the quartz were simply placed in contact with high-strength materials, micro fractures occurred at the high-stress points of contact, resulting in poor repeatability of frequency shifts with force. On the other hand, if soft materials were used in contacting the quartz resonator, plastic deformation caused significant modification of the contact with the resonator, resulting in frequency instabilities. If bonding techniques such as solders or epoxies were used, they exhibited sufficient plastic flow to cause appreciable zero-frequency offset after the application of full load.

The next approach to the problem came after a period of reflection. It was reasoned that, if the joint between quartz and the rest of the mechanism introduced undesirable mechanical problems, the entire mechanism should be made out of crystalline quartz. This resulted in the design shown in the cross section of Fig. 7. Here the quartz resonator is a thick plate across a quartz cylindrical diaphragm. This entire structure is fabricated from a single piece of crystalline quartz. The design is a 5-MHz fifth-overtone doubly convex resonator originally introduced by Arthur Warner of Bell Telephone Laboratories.[9] It is the basic resonator for all present precision frequency standards and exhibits near-optimum performance in short-term and long-term frequency stability. Its choice was based on the assumption that if one sets out to pervert something, it should be the very best.

Because the compressive strength of quartz is more than an order of magnitude greater than the tensile strength of quartz, this design substantially increases the dynamic range. For the 5-MHz fifth-overtone resonator and a wall thickness comparable in thickness to the resonator, failure should begin at a pressure level of 10 000 lbf/in². In this configuration, a frequency change of 2.5 Hz per lbf/in² is introduced by pressure. The frequency change at 10 000 lbf/in² is 0.05 percent of the resonator frequency, which has an inherent frequency stability of 10^{-10} or better, indicating a dynamic range of 5×10^6. Actually, a sensitivity of 10^{-3} lbf/in² out of a full-scale range of 10^4 lbf/in² has been observed. Linearity of 0.08 percent has been observed in several experiments. These measurements are continuing and improvments in linearity are anticipated.

Even with the precaution of fabricating the entire device out of quartz, there were still significant problems in repeatability. After application of several thousand lbf/in², the shift in the zero-pressure frequency corresponds to a pressure change of as much as a tenth

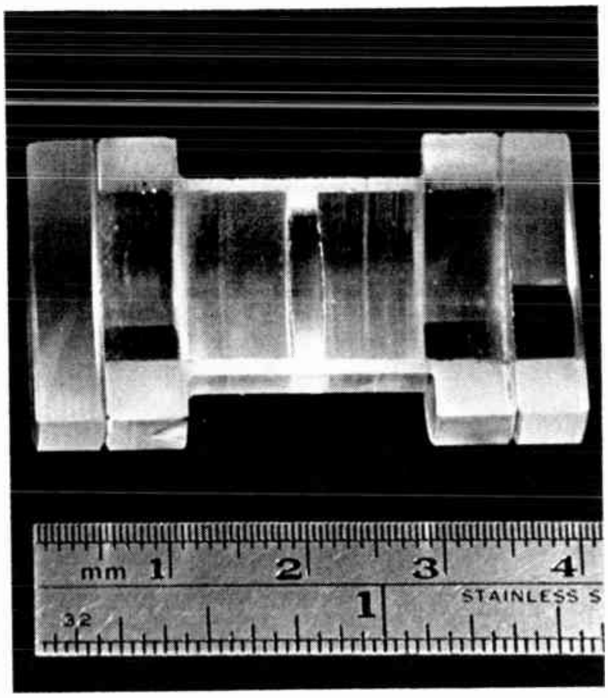


FIGURE 7. quartz pressure transducer cross section.

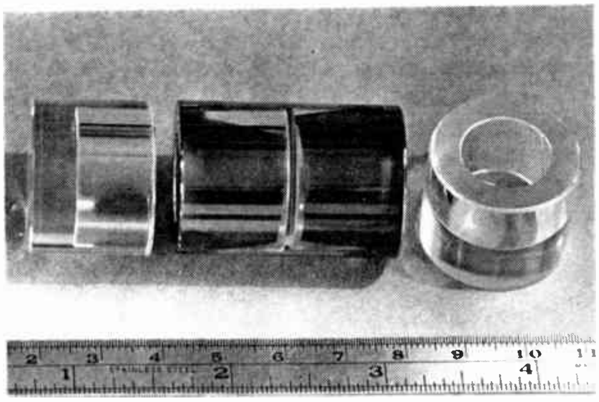


FIGURE 8. Quartz pressure transducer prior to sealing.

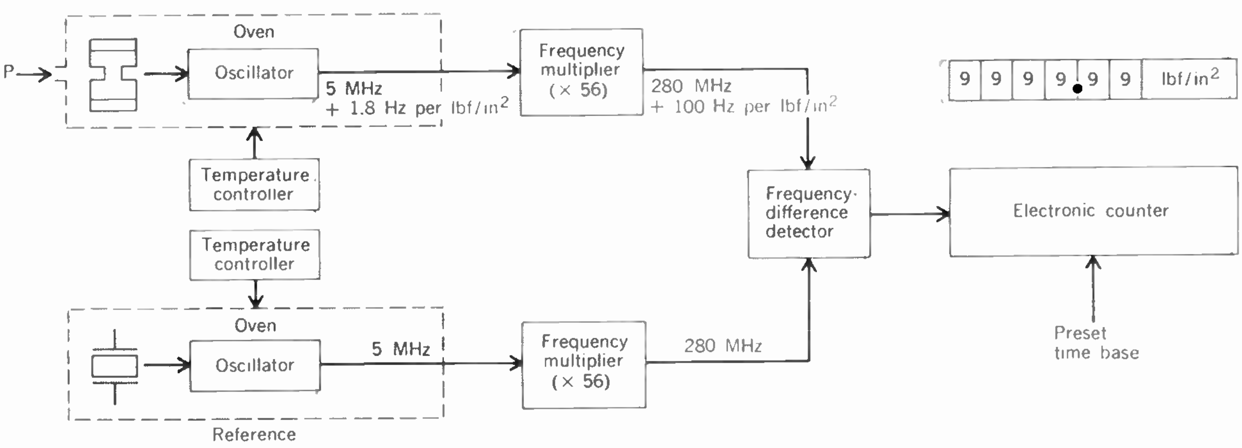
of a lbf/in^2. These effects again were traced to defects in the bond joining the quartz end caps to the quartz cylinder. Extensive analysis and design modification resulted in greater acoustic attenuation at the bond, reduction of stress across the joint, and the elimination of tensile stresses within the transducer.[10] The final version is displayed in Fig. 8, prior to sealing the end caps to the cylinder. With these improvements, repeatability of better than 0.01 lbf/in^2 was obtained, and a sensitivity to pressure changes of at least 0.001 lbf/in^2 has been observed.

The design of quartz resonators with zero temperature coefficients was completed in 1933, resulting in the *AT* and *BT* thickness shear resonators. Fabricating the quartz pressure transducer from an *AT* orientation results in a zero temperature coefficient of frequency at 0 lbf/in^2. However, we found that the scale factor of 2.5 Hz/lbf/in^2 changed by 800 parts per million for each °C. Thus, to obtain a pressure reading with an accuracy of one part per million requires temperature control of the resonator to 1.2×10^{-3} °C. Hence, the dependence of frequency upon temperature at zero lbf/in^2 can be made zero by using the *AT* resonator. However, the dependence of frequency upon temperature at 10 000 lbf/in^2 has a negative temperature coefficient of 8 lbf/in^2 per °C.

Since there are two degrees of orientational freedom in crystalline quartz, it should be possible to use one degree of freedom to obtain a zero temperature coefficient at 0 lbf/in^2, and the other degree of freedom to obtain a zero temperature coefficient of the scale factor. We were fortunate in finding an orientation that nearly satisfied these conditions. The temperature coefficient at 0 lbf/in^2 has been made zero and the temperature coefficient of the scale factor has been reduced to 20 parts per million per °C, reducing by a factor of 40 the requirements for temperature control.

A simplified block diagram of the quartz pressure-measuring system is shown in Fig. 9. At this new orientation, the frequency sensitivity to pressure is slightly less at 1.8 Hz/lbf/in^2. Multiplying this frequency by 56 and subtracting the frequency of an identical oscillator, also multiplied by 56, results in a frequency that can be made to directly read in lbf/in^2

FIGURE 9. Block diagram of quartz resonator pressure transducer system.



and, for one-second sample times, provide a sensitivity of 0.01 lbf/in^2.

The transient response characteristics of the pressure transducer are limited in part by the thermal transients introduced by the adiabatic heating of the fluid medium. For example, if the pressure is rapidly increased from 0 to 10 000 lbf/in^2, the fluid, such as silicon oil, will be adiabatically heated, causing the outer portion of the quartz crystal to rise in temperature above the central portion of the resonator. This will cause a significant frequency decrease even though the resonator is designed to have a zero temperature coefficient of frequency. This effect can be reduced by decreasing the amount of fluid in the cavity and by using a fluid with a small bulk compressibility.

One of the most interesting applications of the quartz pressure transducer is in the field of oceanography. The full-scale pressure range of the quartz transducer is compatible with submersion to the bottom of the ocean with the possible exception of the deepest trenches. The thermal environment at the ocean bottom is nearly ideal. The temperature at the bottom in nearly all parts of the world is zero degrees plus or minus two degrees Celsius, and the stability with time is a few hundredths of a degree. In this environment, the quartz pressure transducer provides a measurement of pressure to 0.01 lbf/in^2 or better under pressures of 10^4 lbf/in^2. This type of transducer is ideal for tidal studies and provides an ideal sensor for an early warning system to detect tsunami waves. Another interesting application is in oil-well logging, where changes in pressure with flow rate and time are to be monitored. Although the temperature is higher than that in the ocean, it is equally stable.

The quartz pressure transducer provides a rugged sensor with the advantages of wide dynamic range, high linearity, and direct digital presentation combined with small temperature sensitivity, permanent calibration, high portability, and high sensitivity.

Conclusion

These two sensors for temperature and pressure utilize the extreme precision of the quartz crystal and make possible direct digital presentation of these physical parameters with extremely small dependence upon other environmental parameters. They are not, however, digital devices in the purest sense. Strictly speaking, they provide for the conversion of information from the analog physical parameters of pressure and temperature into the analog parameter of frequency, which in turn can be measured with great precision using the highly developed techniques of digital counters. Needless to say, exploration of other quartz-crystal applications in the measurement of physical parameters is continually being pursued.

The authors wish to acknowledge the participation of Charles Adams, Jerry Leach, and Edward Karrer in these investigations, and the support of Henry Yoshida, Darlene Watson, and George Nelson. Revised text of a paper presented at the 1968 Solid State Sensors Symposium, Minneapolis, Minn., September 12–13.

REFERENCES

1. Wade, W. H., and Slutsky, L. J., *Rev. Sci. Instr.*, vol. 33, p. 212, 1962.
2. Gorini, I., and Sartori, S., "Quartz thermometer," *Rev. Sci. Instr.*, vol. 33, pp. 883–884, Aug. 1962.
3. Flynn, T. M., Hinnah, H., and Newell, D. E., "An improved cryogenic thermometer," *Proc. Cryogenic Eng. Conf.*, vol. 8, pp. 334–339, Aug. 1962.
4. Smith, W. L., and Spencer, W. J., "Quartz crystal thermometer for measuring temperature deviations in the 10^{-3} to 10^{-6} °C range," *Rev. Sci. Instr.*, vol. 34, pp. 268–270, Mar. 1963.
5. Hammond, D. L., Adams, C. A., and Schmidt, P., "A linear quartz-crystal, temperature-sensing element," *Trans. Instr. Soc. Am*, vol. 4, pp. 349–354, Oct. 1965.
6. Bottom, V. E., *et. al.*, "Mounting techniques for improved heat dissipation in quartz crystal units," U.S. Army Signal Corps Final Rpt., Cont. No. DA36-039-SC-5485, Dept. of Physics, Colorado A&M College, Feb. 28, 1951, to Feb. 29, 1952.
7. Mingins, C. R., Barcus, L. C., and Perry, R. W., "Effects of external forces on the frequency of vibrating crystal plates," *Proc. 16th Ann. Freq. Cont. Symp.*, pp. 47–76, Apr. 1962; also *Proc. 17th Ann. Freq. Cont. Symp.*, pp. 51–87, May 1963.
8. Thurston, R. N., McSkimin, H. J., and Andreatch, P., Jr., "Third-order elastic coefficients of quartz," *J. Appl. Phys.*, vol. 37, pp. 267–275, Jan. 1966.
9. Warner, A. W., "High-frequency crystal units for primary frequency standards," *Proc. IRE*, vol. 40, pp. 1030–1033, Sept. 1952.
10. Karrer, H. E., and Leach, J., "A quartz resonator pressure transducer," presented at the Feb. 1969 Transducer Conf.; to be publishe in *IEEE Trans. Industrial Electronics and Control Instrumentation.*

Donald L. Hammond (SM) holds bachelor's and master's degrees in physics from Colorado State University at Fort Collins, and has pursued graduate studies in elasticity at Columbia University, New York City. Joining the Hewlett-Packard Company in 1959 as manager of the Quartz Crystal Department, he established a facility for research, development, and production of precision quartz-crystal resonators. Promoted to general manager of physics research and development in 1964, he directed activities in the areas of quantum electronics, electroacoustics, and high-vacuum devices. When Hewlett-Packard Laboratories was established in 1966, he was made director of its Physical Electronics Laboratory, which is concerned with the research and development of lasers, optics, analytical instrumentation, piezoelectricity, thin-film devices, and cathode-ray tubes. A member of the American Physical Society and Sigma Xi, Mr. Hammond is the author of numerous technical articles and papers.

Albert Benjaminson received the B.S.E.E. degree from the University of Adelaide, South Australia, in 1950; prior to this, he pursued a course of study at the Polytechnic Institute of Brooklyn in New York. His early engineering experience was obtained at the Pilot Radio Corporation and Granco Products in New York City, as well as the Sperry Gyroscope Company, where he participated in the development of radar systems, Loran-C, and tactical ground support equipment for military aircraft. A development engineer with Hewlett-Packard's Dymec Division since November 1959, he became manager for RF systems and transducer development in December 1967, responsible for the company's line of phase-locking synchronizers, which are used for VLF phase comparator receivers and crystal control of microwave oscillators. He has since developed the quartz thermometer and is currently employed as a research engineer in the Hewlett-Packard Physical Electronics Laboratory, where he is working on a pressure transducer system. He has authored many technical articles.

"VORTEX," 1966, by Harold Tovish. (Collection The Whitney Museum of American Art, New York)

Art and technology

I. Steps toward a new synergism

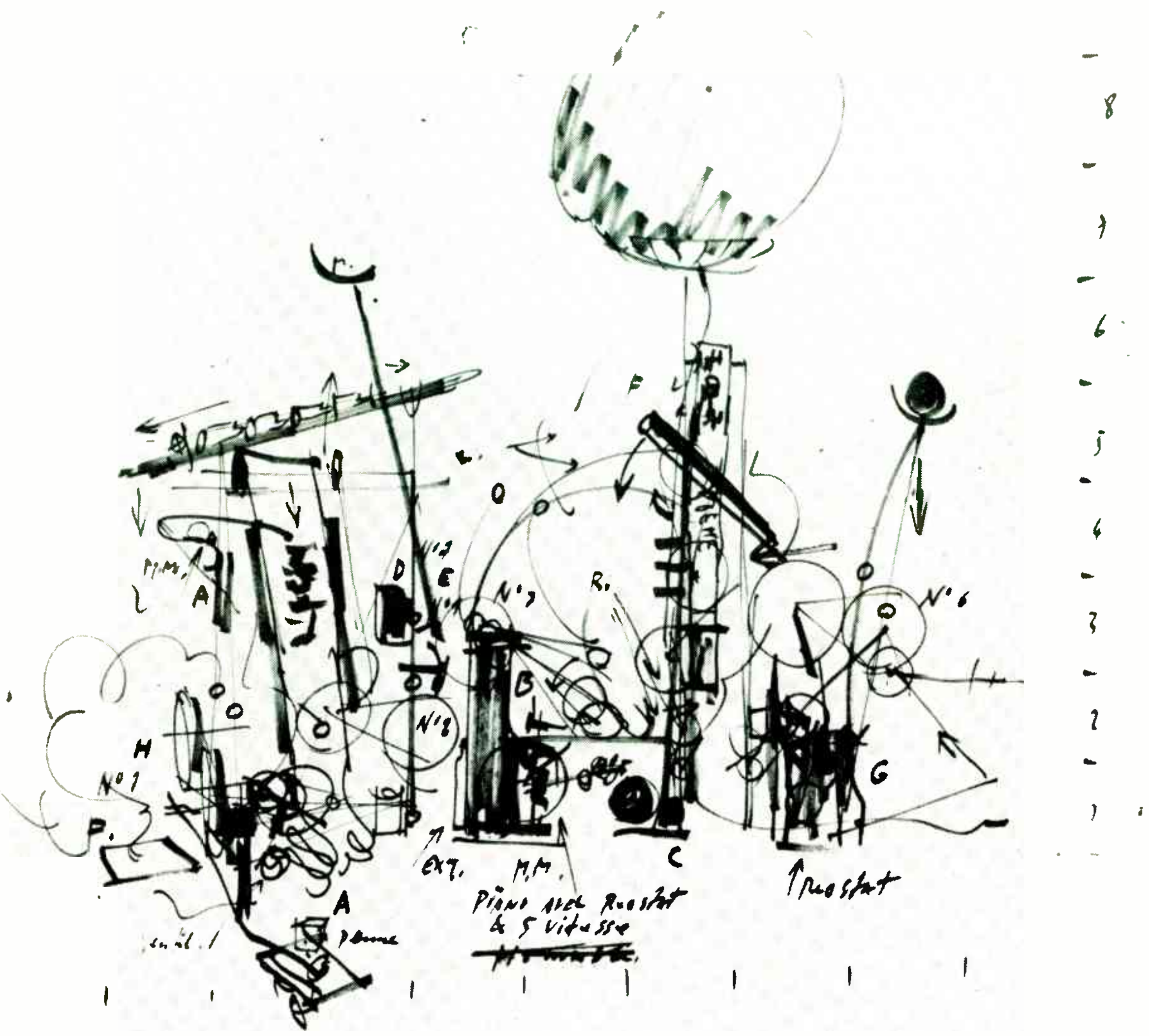
In the heyday of the Dada movement, the artist Francis Picabia said, "We wanted to make something new, something that nobody had ever seen before!" His is the voice of this century—of both art and technology. Now the emergence of collaborations between artists and engineers compounds the possibilities for even more radical changes in our environment through the "new combine," art and technology

Nilo Lindgren *Staff Writer*

Many artists today, especially those in the younger generation, are looking to modern technology as a new medium in which to delve, and they are looking, as well, for engineers to assist them and to collaborate with them on their works. This two-part article describes certain aspects of the art-and-technology movement, and speculates on some of its possible consequences. Part I traces the roots of current artistic interest in technology and describes modes of collaboration. Part II will describe some of the experiences of those who have already engaged in collaborative projects, discuss organizations that have been formed to facilitate the collaborative process, and ask you to consider the possibility of jumping in yourself or lending your organization's support.

On March 17, 1960, the fantastic machine shown on this month's cover was put on display in the garden of the Museum of Modern Art in New York. The display, shown to a chic audience of specially invited guests, lasted for half an hour, and then it was gone. The machine sculpture, called "Homage to New York," had destroyed itself.

During its short life, this artful contraption, programmed for a symphony of suicide, sent out smoke flashes, rang bells, played a piano with mechanical arms made of bicycle parts, poured gasoline on itself, set itself on fire, crushed bottles filled with evil-smelling gas, turned on a radio, spilled cans of paint on rolling scrolls, honked a loud klaxon, fired out a smaller suicide carriage at terrific speed, threw out silver dollars, melted its supports, sagged, and nearly collapsed. The



PRELIMINARY drawing for "Homage to New York," 1960, by Jean Tinguely.* Engineer Billy Klüver who worked with Tinguely, recalls that when the public arrived to see the machine destroy itself, their work was still in progress. "At 7:30 I was finished. 'On va?' 'On va,' said Jean. He looked as calm as if he were about to take a bus. Not once did we go over and check everything. The construction and the beginning of the destruction were indistinguishable." (Collection The Museum of Modern Art, New York, gift of Peter Selz)

machine was supposed to have crawled to the museum pond and thrown itself in, but it didn't. This self-destruction sculpture, *l'art éphémère*, the work of sculptor Jean Tinguely, was to create "a direct connection between the creative act of the artist and the receptive act of the audience." It was a direct and delightful assault on the traditional idea that all art must be "lasting."

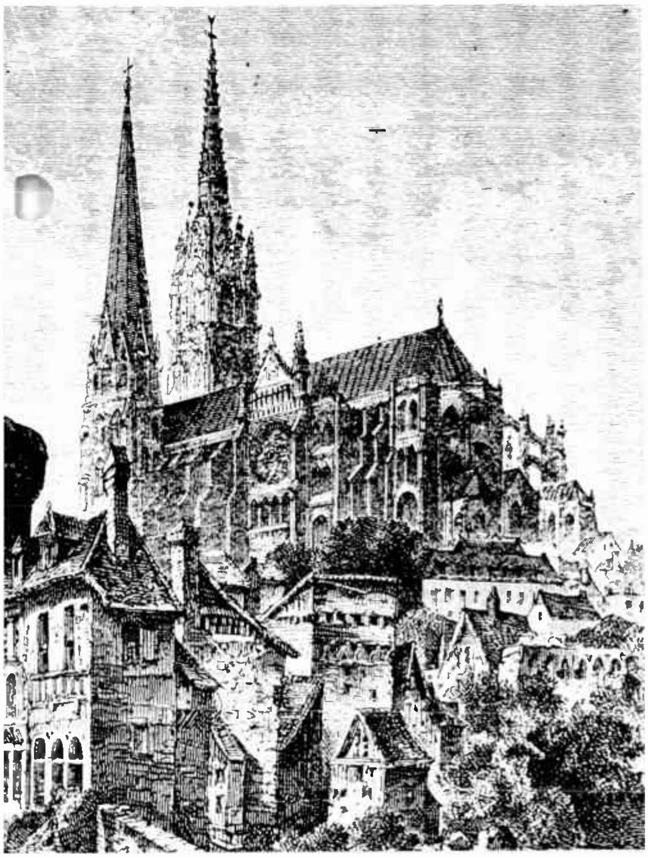
"Ephemeral art," wrote Billy Klüver, the Bell Labs engineer who helped construct "Homage to New York," "forces us into contact with the ever-changing reality." [1] For engineer Klüver and artist Tinguely, the self-destroying machine "was not a functional object and was never treated like one." Its parts—bicycle and baby-carriage wheels, motors, and so on—many of which had been dredged from the dumps of New Jersey, went back to the dump. The self-destruction of the machine, they said, was "the ideal of good machine behavior."

Was it nonsense? A put-on? A Dadaist *geste*? An "expression of nihilism and despair"? Was it art?

Another of Tinguely's "philosophical machines" appeared in New York at the 1967 Vision Conference sponsored by the International Center for the Communication Arts and Sciences. This machine did not destroy itself; instead, it destroyed beer bottles, full of beer, which were fed to it on a chain conveyor. Introduced by a trombone player and a singer, who was bewigged and dressed in the costume of a bygone era, and whose lyrics were puckishly woven round the theme "There are too many machines," this machine, rather like a giant black grasshopper, and equipped with whirling wheels of all sizes, drew the beer bottles up onto a

---

* In Exhibition: "The Machine—As seen at the End of the Mechanical Age," The Museum of Modern Art, New York, November 27, 1968–February 9, 1969.



CHARTRES, one of the most magnificent of the Gothic cathedrals, was created at a time when art and engineering still worked hand in hand.

tonguelike apparatus, where it smashed them with synchronized blows of an axlike implement. Enslaved to the machine was a small man in an oriental costume, wielding a broom and dustpan, who swept up the broken glass. People laughed uproariously at first, but then—as the machine devoured bottle after bottle, smashing them into smithereens, as the action stretched out in time—the tension mounted. Some grew irritated, others somber, and others sad. The initial delight—manifested by cheers when the machine made a direct strike, and groans when it missed—gave way to deeper feelings as the action was sustained.

This infernally humorous "garbage disposal" machine, fed bottles by people and cleaned up after by people, made one yearn for it to desist. At the same time, one yearned to have it go on. Is that a function of art? Was it art?

Just this past fall, there went on display at the Museum of Modern Art in New York an "environmental" work called "Soundings," created by the U.S. artist Robert Rauschenberg. The piece—"hung" or "erected," depending upon one's professional point of view—is a construction of nine 8-foot-high silvered panels forming, for the spectator who enters the gallery, a 36-by 8-foot mirror object, in which all he sees at first is his own reflection. Only when the spectator speaks does the piece come to life. It responds to the varying resonant frequencies of a speaker's voice by displaying (see page 62) chairs that are lit up here and there, at different places and at different levels from the foreground mirror. For different speakers, and for different combinations of words and sounds, different combinations of chairs appear and disappear. Observers say that delightful things have happened with "Soundings." One day, for instance, a man came into the gallery with a guitar, and he played and sang to the work while many other spectators gathered around him to watch and listen.

Working with Rauschenberg on "Soundings" was a team of engineers who designed and wired up the panel, which is actually three levels deep, for 200 different combinations of response.

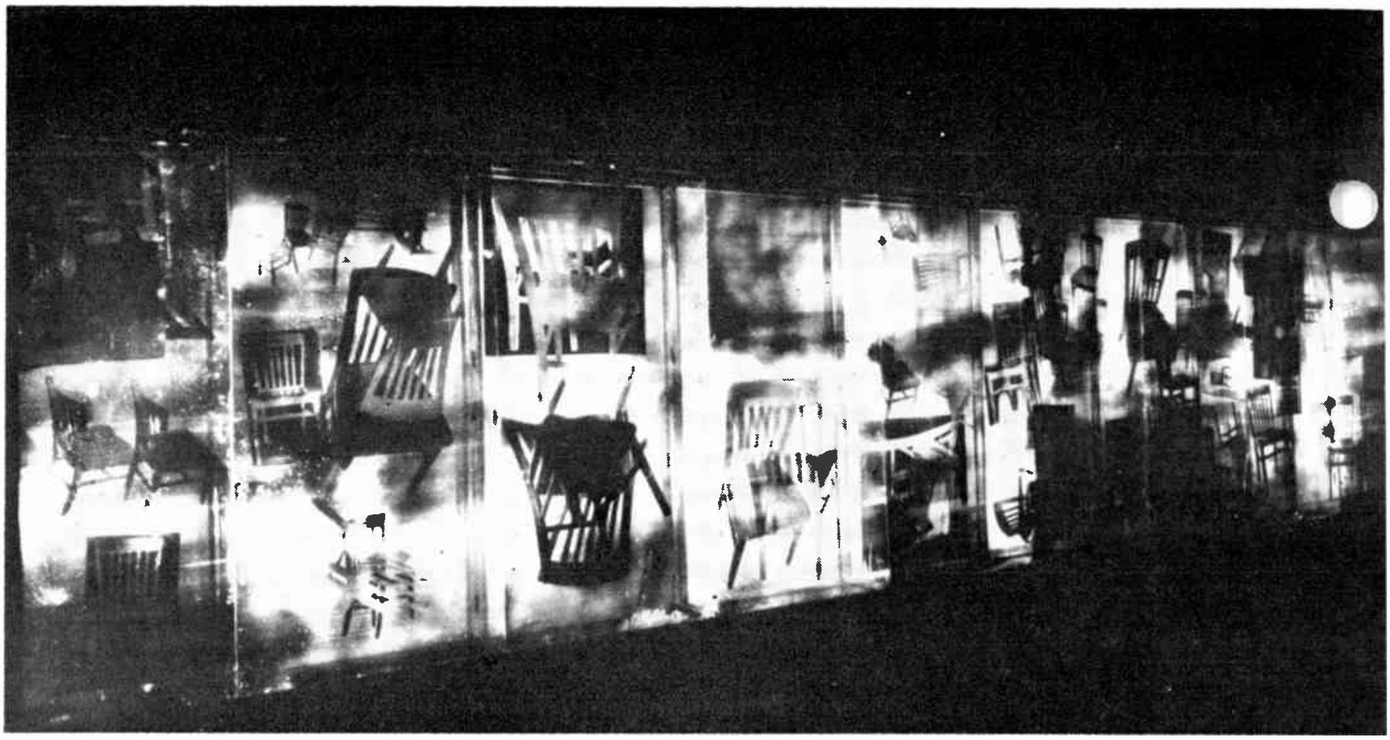
Clearly enough, these works depart from traditional sculpture and painting. People respond to them in different ways. Without defining how, we know that the intention of the artist was different in each case. Whether or not one calls these works art is not, however, the question we take up in this article.

We single out these pieces not to dramatize the radical changes that have taken place in the arts, but to point out what they have in common. Artists wanted to make these works and, to do them, they needed engineering assistance. These works all used elements of 20th century technology. More important, they represent collaborations between artists and engineers.

Nor are these isolated instances. Many artists today, especially in the younger generation, look to modern technology as a new medium in which to delve. However, most of their experiments thus far have proceeded with little or no contact with engineers. The artists have used, as it were, the spilloff of technological production, most predominantly in the United States, where the cataract of products and devices springing from technology has had the most profound and pervasive sociological effects. Nor should this be surprising. Where people's lives are most deeply touched is where the artist moves. (With respect to this point, McLuhan's view, quoted in the box on page 62, is worth contemplating.)

One trouble is, of course, that technology exercises a two-pronged effect. Its wonderful "benefits to man," in their unalloyed forms, have also left people acutely suffering from new devastations. The automobile gives people more mobility, but its exhaust poisons the air, and the superhighways destroy the countryside. So people now embrace technology with deeply ambivalent feelings. The role of the artist as a mediator between technology and people—Rauschenberg, for instance, speaks of making the technology less mysterious and "more friendly"—could be of great consequence, *provided* the artist gains entrée to the sources of technology, to the brains of the scientists and engineers who carry it. Otherwise, the artists (misjudging the products of technology for its essence) will work against the constant danger of achieving a few "effects" in a technological idiom, but at an incommensurate cost—that of trivializing the real grandeur of the technical resources they wish to reveal.

The artists themselves, acutely aware of this danger, need and are asking for the collaboration of engineers. At the 1968 IEEE International Convention, filmmaker Stan VanDerBeek, his arms outstretched, cried, "This incredible geography of ideas that is unfolding around us! Where do we (the artists) start with all this?"



"SOUNDINGS," 1968, an environmental work by Robert Rauschenberg, composed of silk-screen ink on Plexiglas with electronic equipment. A team of engineers, coordinated by Leonard R. Robinson, worked with the artist. The electronics was chiefly the work of Ralph Flynn with the assistance of Anthony Tedona and Logan Hargrove. The mechanical design was by Per Biorn and others. In terms of other technological art works, "Soundings" is an elegant and sophisticated piece of systems engineering. (Courtesy Leo Castelli Gallery, New York)

The artist and technology

The effects of technology do not occur at the level of opinion or concepts, but alter sense ratios or patterns of perception steadily and without any resistance. The serious artist is the only person able to encounter technology with impunity, just because he is an expert aware of the changes in sense perception.

Marshall McLuhan
From *Understanding Media: The Extensions of Man*

And it is clear that there are real barriers that must be overcome. Engineers and their industries have it in their power to reject or to respond to the call for collaboration. If they do not heed the artists' pleas, if they feel threatened by something that seems a little "too far out," if they are contemptuous of something that seems merely gimmicky, if they rely *only* on the narrowest economic judgments and interests, well . . .

But if engineers do respond seriously, and begin to work in deep-going collaborations, there is no telling what interesting issue will come from such unions. The power of the artist to stun, surprise, shock, move, when coupled with the real power of technology, could be transfiguring. As one art historian recently quipped, "After Picasso, women and owls never looked the same." After an art-and-technology movement, our whole environment might look radically and revitalizingly different.

From past to present

Before going further, we should consider for a moment a question of interpretation. In this article, we are dealing with the art-and-technology movement as a new phenomenon that contains within it the seeds of radical possibilities. But there are those observers, mindful of the millenia during which the arts and crafts were not separated, who see this reaching out of artists and technologists as simply the signal that certain streams of human endeavor, which somewhat accidentally and temporarily diverged a bare 200 years ago, are converging once again. As they point out, the works of antiquity, of the Middle Ages, and even of the Renaissance, were expressions of a creative spirit that did not yet distinguish, as we have learned so well to do, between the fine and the mechanical arts. (See, for instance, Singer and Holmyard's *History of Technology*.[2]) From this point of view, the tentative but conscious cooperation of the artist and technologist may seem to be a deliberate and contrived act only because it comes at a time when the functions of these two artisans are furthest removed from one another. We do not really quarrel with the scholarly interpretation.

Without denying history, however, we may lay our stress on a more journalistic interpretation. We may, with Marshall McLuhan and others, see the mechanical age as ended, and we may see ourselves as having passed the threshold into an electric technology that throws open totally new possibilities.[1-3] The journalist looks at the artist-and-technologist synergetic movement as a unique and potent event, and although he exaggerates the magnitude of certain of its aspects, his efforts may add, however slightly, to its impetus. If the confluence of art and technology in the '60s and '70s proves to be vigorous and thoroughgoing, both academics and journalists may take comfort in the fact that a 200-year-old breach has begun to heal at last.

With this caveat in mind, let us leap across the centuries of art history to trace a few of the roots of the current artistic interest in technology—to record, as it were, the immediate legitimacy of the art-and-technology movement. For the engineer who feels attracted to the idea of collaborating with an artist, but who also feels intimidated because he has no idea of what has been going on in art, a few points may prove unburdening, and perhaps even decisive.

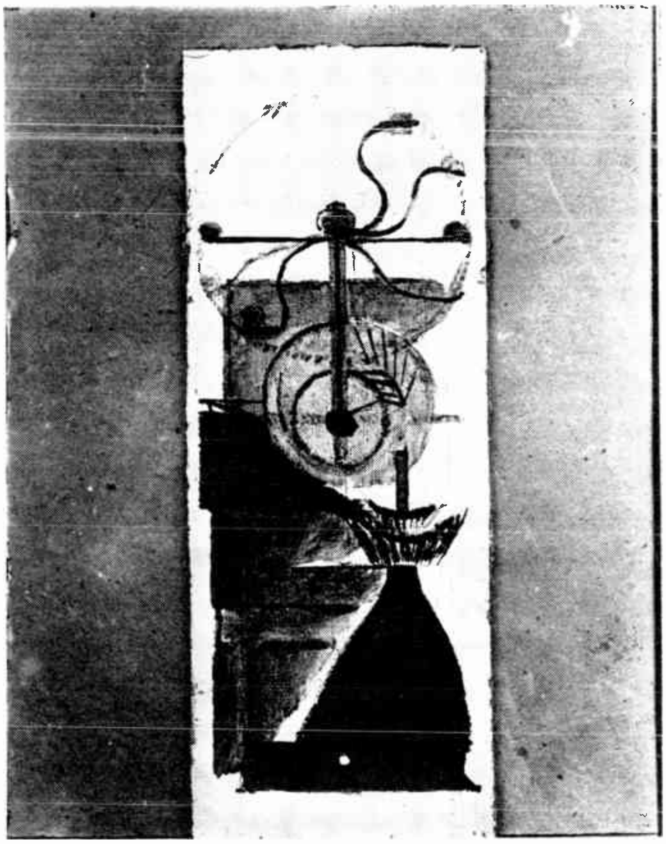
Toward the radically new

Even within the context of the explosive art of this century, the recent shift in artistic aims toward works sometimes labeled "anti-art," represents a radical departure.[4-6] This art differs in its manifestos and in its philosophy of esthetics as much as it does in the visible objects it creates. It abandons illusionism in painting, those make-believe windows looking out on the world. It creates works that attempt to be interactive with their observers, and that incorporate elements of variability and randomness. The artists who make these works say that they are attempting to elicit new responses of the viewers, either by thrusting them into states of sustained meditation, and even boredom, or by forcing the viewers out of a passive state by contriving to make them act or participate in order for anything to happen.

Conscious of the gap that has steadily widened between current art and current technology, especially in our time, the new art has deliberately begun to explore the processes, materials, and systems coming from modern technology. The new art is exploring the *dynamics* of a person's engagement with a work rather than depending upon his passive contemplation. This shift of focus from the statics to the dynamics of perception and feeling, perhaps one of the most characteristic features of recent 20th century art in contrast to 19th century art, has a parallel in many fields of science and engineering. The interest of engineering and science in dynamics, however, stems from the need to deal with ever more complex and interlocking systems (all the way up to the level of complexity of the living organism); artistic interest in dynamic interactions draws its strength from many different fibers of 20th century art.

Readymades—the first step

One of the significant figures in the art of this century, and the man whom many credit with the first deeply influential steps, is Marcel Duchamp, most popularly known for his painting "Nude Descending a Staircase." He departed from the Cubist context of Parisian painting in 1912, which itself was on the verge of total abstraction—he departed from brushes, canvas and paints, to create "an anti-art of Readymade objects and images on glass." [4] Centuries of illusionistic iconography were beginning to give way to a new impulse.



"COFFEE GRINDER," painted by Marcel Duchamp in 1911,* heralded his subsequent fascination with real machines and real objects. He wrote: "But instead of making a figurative coffee grinder, I used the mechanism as a description of what happens. You see the handle turning, the coffee after it is ground—all the possibilities of that machine." He regarded it as the key picture to his complete work. (Collection Maria Martins, Rio de Janeiro)

In 1913, Duchamp became fascinated with a chocolate grinder in action and painted a strict image of it, but his impulse was not satisfied until he abandoned the painting of *images* of real machines and began to make real machines. Thus, he made the "logical" move of selecting manufactured, commercially available objects, and by mounting them in art shows, out of their normal contexts, revealed them as having unexpected and disturbing expressive possibilities. His first "Readymade," a bicycle wheel mounted upside down on a stool, which could be set idly spinning by the spectator, was exhibited in 1913; another of his early Readymades was a snow shovel, which he purchased in a hardware store and called "In Advance of a Broken Arm." The joke was a serious one: it was, for Duchamp, a part of his search to escape image-making, the pretense of illusionism, and a given esthetic. The process of "radical juxtaposition," or displacement of ordinary man-made objects, made people look at such objects with new eyes; but more than anything else, such anti-art gestures by artists were tests of the given or accepted definitions of what constitutes art.

Our century, in fact, has seen many tests of what constitutes art. Successive waves of new works have continually stretched, or even attacked, the definitions of art. Barely was a set of esthetics accepted than it was challenged. Consequently, most people today, it is safe to assume, are somewhat uneasy as to what art is about.

A shotgun of influences

In his interest in new materials, in real objects, in real machines, in movement, in optical questions, Duchamp outstandingly foreshadowed the concerns of artists down to the present day. But there were many others, both before and after him, whose *nonromantic* concern with technology have given dimension to the art-and-technology movement. Even a fragmentary nonchronological listing must include the following influences: Cezanne, with his concept of breaking nature into component spheres and cylinders; Seurat, with his stress on a science of color, practiced with his pointillist technique; Eakins, with his interest in photography; Balla, who tried to paint motion; the development of movies, which were the first new art form to grow out of artistic interest in modern technology; the sculptor Calder with his mobiles; the entire Dada movement, including such artists as Max Ernst and Kurt Schwitters, whose collages are regarded as having provided, as the critic Hilton Kramer says, the "basic visual syntax" for recent artistic explorations such as the Happenings and the art-and-technology movement itself; the physicist-composer Edgar Varese, who groped for a technology to make a new "sound space." (The Futurist Movement of the '20s, although it was greatly impressed with the machine is generally excluded since it celebrated the machine and made of it a romantic concept—the Futurists did not wish to work directly with machines.) More recent influences include: Rauschenberg; Jasper Jones, who showed with his flag paintings what to do with the surfaces of paintings treated as objects; the composer John Cage, with his stress on silences and randomness, so as to draw the listener into a new sound space, a space in which the listener himself becomes a composer or a "realizer"—many of the newest artists acknowledge a keen debt to Cage (think about the observation he makes that appears on page 65); the dancer–choreographer Merce Cunningham, who has explored new directions in dance along with Cage; the sculptor Jean Tinguely whose "philosophical machines" work wondrously in a kind of ecological niche between classically oriented sculptures at the end of the mechanical age and the new interactive works. Still another important influence on the thinking of modern artists has been the philosophy of Wittgenstein.

As we come to the present day, we see a profusion of brilliant and often bewildering works, and many ongoing experiments in the art-and-technology mode. Five of the figures shown on these pages, for instance, were included in a recent exhibition on "The Machine—As Seen at the End of the Mechanical Age" at New York's Museum of Modern Art. Two of these pieces were the results of collaborations between artists and engineers stimulated by a contest sponsored by a new organization for artists and engineers called Experiments in Art and Technology.

Modes of collaboration

When Rauschenberg began in the '50s to combine ordinary objects with his paintings—he used his bed, old doors, objects found on the street, stuffed animals, metals, old radios, automobile tires—critics were at odds as to what to call these new art works. Rauschenberg himself speaks of his search for materials in the streets as a collaboration with the neighborhood in which he was living. For a time, the critics called him a neo-Dadaist, but this label didn't stick. Then, Rauschenberg states, he decided almost as a joke to call these works "combines," a name that did stick, and that he feels allowed people to look at the actual works without getting hung up on questions of what they were supposed to be—paintings, sculpture, collages, or whatever.[7]

Art-and-technology has been called the "new combine,"[8] but until the merger of materials and people has a proper chance to grow its own thing, any name will be premature. As it is now, many of the technological parts and materials still look like the old familiar-shaped actors of the industrial world who have strayed onto a stage in a theater governed by sets of rules that are alien to them, and even comic. They have been set into motion, given parts to play, but haven't yet been mastered by the stage directors.

Standards for the new combine, art-and-technology, do not yet exist. Nonetheless, one can distinguish a number of different modes of juncture between artists and technology, and between artists and engineers.

Artists and new materials. First of all, there are the kinds of works for which an artist does not need an engineer. That is, the artist can acquire the kinds of hardware, materials, and even processes that already exist, and which he can adapt and transform according to his desires and intentions. He may buy such hardware in war surplus stores. Moviemaker VanDerBeek, for instance, states that for artists the greatest boon to come out of the Korean War were the monumental piles of surplus of all kinds of sophisticated equipment. Composer–musician Max Neuhaus, an avid student of engineering magazines, combs the electronics stores for parts with which he constructs his new musical "compositions." The artist may order variations of existing products from industrial suppliers, as Rauschenberg did recently in acquiring new sets of sliding doors for an "environmental" piece called "Solstice."

An elegant and almost invisible mode of using available new materials appears in the wind and water sculptures of the artist Hans Haacke, who seeks in his quietly moving pieces an analog of life.

Haacke worked for several years perfecting techniques of using acrylic cements and simple Plexiglas to create water sculptures—precision-made Plexiglas containers holding liquids—in which the liquids evaporate and condense on the walls, forming condensation trails, which change with time and with environmental changes. In his search for works that are so lightweight as to be nearly immaterial, Haacke has gone on to make wind sculptures—he has become a "wind player," through manipulation of wind sources (fans, etc.), suspensions of fabric and balloons. The works are deceptively simple. One might comment on them by noting that Haacke's bookshelf is crammed with textbooks on hydraulics and aerodynamics, sub-

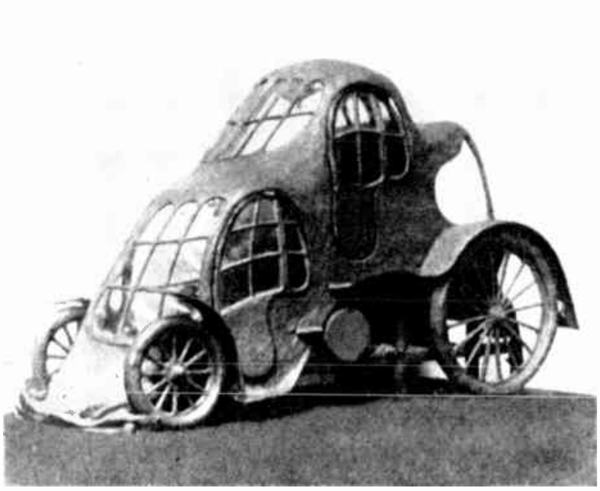
Art instead of being an object made by one person is a process set in motion by a group of people. Art's socialized. It isn't someone saying something, but people doing things, giving everyone (including those involved) the opportunity to have experiences they would not otherwise have had.

jects in which he has become an earnest student.

Artists and new processes. Another approach is to engage the available processing and productive capabilities of industry to make works according to the artist's specifications. For instance, artists such as Donald Judd, Robert Morris, Ronald Bladen, and others, who work in the "minimal" or "reductive" school, create a kind of factory sculpture, and consequently are involved in essentially managerial problems. For the kind of austere boxlike forms that he has made, Judd specifies the materials, schedules shipments, and oversees production at the plant rather than in his studio. Art critic James R. Mellow, in writing about Judd's work, notes that while such a procedure suggests "a correspondence between his work and that of the industrial designer who encounters similar problems, Judd sees nothing special in the comparison. Given the nature of the work he wants to produce and the nature of the materials he is dealing with, these are simply the conditions of his art."[9]

Still other artists have gone directly into the factories and industrial and research organizations to create works sometimes with the assistance and advice of technicians and engineers. In such a collaborative venture, the sculptor Piotr Kowalski made a stainless-steel flowerlike work by an underwater-controlled explosion technique that is used by North American Aviation, Inc., to weld and shape high-density metals.[8]

EARLY efforts to effect a remarriage between art and technology brought about some curious and sometimes fascinating results. This maquette of a car by Pierre Selmersheim won a prize in a contest in 1895.*

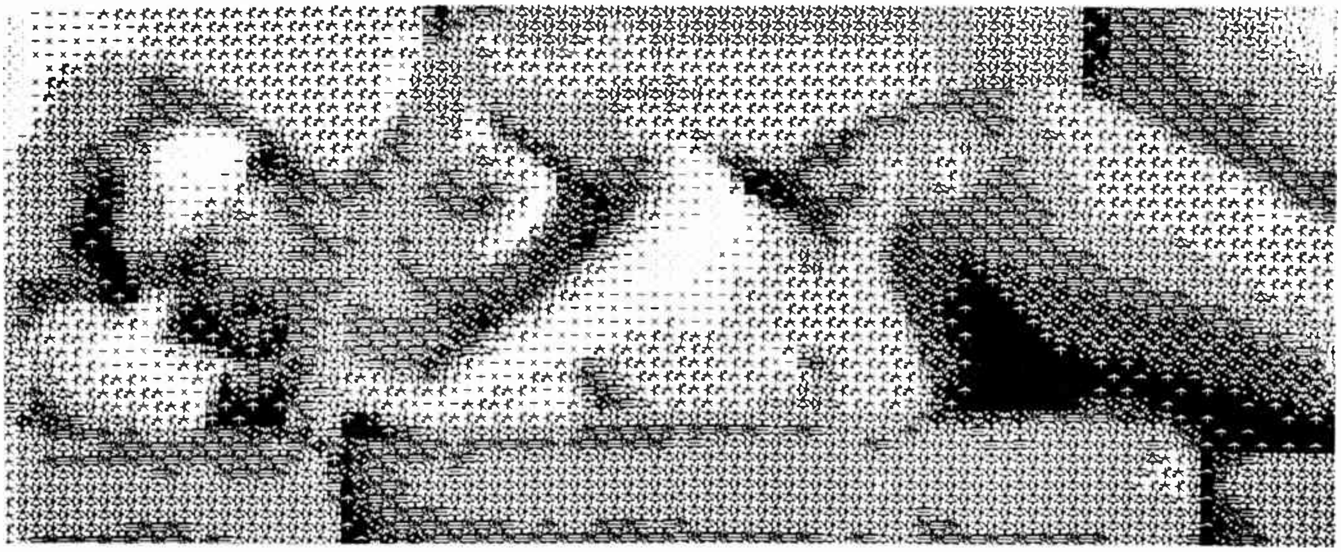


Artistic interest in industrial processes is often thwarted, however, as in Harold Tovish's "Vortex" (frontispiece). Although "Vortex" looks machine-turned, the artist was obliged to turn it by hand with wet plaster. Then the form was cast in bronze. Blocked by the high cost of precise machine work, Tovish, now a Fellow at M.I.T.'s Center for Advanced Visual Studies, remarks ironically, "An artist's time being worthless, the only thing I could squander generously was my time in laborious handwork."[10]

To combat this sense of social worthlessness, artists have begun to band together to gain the industrial cooperation they cannot gain easily as individuals. One such group was formed more than a year ago by 20 arists in southern California, the home of some of the largest aerospace companies. Called the Aesthetic Research Center (ARC), it grew out of the artists realization that they had to move out of the studio and into the industrial environment.

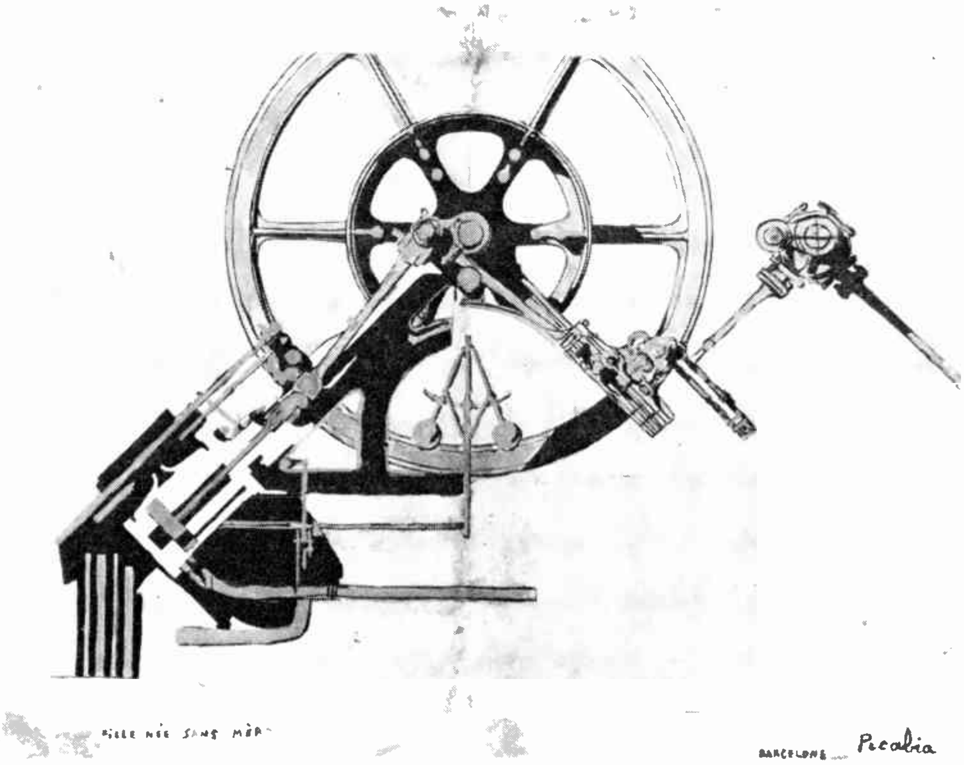
One of the most recent artist–industrial cooperative programs, which was announced late in December, is that organized by the Los Angeles County Museum of Art, which aims at placing 20 major artists in residence in some of the world's largest and most advanced technologically oriented corporations. The Art and Technology Program of the museum, directed by Maurice Tuchman, has already brought about a number of artist–industry commitments: Pop artist Claes Oldenburg, the creator of "soft sculptures," will take up residence and use the facilities of Walt Disney Productions; painter Robert Irwin and light artist James Turrell will collaborate on an experimental project at the Life Science Division of Garrett Aerospace; Robert Morris, influential in founding minimal art, will take up residence in a major aerospace industry; Robert Rauschenberg will take up residence in a Los Angeles electronics firm; Pop painter Roy Lichtenstein will work with Universal City Studios; sculptor Larry Bell will take up residence at the RAND Corporation, where he will work on theoretical artistic problems involved in decision making; and the French artist Jean Dubuffet is to take up residence at the American Cement Company in California, where he proposes creating an ambitious sculptural monument involving extraordinary architectural and engineering resources. Among the twenty-six corporations participating in Tuchman's Art and Technology Program are the Container Corporation of America, Jet Propulsion Laboratory, International Chemical and Nuclear Corporation, Kleiner-Bell Foundation, Ampex Corporation, IBM, Litton Industries, Lockheed Aircraft, Teledyne, Kaiser Steel, Philco-Ford, Hewlett-Packard, TRW Systems, and others.

Artists and computers. The computer holds so extraordinary a place in our modern technological and intellectual environment—it is not only a unique but also a deeply symbolic system development—that it holds special attractions for the artist. One could, without much reservation, compare its potential now to the potential of the movie camera at the beginning of this century. In its "plasticity," the computer resembles older media; in its active and interactive capabilities,



TWO PERCEPTIONS of the female.

"Studies in Perception, I," 1968, above,* a computer-processed photographic print by Leon Harmon and Kenneth Knowlton of Bell Labs, was produced by scanning a photo with a televisionlike camera and converting the output into levels of brightness that were printed out with patterns rather than dots. (Owned by the artist and engineer.)

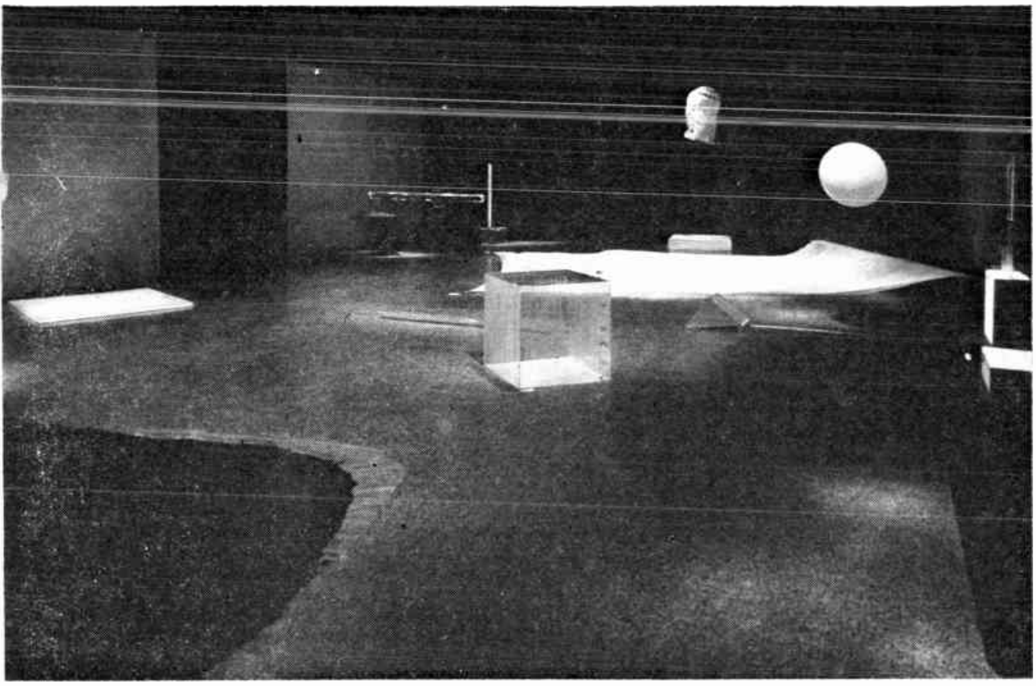


"Fille née sans mère," c. 1917, right,* by Francis Picabia is a gouache on a railway-machine diagram. After a visit to the U.S. in 1915, Picabia said: "Almost immediately . . . it flashed on me that the genius of the modern world is in machinery and that through machinery art ought to find a most vivid expression." The "girl born without a mother" perhaps refers to the machine, a female being, created by man. (Collection Mr. and Mrs. Arthur A. Cohen, New York)

it opens the door for wholly new artistic directions.[11] We already hear a lot about "computer art" of a limited kind: as computers get cheaper and as their interactive languages grow richer, we shall be hearing much more.

At the Bell Telephone Laboratories, for instance, which has been particularly prominent in its receptivity to and sponsorship of collaborations with artists, the computer facilities have already been used for music composition, for movies, and for graphical output designs.

Among the composers and musicians who have gone to Bell Labs to explore the potentials of the computer medium have been Edgar Varese, Herman Scherchen, Milton Babbitt, and James Randall of Princeton's music faculty, and Vladimir Ussachevsky of Columbia University, who is currently doing research in all aspects of electronic music. Another composer who worked on computer music at Bell is James Tenney, who has the rather unusual double talent and advantage of being an electrical engineer as well as a composer: he is now a faculty member in the Electrical Engineering Department at the Polytechnic Institute of Brooklyn. Tenney thus embodies a principle that the opponents of collaboration urge—they argue, in effect, that collaborations of any significance will be extremely difficult and frustrating to bring about, and that the artist must also be a trained engineer if he is to make anything serious out of the technological medium.



Environmental sculptures by artist Hans Haacke exhibited at M.I.T. in 1967. Haacke says that he thinks of these condensation boxes and wind sculptures, which depend upon a relatively simple technology, as "systems." He describes them as being opaque in the sense that they hold no meaning of their own; that is, one does not seek to "interpret" them. Haacke has also recently created an interactive room, using commercial photocells supplied by an interested manufacturer, Automatic Timing and Controls, Inc. The room senses the presence of a person; lights automatically illuminate him wherever he stands or as he moves about.

The application of computers in the arts no doubt presents both special advantages and special problems. Tenney, for instance, points out that with traditional musical intsruments the composer is limited by certain tones around which he must compose, but with the computer he no longer needs to live with such limitations: theoretically, *any* sound can be produced with computer, so that the composer can in this medium get any kind of continuity of "line" he wishes. This great openness in the sound space, however, means that the artist must discover the limitations of the new medium at the same time that he is trying to discover how he chooses to strive against those limitations. Thus, he is thrown back even more sharply against his own limitations than in the case, say, of painting, where there exists a long history of efforts and esthetic standards, and in which each new painting embodies a consciousness of the history of painting.

Some of the limitations of computers have been discovered already by the artist. He finds, first of all, that he must make some kind of tradeoff in the amount of energy he is willing or able to expend on programming to get the results he might wish. That is, although theoretically any kind of sound is possible, some sounds may cost him far more than they are worth. And although Tenney, an engineer, can offer the contrast that it took him about a year to learn to work and compose in the computer medium, whereas it took him 15 years to learn to work with an orchestra, other artists may find that learning to program computers for their own ends may be considerably more onerous, at least until languages are developed specifically for artistic ends.

The question of suitable computer software for the needs of different kinds of artists futher illuminates the differing depths of collaboration. An artist may learn existing programs, with some guidance from an engineer, or he may work very closely with an engineer in creating special new programs. For instance, Dr. Max V. Mathews, director of the Behavioral Research Laboratory at Bell Labs, has been developing for some time a graphical input language for music composition that a user could employ fairly readily and without the need for in-depth instruction.[12] Artists Stockhausen and Varese have experimented with it—but even children could be taught to use it.

In another instance, engineer Kenneth C. Knowlton, at Bell Labs, has worked for nearly two years with moviemaker Stan VanDerBeek experimenting on new languages specifically geared to the composition of computer-animated movies. VanDerBeek, who has suffered through the effect of learning programming (he says that learning to program was like trying to learn to draw with the pencil attached to the end of his nose), now points out that he is able to work on his moviemaking no matter where he is—simply by writing out his programs. But the process is still clearly experimental. For the one-minute film "Man and His World" that VanDerBeek and Knowlton prepared for Expo 67, the basic designs were done by "a special set of higher-level macros" and subroutines, but the color had to be painted on the film by hand.

In discussing the software problems of compter art at a special session on "Computer Output as Art" at the 1968 IEEE International Convention, Knowlton judged that the present experiment would lead eventually to a "number of relatively stable languages" for artists, each for a distinct medium. Such a medium he stated, may "become sufficiently established and familiar—no longer a cute gimmick—that artists can use it to 'say something' without the medium itself arousing such curiosity, acclaim or disdain as to severely distract from the artistic content of the work."[13]

Experiments on other forms of computer-animated films are going on also at M.I.T.; these have an immedi-

"Toy-Pet Plexi-Ball," 1968, created by Robin Parkinson and Eric Martin, utilizing Plexiglas, electric equipment, motor, microphone, and a synthetic fur bag, was included in a recent exhibition at New York's Museum of Modern Art.* The result of a collaboration between an artist and engineer, it was one of nearly 200 works stimulated by a contest sponsored by a new organization for artists and engineers called Experiments in Art and Technology. The "Toy-Pet Plexi-Ball," which has three "eyes" and one "ear" that respond to light and sound, rolls about through complex evasive maneuvers. Only when it is in its fur bag does it remain dormant.

ate-playback feature, so that the frames may be edited with a light pen at will. In this sense, the computer (the TX-2) is being used rather more like traditional pencil and paper.

At the same 1968 IEEE session, George A. Michael, of the University of California Lawrence Radiation Laboratory, in discussing hardware trends for visual displays that might be used by artists, noted that the simple display graphic devices being developed for scientific and office purposes are not very versatile and he wondered whether or not artists would wish to use them seriously in their evolving forms. He noted, for instance, the limitations of such displays: colors cannot be used; large murals are impossible; modifications of texture through light adjustment is difficult; and high resolution for shading effects is still out of reach. Michael noted that cleverer programming will be needed to exploit the maximum capabilities of the computer medium, and that if artists are "to do their thing" they must become embroiled with the engineers and programmers in the details of creating programs.[14] Or else, as with VanDerBeek, who found in Knowlton a willing and compatible collaborator, they must find engineers who are willing and able to engage in a deep-going collaboration in which each partner learns from the other as they go along.

Artists and engineers. Potentially the most interesting collaboration is that in which the artist and engineer create together, with unconstrained choice, guided by the artist's sensibility and concepts and by the engineer's knowledge, insight, and mastery of the technological components and physical laws. Such collaboration stands in contrast to the mere exploitation of new material and equipments, which has slowly and steadily and sometimes almost imperceptibly been going on anyway.

On this kind of collaboration, the testimony of an engineer who has been working with many artists over many years is well worth considering.[15] Dr. Billy Klüver, mentioned in the introduction for his role on "Homage to New York," who has worked with filmmaker Andy Warhol, composer John Cage, painter Jasper Johns, and others, concludes: "You do things you wouldn't normally do, because you're in touch with a mind whose vision is totally different from yours. The artist's vision and concern relate to other aspects of human activity, and that's the end that particularly interests me. I'm not so much interested in helping artists as I am in seeing what effect the artist could have on technology. In the future, I see the artist having more and more impact, as he learns more about technical processes. The contribution of the artist could conceivably lead to an increased awareness, a new view of the problems the engineer, designer, scientist has to deal with. For instance, it might reflect on questions like: What should the next mass media look like? We will have the megalopolis: what is it going to look like? I think that the main influence of art and technology together will come in the area of environment." [16]

Whether or not one prefers to regard such deep-going collaborations as the reassertion of an older tradition or as a genuine "new combine," one must recognize that this the art-and-technology movement does hold the seeds of radical possibilities, not just for the artist but for the engineer as well. Some recent experiments and experiences with this form of collaboration will be treated in Part II.

REFERENCES

1. Pontus Hultén, K. G., *The Machine as Seen at the End of the Mechanical Age.* New York: The Museum of Modern Art, 1968.

2. Singer, C., Holmyard, E. J., and Hall, A. R., *A History of Technology.* London: Oxford Univ. Press, 1954.

3. McLuhan, M., *Understanding Media: The Extensions of Man.* New York: McGraw-Hill, 1965.

4. Rubin, W. S., *Dada, Surrealism, and Their Heritage.* New York: The Museum of Modern Art, 1968.

5. Sontag, Susan, "Against interpretation," in *Against Interpretation.* New York: Dell, 1965, pp. 3–14.

6. Sontag, Susan, "One culture and the new sensibility," in *Against Interpretation.* New York: Dell, 1965, pp. 293–304.

7. Kostelanetz, R., "A conversation with Robert Rauschenberg," *Partisan Rev.*, vol. 35, pp. 92–106, Winter 1968.

8. Davis, D. M., "Art and technology," *Art in America*, vol. 56, pp. 28–47, Jan.–Feb. 1968.

9. Mellow, J. R., "Everything sculpture has, my work doesn't," *The New York Times*, p. D21, Mar. 10, 1968.

10. Tovish, Harold, Private communication.

11. Noll, M., "The digital computer as a creative medium," *IEEE Spectrum*, vol. 4, pp. 89–95, Oct. 1967.

12. Brodey, W. M., and Lindgren, N., "Human enhancement—Part II: Beyond the machine age," *IEEE Spectrum*, vol. 5, pp. 79–93, Feb. 1968.

13. Knowlton, K. C., "The problem of software for computer art," *1968 IEEE Internat'l Conv. Digest*, Mar. 20, 1968.

14. Michael, G. A., "Hardware trends for artistic visual displays," *1968 IEEE Internat'l Conv. Digest*, Mar. 20, 1968.

15. Kluver, J. W., "Nine evenings of theatre and engineering: notes of a participant," *Artforum*, vol. 5, Feb. 1967.

16. "Art and science: two worlds merge," *Bell Telephone Mag.*, pp. 13–19, Nov. Dec. 1967.

The early history of electronics

III. Prehistory of radiotelegraphy

Before radiotelegraphy could become a reality, it had to be thought of; a workable detector of electromagnetic waves had to be discovered; and the results obtained by Hertz had to be disseminated to a wider audience

Charles Süsskind *University of California*

In earlier installments in this series, we saw how the work of Hertz and his predecessors had led an electrical engineer to ask whether waves could not be used for the transmission of power or of telephone messages.[1] That suggestion had been made in private in 1889, but other, better-founded proposals began to find their way into public print almost immediately.

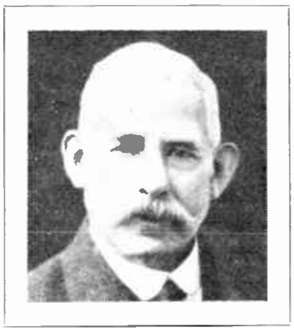
First proposals

The first such proposal to come to our attention was made by Threlfall in Australia in 1890. Upon being elected to the presidency of Section A of the Australasian Association for the Advancement of Science, Threlfall delivered an address at Melbourne in which he reviewed the state of electrical science, described Hertz's results in connection with the parabolic cylindrical mirror experiments, and suggested that the difficulties of the large receiver resonator size might be obviated. "It may be questioned whether greater accuracy might not be obtained by the use of Geissler tubes, coupled with some system of photography," he said. "These tubes have been already successfully applied in Dr. Lodge's laboratory, and if it be permissible to prophesy wildly we may see in this observation the germ of a great future development. Signalling, for instance, might be accomplished secretly by means of a sort of electric ray flasher, the signals being invisible to anyone not provided with a properly tuned tube."[2]

In the following year, an editorial in the influential British periodical, *The Electrician*, commented anonymously on proposals that communication with lighthouses and lightships should be maintained by "wireless" means employing conduction through water. The editor during 1890–1895 was Alexander Pelham Trotter (1857–1947), a Cambridge graduate who was primarily an illumination engineer. "The obvious way of communicating between a lightship or a lighthouse and the coast would be by flash signals," the editorial stated, "but a slight mist would effectually stop the beam of light, since mist is opaque to light . . . because the frequency of the vibrations is high. . . . Hertz has shown how slower vibrations may be received and observed, but the wave length of such vibrations is at least a foot. . . . But if we could reduce the frequency about 2,000 times and produce vibrations at about two hundred thousand million per second, we should get waves about one millimetre long. These radiations would probably pierce not only a fog, but a brick wall. *When* we get such vibrations there will be many interesting uses for them. One, at all events, would be the possibility of communicating between lightships and the shore."[3]

It is not clear what calculation led the writer to conclude that millimetric wavelengths would be required to pierce a fog. However, it should not be thought that the employment of what (much later) came to be called microwaves was considered to be such an exotic idea at that time as it would have been a little later. In a survey of microwave antenna and waveguide techniques before 1900, Ramsey has shown that Hertz and his successors developed devices and techniques that anticipated much of the practice of microwave engineering.[4]

An even more accurate prediction was made in 1892 by Crookes. Writing on "Some Possibilities of Electricity" in a London nontechnical periodical edited by the bon vivant and littérateur Frank Harris (1854–1931), Crookes proposed the use of electromagnetic waves for telegraphy and enumerated the three main requirements—reliable transmitters, sensitive tuned receivers, and directional antennas—stopping well short



(Sir) Richard Threlfall (1861–1932) was born in England and studied at Cambridge, where he became J. J. Thomson's assistant before accepting the professorship of physics at the University of Sydney (1886–1898). He then returned to England, where he worked as a consulting engineer and electrochemist and also served on several government boards. He was knighted in 1917 for his work on munitions and food preservation during the 1914–18 war.

(Sir) William Crookes (1832–1919) was born in London and was trained in chemistry, a field in which his reputation became firmly established when he discovered thallium in 1861. His invention of the radiometer, in which a set of vanes blackened on one side and polished on the other is caused to turn under irradiation, attracted much attention, as did his investigations of electrical discharges in rarefied gases. He noted the dark space (now named for him) around the negative electrode and investigated the nature of the emanating cathode rays, which he believed to be charge-carrying gaseous molecules. His observation that the impact of alpha particles is accompanied by a scintillation is the basic principle on which scintillation counters are based. He was knighted in 1897. The breadth of his interests is reflected in his service as president, at various times, of the IEE, the Society of Chemical Industry, the Chemical Society, the British Association, and the Royal Society. (See Fournier D'Albe, E. E., *Life of Sir William Crookes.* London, England: Fisher Unwin, 1923.)

of millimetric wavelengths. This prescient analysis deserves to be cited at some length:

"Rays of light will not pierce through a wall, nor, as we know only too well, through a London fog. But the electrical vibrations of a yard or more in wavelength of which I have spoken will easily pierce such mediums, which to them will be transparent. Here, then, is revealed the bewildering possibility of telegraphy without wires, posts, cables, or any of our present costly appliances. Granted a few reasonable postulates, the whole thing comes well within the bounds of possible fulfilment. At the present time . . . an experimentalist at a distance can receive some, if not all, of these rays on a properly-constituted instrument, and by concerted signals messages in the Morse code can thus pass from one operator to another. What, therefore, remains to be discovered is—firstly, simpler and more certain means of generating electric waves of any desired wavelength, from the shortest, say of a few feet in length, which will easily pass through buildings and fogs, to those long waves whose lengths are measured by tens, hundreds, and thousands of miles; secondly, more delicate receivers which will respond to wavelengths between certain defined limits and be silent to all others; thirdly, means of darting the sheaf of rays in any desired direction . . .

"At first sight an objection to this plan would be its want of secrecy. Assuming that the correspondents were a mile apart, the transmitter would send out the waves in all directions, filling a sphere a mile in radius, and it would therefore be possible for anyone living within a mile of the sender to receive the communication. This could be got over in two ways. If the exact position of both sending and receiving instruments were accurately known, the rays could be concentrated with more or less exactness on the receiver. If, however, the sender and receiver were moving about, so that the lens device could not be adopted, the correspondents must attune their instruments to a definite wavelength, say, for example, fifty yards. I assume here that the progress of discovery would give instruments capable of adjustment by turning a screw or altering the length of a wire, so as to become receptive of wavelengths of any preconcerted length . . .

"This is no mere dream of a visionary philosopher. All the requisites needed to bring it within the grasp of daily life are well within the possibilities of discovery, and are so reasonable and so clearly in the path of researches which are now being actively prosecuted in every capital of Europe that we may any day expect to hear that they have emerged from the realms of speculation into those of sober fact. Even now, indeed, telegraphing without wires is possible within a restricted radius of a few hundred yards, and some years ago I assisted at experiments where messages were transmitted from one part of the house to another without an intervening wire by almost the identical means here described."[5] (That was a reference to the experiments of Hughes described in our first installment; it was as a result of this mention that Fahie first learned of the experiments and brought them to light.)

This 1892 article was a remarkable prediction and a correct analysis of the principal obstacles that would have to be overcome before radiotelegraphy could become a reality. The low sensitivity of the receivers was the greatest obstacle; no progress was possible until a better means of detecting electromagnetic waves would become available. Such a device soon appeared. It was the *coherer*, which was to determine much of the progress of the new technology for a couple of decades.

During the early 1890s, many experiments based on the researches of Hertz were performed, notably in the field of what later came to be called microwave optics. Perhaps the most important from the viewpoint of the future development of radiotelegraphy was the work of Augusto Righi (1850–1920) at Bologna, who published an entire book in 1897 on the subject of the optics of electrical oscillations, *L'ottica delle oscillazioni elettriche*.[6] A German translation by Bernhard Dessau (1893–1936), *Die Optik der elektrischen Schwingungen*, appeared the next year (O. R. Reisland, Leipzig, 1898) and contains two additional papers on dielectric measurements at centimetric wavelengths. The two men later coauthored one of the early textbooks on radiotelegraphy, which was published simultaneously in Italian and German versions in 1903:

Telegrafia senza filo, Nicola Zanichelli, Bologna: and *Die Telegraphie ohne Draht*, Friedrich Vieweg und Sohn, Braunschweig.

Other quasi-optical experiments were carried out by Édouard Sarasin (1843–1917) and Lucien de La Rive (1834–1924) at Geneva,[7] Antonio Giorgio Garbasso (1871–1933) and Emil Aschkinass (1873–1909) at Berlin,[8] and J. C. Bose at Calcutta,[9] and several other investigators; the aforementioned survey by Ramsay[1] contains accounts of a number of additional investigations. But these experiments did not directly affect the development of radiotelegraphy; nor did a suggestion made by Nikola Tesla (1856–1943) that it might be possible to transmit "intelligible signals or perhaps even power" over large distances without wires if sufficiently high frequencies were employed.[10] Tesla had the curious idea to use changes in the electrostatic potential of the earth's surface for signaling, an impractical proposal reminiscent of the Dolbear scheme described in our first installment.[1]

The coherer

The history of the coherer is one of discovery and rediscovery. The basic phenomenon was known for many years before it was applied for the detection of electromagnetic waves. The observation that dust particles cohered into strings or flakes in the vicinity of electrostatic generators was apparently made as early as 1850.[11] The problem next engaged the attention of Samuel Alfred Varley (1832–1921), who noted in 1866 that finely powdered metal presented a high resistance to the passage of small currents but became a very good conductor when the voltage was increased. Varley constructed a lightning protector that consisted of a box of carbon powder mixed with a nonconducting powder; two wires were inserted into the powder until they almost touched. Placed across a telegraph instrument, the device had no perceptible effect under normal operating conditions, but became a bypass shunt when lightning struck. Varley said that he had "no doubt, from an examination of the bridge afterwards, that under the influence of a high tension discharge connexion was made between the two metallic points by a bridge of conducting matter, arranged closely together, and that if the instruments had been shaken to loosen the powder, all would have been put right."[12]

In 1879, Lord Rayleigh (1842–1919) investigated the behavior of water particles when the vicinity of the breakup point of a vertical jet of water was electrified. He found that when an electric charge was brought near, the fine spray changed into much larger drops. In addition to its relevance to coherer action, this experiment served to elucidate such meteorological phenomena as the formation of thundershowers.

Still another experiment was performed in 1884 by Temistocle Calzecchi Onesti (1853–1922), who placed copper filings between two brass plates and found that the resistance changed from a very large value to a low one when a substantial current was passed through the device.[13] He next constructed it in the form of a hollow glass tube with metal plugs at both ends and filled with filings of various metals; he found that if the tube were rotated about its axis after each passage of current, the filings "decohered" and the experiment could be repeated. He also showed that contact with an electrified body or an electrostatic discharge made the device conductive.

(Sir) Jagadis Chunder Bose (1858–1937) was born in Bengal and was educated in India and at Cambridge. He was professor of physical science at the Presidency College, Calcutta, during 1884–1915, and experimented (among other things) with electromagnetic waves and their detection. He designed a generator of millimetric waves and systematically examined many materials for coherer action, which he believed depended on a molecular stress-and-strain effect that also served to explain fatigue phenomena in biological substances. These considerations led him to postulate a theory of the similarity of response in the living and nonliving. He was knighted in 1917 and was made a Fellow of the Royal Society in 1920. The Bose Premium of the Institution of Electronic and Radio Engineers is named in his honor.

Other possibly relevant experiments were those of Clark and Lodge, who were studying the phenomenon, first discovered by John Tyndall (1820–1893), that when a hot body is brought into a dusty chamber, a dust-free region appears over the body.[14] Tyndall thought that the phenomenon depended on combustion, but it was later found that a more subtle molecular action was the cause.[15] Clark and Lodge now investigated the possibility that electrical forces might be at work. They found that if a metal rod inserted into a smoky atmosphere was electrified, the atmosphere cleared up at once. They had discovered (or rediscovered) an electrostatic dust precipitator.[16]

A little later Lodge was working on a transmission-line lightning protector formed by the spark gap between the two metal spheres. When a spark passed between them they were sometimes fused together. Lodge suspected that some action other than a purely thermal one was involved, but neither in his paper on the subject[17] nor in the long discussion that followed its presentation[18] did he put cohesion down as an explanation; nor did he mention it in a subsequent paper[19] suggesting how spherical radiators may be excited by sparks. Not until 1893 did he mention cohesion, in a discussion of a paper by George Minchin

(Sir) Oliver Joseph Lodge (1851–1940) was born in Penkhull in Staffordshire. He was the brother of two distinguished historians, Eleanor Constance Lodge (1869–1936) and Sir Richard Lodge (1855–1936). He studied in London, receiving the D.Sc. in 1877 and serving as demonstrator at University College until 1881, when he became professor of physics at the newly established University of Liverpool. His principal contributions to electricity were made during his tenure at Liverpool, where Joseph Warner Clark (1856–1885) was demonstrator and later assistant professor during 1883–1885. In 1900 Lodge became principal of the new University of Birmingham, a post he held until his retirement in 1919. He was knighted in 1902 and served at various times as president of the Physical Society, the British Association, and the Society for Psychical Research, an interest that dominated his life after his young son Raymond was killed early in the war of 1914–1918. Lodge died at Lake near Salisbury, England, on August 22, 1940.

Minchin (1845–1914) on the action of electromagnetic radiations on films containing metallic powders.[20] In the following year, on June 1, 1894, Lodge gave a historic lecture before the Royal Institution on "The Work of Hertz and Some of His Successors," in commemoration of Hertz, who had died five months earlier. The lecture was widely reported,[21] reprinted verbatim in *The Electrician*,[22] and afterwards issued as part of a book that went into several editions.[23] It was in this lecture that Lodge first used the term *coherer* and, by referring to this earlier work on the measurement of standing waves along wires, on the dust precipitators, and on lightning protectors, managed to suggest that he had not only come close to anticipating Hertz, but had also been the first to employ a coherer for the detection of nearby sparks.

It is quite clear that the man who really did so was the French physicist Branly and that Lodge, for all his brilliant popularizations, was a relative latecomer. Lodge's principal contribution to radiotelegraphy, a method of resonant tuning described below, would have been sufficient to ensure him a lasting place in the history of electronics; he need not have attempted to carve out a greater niche for himself.

Branly noticed that the resistance of a glass tube filled with metal filings, normally of the order of several million ohms, was reduced to a few hundred ohms when an electric discharge occurred in the vicinity. He was a meticulous experimenter. He noted that the high resistance could be restored by heating or mechanical shock, and that the change did not occur when the equipment was enclosed in a metal box. After investigating these effects in a number of materials, he reported his findings in both a scientific and a technical journal, but in neither article did he propose the use of the effect for the detection of electromagnetic waves, much less for communications, although he worked on such applications later.[24] It should be remembered that Branly's first experiments took place only two or three years after those of Hertz.

Branly's technical article appeared almost at once in an abbreviated form in English translation,[25] but seems to have produced little immediate interest: physicists do not habitually read engineering journals and other readers might have overlooked the articles, as Lodge ruefully confessed later, "probably by reason of their coincidence with the holiday season."[26] But a little later, Dawson Turner (1857–1928) visited Paris, *saw* Branly's experiments, repeated some of them, and reported on the results at the British Association meeting in Edinburgh.[27] This report, and a subsequent demonstration in London,[28] excited the interest of Minchin, who saw an analogy with the resistance drop owing to the action of light on photoconductive cells. He now recalled that electromagnetic radiation from electric sparks had also affected the sensitivity to light of such cells in some experiments that he had made some years before,[29] in a way that varied inversely with wavelength; and that the effect ceased when the cell, *together with its terminal wires*, was enclosed in a metal box. He proposed the use of thin films of fine metallic powders embedded in gelatine on collodion as being more reliable. Lodge saw the paper before it was published and wrote the aforementioned quick note that appeared appended to Minchin's paper.[30] It was in this note, communicated on November 23, 1893, that Lodge first mentioned cohesion.

The important part that Branly's coherer played in the development of radiotelegraphy was certainly appreciated by his contemporaries. One of the first radiotelegraphic messages to cross the English Channel (on March 28, 1899) was addressed to Branly and sent him "respectful compliments across the Channel, this fine achievement being partly due to the remarkable researches of M. Branly."

An interesting sidelight is that Branly was the first to use the word *radio* in connection with electrical communications. When Lodge, in his Royal Institution lecture, lumped Branly's device with some others of his own devising under the name *coherer*, the name stuck, much to the discomfiture of Branly: "My tubes of filing have been named *coherers* by Lodge, a name that has been generally accepted," he wrote. "This expression is based on an incomplete investigation of the phenomenon and on an inexact interpretation; I have proposed the name *radioconductors*, which recalls the essential property of discontinuous conductors to be *excited* by electric radiation." This note appeared on

Edouard Branly (left) (1846–1940) was born at Amiens, the son of a professor. He was graduated from the Ecole Normale Superieure in 1865 and completed postgraduate study in 1867. After teaching in secondary schools and being in charge of the physics laboratories at the Sorbonne, he became the first professor of physics at the Institut Catholique when it was established in Paris in 1875. He also qualified as a physician and practiced for several years. His electrical researches, begun in 1889, earned him the Houllevigue Prize of the Academy of Sciences in 1898, and he was elected a member in 1911, narrowly edging out Marie Curie (1867–1934), the discoverer of radium, who received 28 votes to Branly's 30. When Branly died, he received a state funeral.

(Sir) Henry Bradwardine Jackson (1855–1929) experimented with ship-to-ship radiotelegraphic communications while captain of *Defiance* in 1895–1896. He was subsequently the first to observe interference between signals traveling from transmitter to receiver over different paths. He was First Sea Lord of the Admiralty in 1915–1916, president of the Royal Naval College at Greenwich in 1916–1919, and in 1920 he became the first chairman of the Radio Research Board of the Department of Scientific and Industrial Research. He became a Fellow of the Royal Society in 1901 and received its Hughes medal in 1926. He was knighted in 1906.

December 6, 1897, and is very likely the first use of the term *radio* in the aforementioned context.[31]

The Branly coherer was used in the early demonstrations of Hertzian waves. Its derivatives dominated radiotelegraphy during the first decade of its history and continued to be used in many installations for years after that. Several improvements were subsequently made in coherer design, notably with a view to evolving one that would be "self-cohering" without tapping. Minchin proposed a coherer that employed a platinum wire dipping into a pool of mercury, the whole contained in an exhausted glass envelope.[32] Bose tested many metals and found that most cohered, but only potassium was also self-decohering, provided it was immersed in kerosene to avoid oxidation.[33] A self-restoring carbon coherer was devised by Thomas Tommasina and saw much use.[34]

There is no doubt that it was the coherer that turned the simple scheme of Hertz into a technologically significant method of communication.

Early experiments

In the summer of 1894, Lodge followed his Royal Institution lecture with a similar demonstration lecture when the British Association met at Oxford.[35] This second lecture was entitled "Experiments Illustrating Clerk Maxwell's Theory of Light" and featured essentially the same experiments as the Royal Institution lecture on "The Work of Hertz." Lodge described various detectors and again demonstrated several, including an electroscope method proposed by Boltzmann,[36] coherers of his own design, and a Branly coherer. Among his own coherers was the spark-gap coherer consisting of a pair of spheres, whose action was demonstrated in a remarkably well-designed experiment. The coherer was mounted in series with a battery and an electric bell. When a spark occurred in the vicinity, the spheres cohered together, the circuit was closed, and the bell began to ring and continued to ring until the spheres were separated by a gentle tap. Lodge found that if the bell were mounted on the same stand as the coherer, the mechanical vibration of the first stroke of the bell would be sufficient to tap the spheres apart, so that each spark would be signaled by a single stroke of the bell. This method of automatically "decohering" was to be widely copied in early apparatus. But even more important from the viewpoint of future development was the method used by Lodge in both lectures to reproduce Hertz's transmitter and receiver. He employed two Leyden jars, the inner and outer coating of each connected to a long wire; the two pairs of wires were of comparable length and one pair was provided with a slider by means of which the length could be accurately adjusted until the two circuits were in resonance.

This arrangement, which Lodge called the *syntonic-jar* experiment, gave a much more striking demonstration of what was later to be called resonant tuning than Hertz's original circuits, which were so heavily damped that the first cycle in a wave train was probably the only one effective in producing radiation. Lodge had been experimenting with resonance for several years[37] and he was well aware of the superiority of his arrangement, which produced persistent vibrations and sharp tuning. During the course of his demonstration, he said, "Here we have in essence a system of a very distinctly syntonic telegraphy, for the jars and their circuits must be accurately tuned together, if there is to be any response. A very little error in tuning, easily made by altering the position of the slider, will make them quite unresponsive, unless the distance between them is reduced. At the maximum distance of response the tuning required is excessively sharp."

But Lodge evidently made no attempt to transmit intelligence; he did not even attempt to measure the distance to which his equipment would remain effective. Referring to the two lectures in a paper on the history of the coherer written a few years later, he wrote: "In both cases signalling was easily carried on from a distance through walls and other obstacles, an emitter being outside and a galvanometer detector inside the room. Distance without obstacle was no difficulty in these experiments, only free distance is not very easy to get in a town, and stupidly enough no attempt was made to apply any but the feeblest power so as to test how far the disturbance could really be detected."[38] Still later, he admitted: "I was too busy with teaching work to take up telegraphic or any other development; nor had I the foresight to perceive, what has turned out to be, its extraordinary importance to the navy, the merchant service, and indeed land and war service, too."[39]

That is not to say that Lodge was not quick to recognize the importance of tuning the receiver to the transmitter once others had applied the transmission of electromagnetic waves to radiotelegraphy. Although the syntonic-jar arrangement was not patentable under British law, since it had been published, Lodge later successfully applied for a U.S. patent (on December 20, 1897); when the patent was finally granted after much discussion, it made specific reference to his work of 1894.[40] Lodge also secured a British patent on a much more elaborate system of tuned or selective telegraphy in 1897.[41] Fourteen years later, this patent became the subject of litigation when the syndicate formed by Lodge and his collaborator Alexander Muirhead (1848–1920) petitioned for a prolongation over the objections of competitors. The petition was successful: the patent was established as the fundamental tuning patent and extended for another seven years, largely owing to the efforts of S. P. Thompson, who had prepared a well-documented chronological account of Lodge's contributions to tuned radiotelegraphy.[42]

In addition to Muirhead, investigators in several countries were inspirited by Lodge's lecture to look into possible extensions and applications of his demonstration experiments, including Righi and a British naval officer, H. B. Jackson. The experiments of Jackson were carried out under the "silent-service" tradition of the Royal Navy and possibly to some extent without the full knowledge of his superiors. The details have only recently come to light.[43] Righi's researches bore fruit not only through his subsequent publications and texts, but also because they excited the interest of one of his young compatriots whose name came to be firmly linked with the development of radiotelegraphy: Guglielmo Marconi.

This study was sponsored in part by Grant GS–1428 from the National Science Foundation.

REFERENCES

1. Süsskind, C., *IEEE Spectrum*, vol. 5, pp. 90–98, Aug. 1968; pp. 57–60, Dec. 1968.
2. Threlfall, R., *Rep. Australasian Assoc.*, vol. 2, pp. 27–54, 1890.
3. *Electrician*, vol. 26, pp. 685–686, 1891.
4. Ramsay, J. F., *Proc. IRE*, vol. 46, pp. 405–415, 780, 1958.
5. Crookes, W., *Fortnightly Rev.*, vol. 51, pp. 173–181, 1892.
6. Righi, A., *L'ottica delle oscilazioni elettriche*. Bologna: Nicola Zanichelli, 1897.
7. Sarasin, É., and La Rive, L. de, *Arch. Sci. Phys. Nat.*, vol. 29, pp. 358–393, 441–470, 1893.
8. Garbasso, A. G., and Aschkinass, E., *Ann. Phys.*, ser. 3, vol. 53, pp. 534–541, 1894.
9. Bose, J. C., *Proc. Roy. Soc.*, sec. A, vol. 59, pp. 160–167, 1895; vol. 60, pp. 167–178, 1896; vol. 63, pp. 146–155, 1898.
10. Tesla, N., *J. Franklin Inst.*, vol. 136, pp. 267*ff*, 1893.
11. Lodge, O., *Electrician*, vol. 10, pp. 269–270, 1870.
12. Varley, S. A., *Engineering*, vol. 10, pp. 269–270, 1870.
13. Calzecchi Onesti, T., *Nuovo Cimento*, ser. 3, vol. 16, pp. 58–64, 1884; vol. 17, pp. 38–42, 1885.
14. Tyndall, J., *Proc. Roy. Inst.*, vol. 6, pp. 1–14, 1872.
15. Reynolds, O., *Trans. Roy. Soc.*, vol. 170, pp. 727–845, 1879.
16. Lodge, O., and Clark, J. W., *Nature*, vol. 28, pp. 297–299, 1883.
17. Lodge, O., *J. IEE*, vol. 19, pp. 346–379, 1890.
18. *Ibid.*, pp. 382–410.
19. Lodge, O., *Nature*, vol. 41, pp. 462–463, 1890.
20. Minchin, G. M., *Phil. Mag.*, ser. 5, vol. 37, pp. 90–94, 1894.
21. See, for instance, *Nature*, vol. 50, pp. 133–139, 1894; *Engineering*, vol. 57, p. 751, 1894.
22. Lodge, O., *Electrician*, vol. 33, pp. 153–155, 186–190, 204–205, and *ff*, 1894.
23. *The Work of Hertz and His Successors*. London: "The Electrician" Printing and Publishing Co., 1894. The foregoing was the original title; the second edition appeared in 1898; and the third and fourth, retitled *Signalling Through Space Without Wires*, came out in 1900 and 1908, respectively. In its final form, the book also contained reprints of several other papers and various correspondence.
24. Branly, É., *Compt. Rend.*, vol. 111, pp. 785–787, 1890; vol. 112, pp. 90–93, 1891; and *Lumière Élec.*, vol. 40, pp. 301–309, 506–511, 1891.
25. Branly, É., *Electrician*, vol. 27, pp. 221–222, 448–449, 1891.
26. Lodge, O., *Electrician*, vol. 40, pp. 87–91, 1897.
27. Turner, D., *Rept. Brit. Assoc.*, vol. 62, p. 637, 1892.
28. Croft, W. B., *Proc. Phys. Soc.*, vol. 12, pp. 421–422, 1893.
29. Minchin, G. M., *Phil. Mag.*, ser. 5, vol. 31, pp. 207–238, 1891.
30. Lodge, O., *Phil. Mag.*, ser. 5, vol. 37, pp. 94–95, 1894.
31. For a history of the word *radio* see Hamburger, F., Jr., *Proc. IRE (Editorial)*, vol. 49, p. 1373, 1961; Süsskind, C., *Proc. IRE (Correspondence)*, vol. 50, pp. 326–327, 1962; Espenschied, L., *Ibid.*, pp. 327–328; Drubba, H., *Ibid.*, pp. 1995–1996; Goodenough, E. F., *Ibid.*, p. 2522.
32. Minchin, G. M., *Rept. Brit. Assoc.*, vol. 72, pp. 547–548, 1902.
33. Bose, J. C., *Proc. Roy. Soc.*, ser. A, vol. 65, pp. 166–172, 1900.
34. Tommasina, T., *Compt. Rend.*, vol. 128, pp. 666–667, 1899.
35. *Engineering*, vol. 58, pp. 382–383, 1894.
36. Boltzmann, L., *Ann. Phys.*, ser. 3, vol. 40, pp. 399–400, 1890.
37. Lodge, O., *Nature*, vol. 41, p. 368, 1890; and *Proc. Roy. Soc.*, ser. A, vol. 50, pp. 2–39, 1891.
38. Lodge, O., *Electrician*, vol. 40, pp. 87–91, 1897.
39. Lodge, O., *Wireless Weekly*, vol. 2, pp. 305–307, 310, 337–339, 412–413, 1923.
40. Lodge, O., U.S. Patent 674 846, May 21, 1901.
41. Lodge, O., British Patent 11 575, May 10, 1897. (Also U.S. Patent 609 154.)
42. Thompson, S. P., *Notes on Sir Oliver Lodge's Patent for Wireless Telegraphy* (private printing). London: Waterlow and Sons, 1911.
43. Pocock, R. F., *Proc. IEEE*, vol. 51, pp. 959–960, 1963; and *J. Roy. Navy Scientific Service*, vol. 20, pp. 1–5, 1965.

Charles Süsskinds' biography appeared on page 33 of the August 1968 issue.

How to extend sampling-oscilloscope versatility

A versatile new sampling-oscilloscope system provides solutions to a number of measurement problems, some of which have been causing difficulties for several years

H. Allen Zimmerman *Tektronix, Inc.*

A recently developed sampling-oscilloscope system consists of some 20 different kinds of units—some plug-ins, some self-contained instruments—which can be used together in literally hundreds of functional configurations. However, rather than giving detailed descriptions of these units, this article approaches the subject by discussing five specific applications in which this measurement system offers a particular advantage.

Application example 1: Looking ahead of the triggering point

Problem: *How do you obtain enough "lead time" in an oscilloscope display to observe adequately the entire triggering event as well as events of interest that may occur just prior to the triggering point?*

The most commonly used method for getting lead time has been by means of a signal delay line in the vertical channel of the oscilloscope. The use of a trigger pick-off circuit ahead of the delay line allows the sweep or time base to have started by the time the input signal reaches the vertical deflection plates of the cathode-ray tube.

Although this is a popular and very practical solution for conventional oscilloscopes, several problems arise when we consider bandwidths above 100 or 200 MHz. These problems include the bandwidth and transient-response limitations of the delay line itself, the length of line required for sufficient lead time in the display, and the size, weight, and cost of a suitable line.

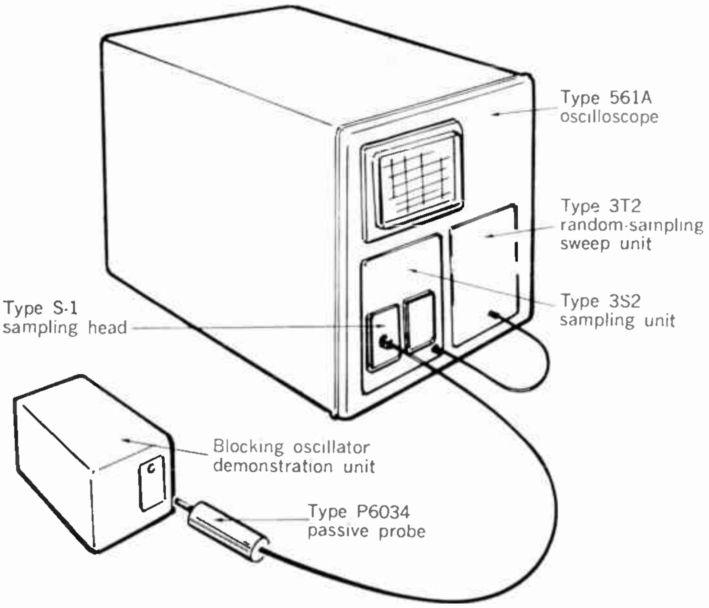


FIGURE 1. Demonstration setup, application example 1.

I. Characteristics of random-sampling oscilloscope

Sweep Rate	Maximum Lead Time	Minimum Signal Repetition Rate
20 to 100 ps/div	50 ns	10 kHz
0.2 to 1 ns/div	500 ns	1 kHz
2 to 10 ns/div	5 μs	100 Hz
20 to 100 ns/div	50 μs	10 Hz
0.2 to 100 μs/div	500 μs	10 Hz

A delaying sweep can be used to provide lead time under certain signal conditions. With this technique, one signal event is used to trigger a delay, which times the oscilloscope sweep to begin just prior to the next signal repetition. The usual limitation of this approach is display jitter unless the signal is extremely stable in its repetition rate. Jitter introduced by the oscilloscope in such a case is typically one part in 10^4 of the delay interval.

A third approach is to use a pretrigger derived from

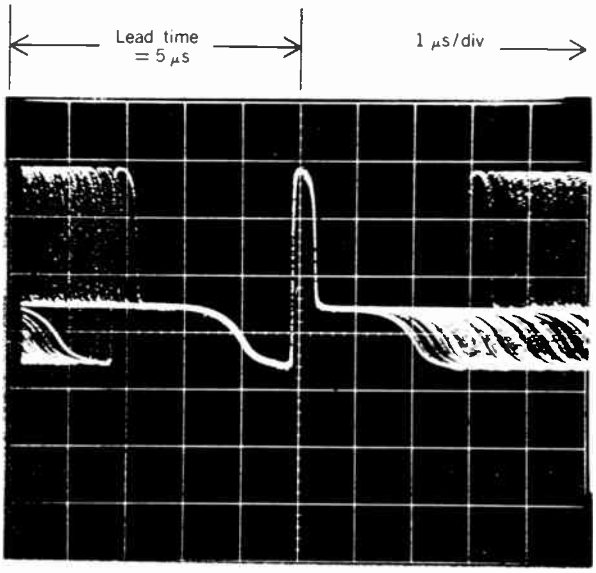


FIGURE 2. Lead time afforded by random sampling.

FIGURE 3. Automated measurement system.

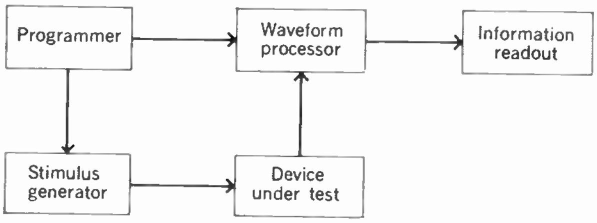
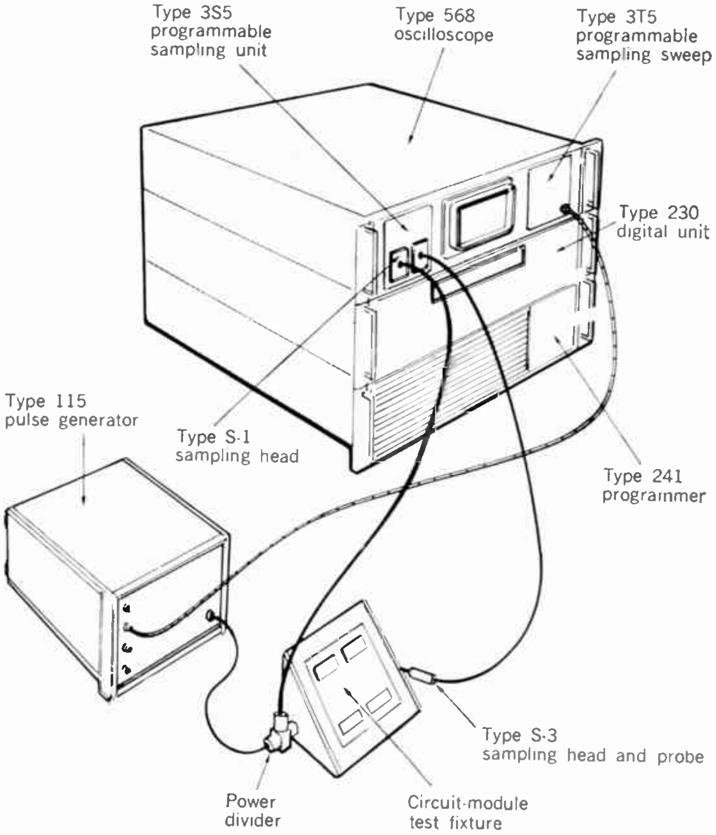


FIGURE 4. Demonstration setup, application example 2.



the signal source itself. The interval by which the pretrigger leads or anticipates the main signal determines lead time in this case. Here, no signal degradation occurs as in the case of the delay line, but many times it may be difficult, inconvenient, or even impossible, to derive a pretrigger that is adequately jitter-free and that sufficiently leads the main signal.

With the recent advent of random-sampling oscillography, an entirely new method for getting lead time was introduced. Neither a signal delay line nor a pretrigger is required in this case. Table I shows how available lead time is related to sweep rate in the oscilloscope display and gives a summary of the minimum signal repetition rates at which the instrument will operate.

The equipment setup chosen to demonstrate this approach is shown in Fig. 1. A simple blocking oscillator with variable repetition rate provides the signal source. A trigger pick-off circuit within the sampling head derives the triggering signal for the random-sampling sweep. The time exposure of Fig. 2 was taken while the blocking oscillator repetition rate was changed. The pulse on which the display is triggered stands still as the repetition rate changes. Note that 5 μs of lead time is indicated for the sweep rate used in the figure: this is far more than would be feasible by a practical delay line—more than 3000 feet (approximately 1000 meters) for 5 μs. In this case, a means for pretriggering is unavailable.

A random-sampling sweep times the sample-taking operation in much the same manner as the delaying sweep. The difference is that the horizontal coordinate of each sample in the display is position-referenced to the *second* repetition itself rather than to the first. The reader is directed to Refs. 1 and 2 for more details on the development and operation of the random-sampling oscilloscope.

Application example 2: Automating test and checkout routines

Problem: *How do you speed up circuit-module or device testing while reducing the probability of operator error?*

"Automation" is the answer, of course. In this case it means programming the test instrumentation through a series of measurements made at specific test points. These may include measurements of pulse amplitude, rise time, duration, or delay, for example. Each measurement is compared with preassigned limit values to determine whether or not it is acceptable. An unacceptable value may either sort or reject a device or call for some action by the operator.

The benefit from such an approach comes directly from the fact that the operator is asked to do less—both in setting instrument controls and in the interpretation of waveforms or data. The degree to which it is desirable to reduce operator intervention may depend on the particular tests being performed. Despite the glamor of mechanized operations, the average human being is still considerably more capable of performing many evaluative functions with greater efficiency than a machine.

The functional requirements of a basic automated measurement system are shown in Fig. 3. The programmer may be as simple as a row of push buttons

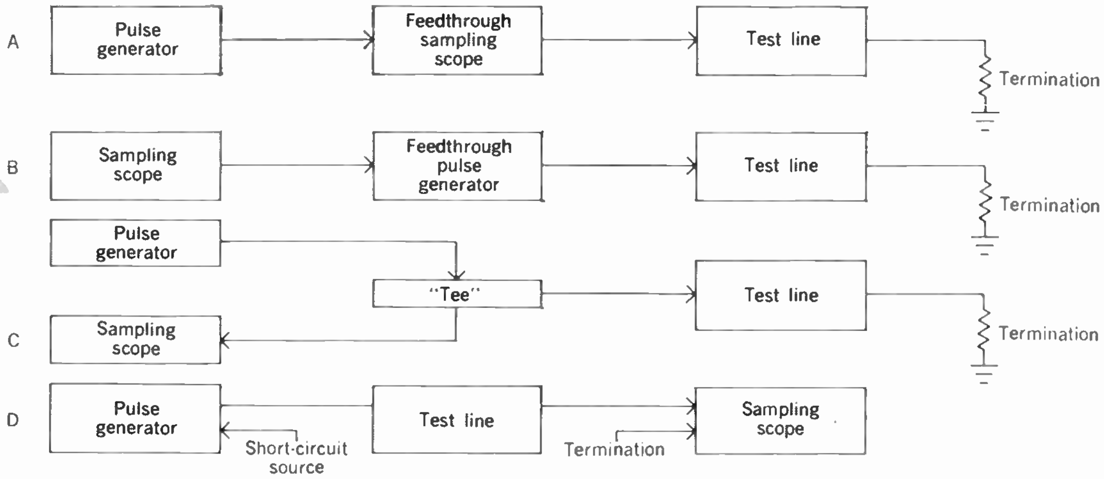


FIGURE 5. TDR system configurations.

or as complicated as a digital computer. The stimulus generator may be a single power supply or a bank of programmable pulse generators. The device under test may be a single resistor or an entire instrument, such as a shipboard radar. The waveform processor may be a simple voltmeter or the front end of a fully programmable digital oscilloscope. The readout may be a cathode-ray tube, a digital indicating system, or a computer interface.

Clearly, the level of sophistication represented by such a measurement system, as well as its functional capability and cost, can vary widely.

Figure 4 shows the equipment setup chosen for the test and checkout procedure of a particular logic module—in this case, a scale-of-ten counter in which four integrated-circuit flip-flops on a plug-in circuit card were employed.

A drive signal is supplied to the test circuit and to one channel of the oscilloscope for triggering and verification purposes. Test points built into the circuit board during its design bring out pulse signals, which are to be measured for amplitude, duration, rise time, and propagation delay with respect to the driving signal in order to verify proper operation.

All of the relevant oscilloscope settings together with the "go–no go" limits for a measurement at a particular test point are selected by a single push button to enable the operator to locate the faulty component or adjustment. Systems of this type are available with widely varying levels of sophistication, all the way from relatively simple measurement-indicating systems to computer-based systems of great power. The system illustrated here is one of the simpler types.

Application example 3: High-resolution time-domain reflectometry

Problem: *How do you achieve maximum resolution in a TDR system?*

First, we need to look at the term "resolution" itself and see what it means in relation to time-domain reflectometry (TDR). Broadly speaking, it means the ability to observe and measure small discontinuities in the characteristic impedance of a transmission system. But TDR resolution really has two dimensions: amplitude and time.

"Amplitude resolution" refers to the ability of a TDR system to display the effect of a discontinuity having a very small reflection coefficient.

"Time resolution," on the other hand, refers to the ability of a TDR system to distinguish between two point discontinuities that are very close together.

In order to give some quantitative significance to these two terms, we may introduce the following definitions:

1. The *amplitude resolution* of a TDR system may be stated as that reflection coefficient which is equivalent to the displayed noise level.

2. The *time resolution* of a TDR system may be stated as the minimum time spacing of two equal point discontinuities that give rise to a 50 percent valley between the two displayed reflections.

Now that we have defined what we mean by "resolution," let us consider some of the contributing factors that limit it.

For a distributed discontinuity—such as a section of 49-ohm transmission line in a 50-ohm TDR system—the amplitude resolution is usually limited by system noise. For a point discontinuity, or one that is distributed over a short distance relative to the equivalent length of the system rise time, the amplitude resolution is usually limited by both system noise and rise time. Unwanted reflections or waveform aberrations that happen to fall on top of the desired reflection can also limit amplitude resolution. It is usually possible to shift these disturbances out of the time region of interest by adding a judiciously chosen length of delay cable in the right place.[3]

Time resolution is limited primarily by system rise time. It must be emphasized that the transmission system or device under test should be considered as part of the overall TDR system and as such may introduce a significant limitation on time resolution. Thus, it may be possible to resolve two closely spaced discontinuities some 3 meters down a coaxial line, whereas the same two discontinuities may appear smeared together at 15 meters.

II. Typical TDR resolution

Example	Type of Configuration (see Fig. 5)	Instrument Types (Tektronix)	Pulse Amplitude	System Rise Time*	ρ, Maximum Amplitude Resolution†	Maximum Time Resolution†
1	A (Feedthrough sampler)	Type 1S2 reflectometer	1.0 V 250 mV	1.1 ns 150 ps	0.0004 0.001	670 ps 85 ps
2	B (Feedthrough pulser)	Type 281 TDR pulser Type 3S1 sampling unit	460 mV	1.0 ns	0.001	600 ps
3	C ("Tee")	Type S-50 pulser Type S-4 sampling head	400 mV	40 ps	0.01	35 ps
4	C ("Tee")	Type 111 pulse generator Type S-1 sampling head	5 V (attenuated to 1.0 V)	610 ps	0.001	400 ps
5‡	D (In-line)	Type S-50 pulser Type S-4 sampling head	400 mV	40 ps	0.004	35 ps

* Includes effects of pulse source, sampler, and configuration.
† Typical performance figures.
‡ Example illustrated.

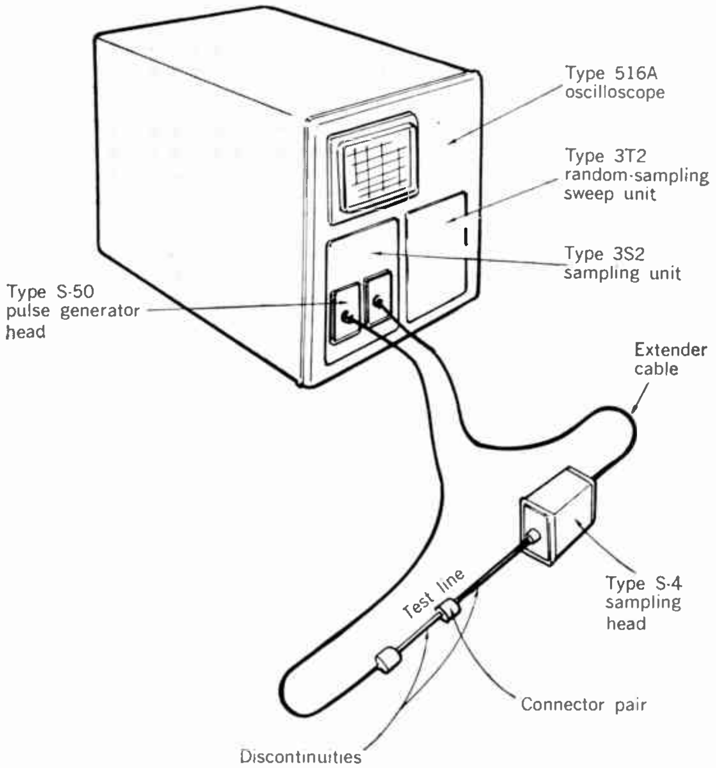
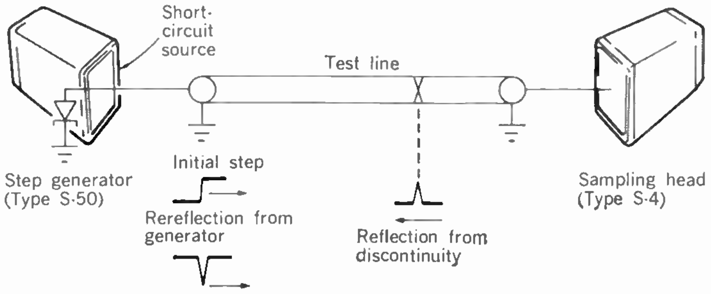


FIGURE 6. Demonstration setup, application example 3.

FIGURE 7. In-line TDR configuration.



From the foregoing discussion, it would appear that one should select the TDR system having the fastest rise time and the least noise. Unfortunately, it is not that easy—partly because these are, to some extent, mutually exclusive characteristics and partly because the *configuration* of the TDR system plays a major role in determining the actual noise and rise time.

Four major TDR system configurations are shov in block-diagram form in Fig. 5. Because of practica instrument considerations, the systems vary widely in their advantages and limitations. It should also be said that the configuration that gives the best amplitude resolution will probably *not* give the best time resolution. Table II compares resolution and configuration for five instrument combinations from one manufacturer.

Example 5 from Table II has been chosen for a typical demonstration. Figures 6 and 7 show the setup. The in-line configuration is particularly well-suited for studying discontinuities in relatively short, high-quality transmission systems.

A tunnel diode supplies the exciting pulse directly into the test line. The pulse propagates down the test line until it encounters a discontinuity, whereupon energy is reflected back toward the generator. The short-circuit (3-ohm) source impedance of the generator then rereflects this energy back through the test line and into the sampler for observation. Since reverse termination in the pulse generator is not required—indeed, not allowed—the full amplitude of the tunnel-diode waveform is available to drive the test line for the best signal-to-noise advantage. And since we have presumed a short, high-quality test line, rise-time losses are held to a minimum.

Figure 8 shows the amplitude resolution of the system. The observed reflection coefficient of 0.004 cc responds to a shunt capacitance of 0.008 picofarad. A similar reflection coefficient of the opposite polarity

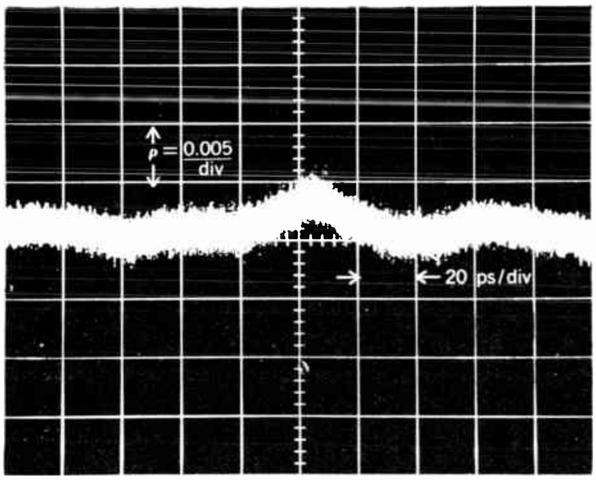
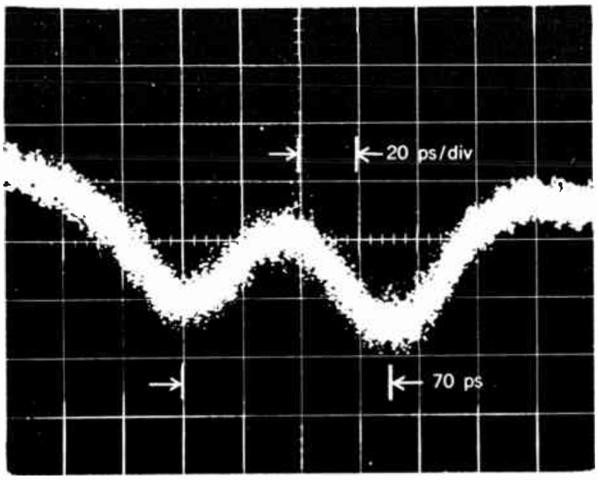


FIGURE 8. Maximum amplitude resolution.

FIGURE 9. Maximum time resolution.



would indicate a series inductance of 20 picohenrys.

Figure 9 illustrates the time resolution of the system. The two discontinuities indicated are separated by 7.4 mm of solid Teflon transmission line. This corresponds to a time separation of 35 picoseconds in the test line which becomes a 70-ps indication in the TDR display due to the round-trip propagation required.[4]

Figure 10 shows a typical reflection from a mated pair of 3-mm connectors from one manufacturer. Obviously, quite a bit of improvement is possible before the TDR resolution limits are reached.

Application example 4: High-speed circuit probing

Problem: *How do you select the best voltage probe for picking signals out of nanosecond circuitry?*

At least three questions need to be asked when any high-speed voltage probe is evaluated:

1. What rise time will the probe allow when being driven by the circuit under test?
2. Does the probe load the circuit under test excessively?
3. Does the probe allow sufficient sensitivity?

The answers to these questions are not found in the probe specifications themselves; they depend on you. That is, they depend on characteristics and requirements of the test circuit and on characteristics of the oscilloscope with which the probe is to be used.

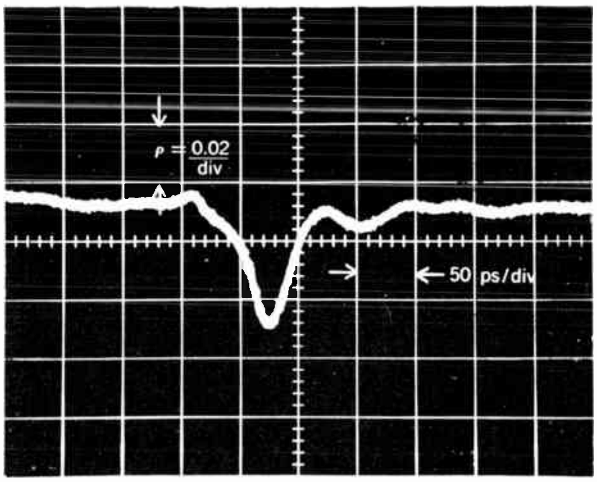


FIGURE 10. Reflection from connector pair.

III. Characteristics of high-speed probes

Probe Type (Tektronix)	Rise Time, ns	Capacitance, pF	DC Resistance	Attenuation
1. Passive (P6034)	0.1	0.7	500 Ω	10×
2. Passive (P6035)	0.2	0.6	5 kΩ	100×
3. Sampling (S-3)	0.35	2.3	100 kΩ	1×
4. Cathode follower (P6032)	0.4	3.6	10 MΩ	10×
5. Field-effect transistor (P6045)	1.5	5.5	10 MΩ	1×

Four important electrical characteristics usually quoted for high-speed voltage probes are:

1. Rise time
2. Capacitance
3. Resistance
4. Attenuation

These factors are heavily interrelated, as shown by Table III for five particular high-speed probes. Although one characteristic can often be traded off for another, in practice a certain *set* of characteristics is needed for a particular measurement. We find that the probe with the fastest rise time, for example, looks like 500 ohms under dc conditions and that the next fastest probe has 100× attenuation.

Probes are either *passive* (being chiefly a resistive attenuator) or *active* (providing electronic means for voltage amplification, isolation, or power gain). In general, passive probes tend to be fast, but either at a sacrifice of sensitivity or at the penalty of rather heavy circuit loading. Active probes, on the other hand, offer less attenuation but tend to be somewhat slower because of the active devices used and the inevitably higher input capacitance. Linearity and signal range are often factors of interest for active probes.

Figure 11 shows waveform photographs taken with

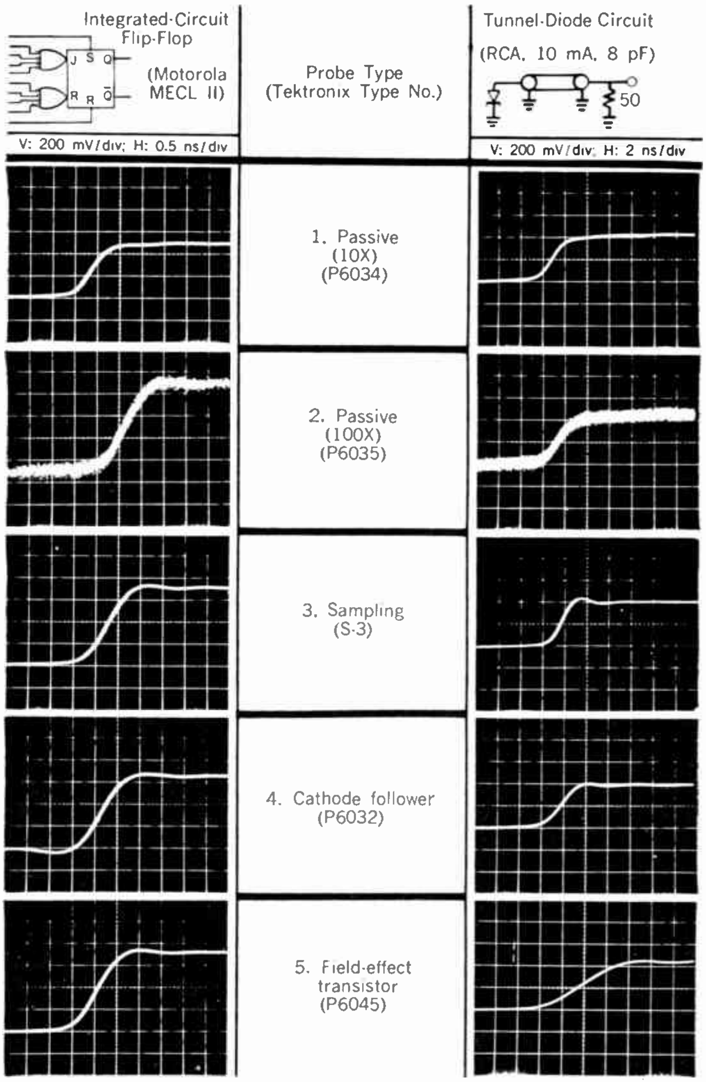


FIGURE 11. Test-circuit waveforms.

the five probes on two different circuit examples—a commercially available integrated-circuit flip-flop and a tunnel diode. These circuits were chosen since they are both commonly encountered signal sources, with switching times in the region of 1 to 10 ns.

The 500-ohm resistance of probe 1 clearly loads the integrated circuit excessively as evidenced by reduced amplitude. The 100× attenuation of probe 2 results in a noisy trace because of the accompanying required increase in sensitivity of the sampler. Probes 3, 4, and 5 all give good results with the integrated circuit but show the effects of slower probe rise time with the tunnel diode.

The advantage of probe 3 over the others rests in its *combination* of characteristics: high sensitivity, fast rise time, and high dc resistance. It is therefore possible to cover a broader range of applications with the single probe.

When using any high-speed probe it is good to keep in mind:

1. That 2 pF of probe capacitance driven from a 500-ohm source produces a 1-ns time constant (or a 10 to 90 percent rise time of 2.2 ns), which may be more significant than the probe rise time itself.
2. That a more sensitive probe can usually be made to accommodate larger signals by simply adding attenuators; however, it is usually impossible to increase the basic sensitivity of a probe to look at smaller signals.
3. That low probe resistance may produce excessive dc loading, whereas the typical dc-blocking capacitor for use with the probe may introduce excessive "droop" or "sag."
4. That probe attenuation degrades the displayed signal-to-noise ratio, since more effective gain will be required from the oscilloscope.

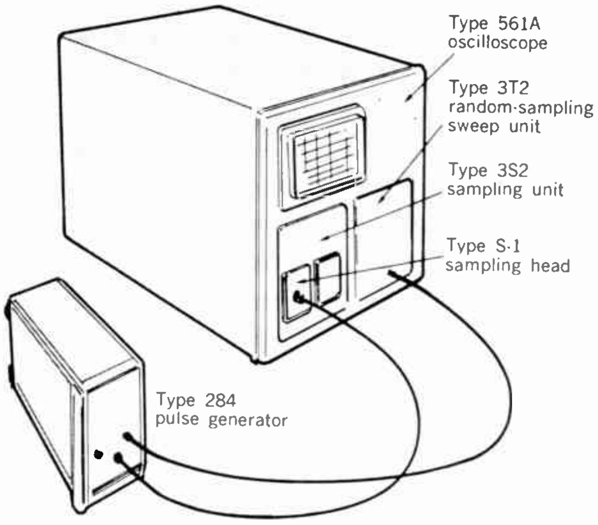


FIGURE 12. Demonstration setup, application example 5.

Application example 5: Toward step-response flatness

Problem: *How do you verify the time-domain response of an oscilloscope?*

It is generally agreed that the characterization of the time-domain response of an oscilloscope should be carried out in the time domain rather than in the frequency domain. Although admittedly better reference standards and techniques exist today in the frequency domain of sine waves, the process of relating this technology to the time domain is so complex and laden with sources of uncertainty that most researchers prefer to establish the best time-domain standards they can and refer their measurements to those standards.

The step function is generally agreed to be the best choice of an excitation signal because of the relative ease with which it can be interpreted and because of its ability to portray directly the effects of both high- and low-speed distortions.

Clearly, other excitation functions such as the impulse, ramp, and square wave have advantages over the step function for certain types of tests. But for purposes of description and specification of oscilloscope behavior in the time domain, the step response has become the generally accepted medium of expression.

The "ideal" step generator, were it available, would have zero rise time and perfect flatness as prime char-

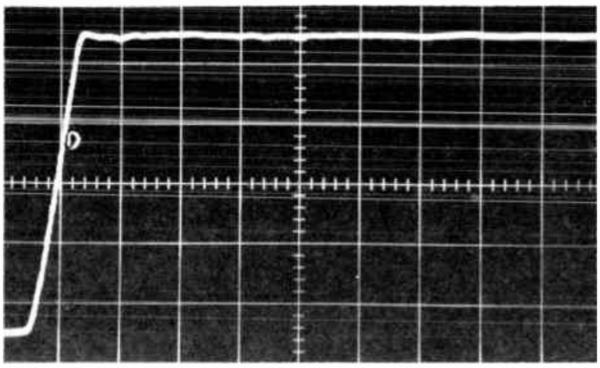
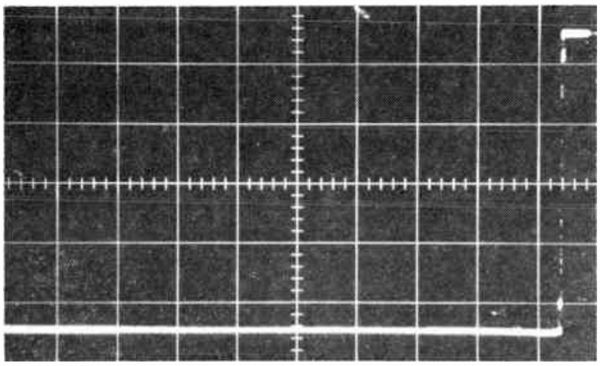


FIGURE 13. Step response. Vertical scale: 100 mV/div; horizontal scale: 0.5 ns/div.

FIGURE 14. Step response. Vertical scale: 100 mV/div; horizontal scale: 50 ns/div.



acteristics. It would also have sufficient amplitude, step duration, and repetition rate for the particular tests being conducted. Whether or not it were a reverse-terminated source might also depend on application, but this characteristic is often high desirable.

Obviously, in a practical step generator we can only approach these ideal characteristics. Unfortunately, zero rise time and perfect flatness are conflicting objectives, since the fastest devices and circuits are the most difficult to observe accurately and to characterize theoretically.

This very problem, of course, is the root of the dilemma: Is the observed aberration caused by the oscilloscope or by the step generator? And even more insidious: If there is no displayed aberration, how can one be sure that the oscilloscope and step generator are not somehow complementary? That is, could a defect in one exactly compensate for the inverse defect in the other?

The equipment chosen for demonstration of state-of-the-art flatness is shown in Fig. 12. The combined step-generator/oscilloscope response is shown in Figs. 13 and 14. In these particular instruments, special attention has been given to the factors affecting the predictability of flat time-domain behavior. System rise time was compromised in favor of flatness and predictability. With care, this combination is capable of displaying a 350-ps rise time, with a total uncertainty of less than ±2 percent.

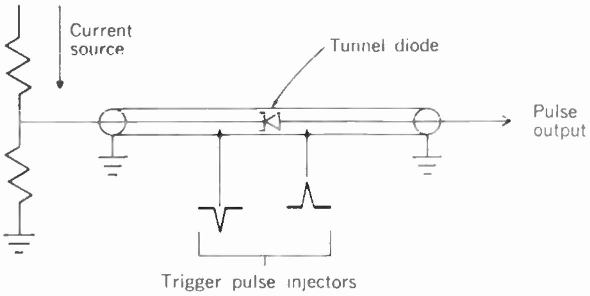


FIGURE 15. Tunnel-diode pulse generator.

The pulse is generated by a tunnel diode mounted in series with the center conductor of a 50-ohm coaxial transmission line.[5] The back end of the line is terminated by a specially designed resistor network, which also serves as a bias supply; see Fig. 15.

Triggering of the tunnel diode is accomplished by a pulse "injected" by two probes very near the tunnel diode itself. Triggering signals at the probes are equal and opposite, so they do not appear in the output.

Time-domain reflection techniques were used extensively during the development stages to ensure that the termination and the tunnel diode electrical environment would not contribute unnecessarily to the pulser's aberrations.

Essentially full text of a paper presented at the 1968 IEEE Northeast Electronic Research and Engineering Meeting (NEREM-68), Boston, Mass., November 6–8.

REFERENCES

1. Frye, G. J., and Nahman, N. S., "Random sampling oscillography," *IEEE Trans. Instrumentation and Measurement*, vol. IM-13, pp. 8–13, Mar. 1964.

2. Zimmerman, H. A., "Random sampling—a new way of fast pulse display," *Service Scope*, Tektronix, Inc., no. 46, Oct. 1967.

3. Strickland, J., and Zimmerman, H. A., "Time domain reflectometry theory and coaxial cable testing," *Broadcast J.*, Mar.–Apr. 1968.

4. Long, G., "Pulse reflections pin down discontinuities," *Electron. Design*, May 10, 1963.

5. Frye, G. J., "A predictable subnanosecond step generator," *IEEE Trans. Instrumentation and Measurement*, vol. IM-17, pp. 324–329, Dec. 1968.

H. Allen Zimmerman (M) is program manager for Sampling and Digital Instruments at Tektronix, Inc., Beaverton, Oreg. He has been involved in the design of sampling oscilloscopes since joining Tektronix in 1961. A native of Kansas, he received the A.B. degree in electrical engineering from Harvard University in 1959 and the M.S. degree, also in electrical engineering, from the University of California, Berkeley, in 1961. Mr. Zimmerman currently holds two U.S. patents covering oscilloscope triggering circuits and has several others pending. At present he is serving on the central executive committee of the IEEE Subcommittee on Pulse Techniques, a group organized for the purpose of updating the terminology and standard definitions associated with pulse measurements.

Operation GIT a success

What is Operation GIT? It is a continuing membership drive (*G*rowth, *I*mprovement, *T*ransfers) launched by the IEEE Membership and Transfers Committee in January 1968 with the objective of promoting membership in the IEEE and achieving a possible ten percent annual net increase in the membership of Sections established throughout the world.

The drive was a success in its first year of operation. Sections in every Region increased membership, and special congratulations are extended to the following 38 Sections, who swelled their memberships by ten percent or more in the calendar year 1968:

Region 1	Mohawk Valley	Vermont
Region 2	Ohio Valley	
Region 3	Central North Carolina Charlotte Gainesville	Miami Palm Beach
Region 4	Central Illinois Central Indiana Madison	Siouxland Southern Minnesota
Region 5	Corpus Christi Fort Worth Houston	Panhandle Permian Basin
Region 6	Alamogordo–Holloman Buenaventura	Santa Barbara Tucson
Region 7	Canadian Atlantic Kitchener–Waterloo Montreal	Northern Alberta Regina
Region 8	Benelux France North Italy	Norway Sweden Switzerland
Region 9	Argentina Chile	Rio de Janeiro São Paulo
Region 10	Tokyo	

Why was Operation GIT a success? The Sections brought to the attention of potential members how enrollment in IEEE would aid them professionally and how they could contribute to the electrical and electronics engineering profession. These candidates for membership were made aware of the many reasons why those in the profession could hardly afford not to join the IEEE and

- attend meetings, local and Institute-wide, to keep up to date on the latest developments in the profession.
- receive IEEE publications, which now constitute over ten percent of the world's literature in the electrical and electronics engineering field.
- participate personally in the rapidly changing technology by becoming involved in the many programs offered by IEEE to help them grow professionally.
- enter into the affairs of the local Sections, where they have the opportunity to meet regularly with engineering associates to learn directly from others and to contribute to these technical discussions themselves.
- join one or more of the 31 Groups established within IEEE in specialty fields related to almost every phase of electrical and electronics engineering: receive publications, attend meetings, and associate with those engaged in their own particular specialty fields.
- read in IEEE SPECTRUM (provided monthly as part of the annual $25.00 membership dues) outstanding technical articles on a limitless variety of subjects, as well as information about the wide diversity of IEEE operations.
- take advantage of subscribing to the monthly PROCEEDINGS OF THE IEEE, for a modest $3.00 per year, to read, and retain for reference, papers of broad significance in the profession.

The Sections have called these facts to the attention of prospective members. How about you as an individual member? You must know at least one engineer or scientist who is not a member of the IEEE. You may have even thought he would join IEEE if he were made aware of the advantages of belonging. The greater participation by engineers and scientists in furthering the objectives of the Institute, the broader the technical exchange. This is the reason IEEE is always concerned with methods of extending the membership, even though we are already the largest engineering society in the world.

You can help yourself, your Institute, and the potential members by jotting down the names and addresses of these prospective members on the prepaid postcard attached to the reader service card in this publication.

IEEE is waiting to hear from you.

Murrell F. Jessen, Chairman
Membership and Transfers Committee

Convolution revisited

That probability theory can serve as an invaluable tool in solving engineering problems is amply demonstrated by the use of convolution techniques providing solutions previously attainable by transform methods

Timothy J. Healy *University of Santa Clara*

Few mathematical operations are more important to the engineer than convolution and transform analysis. In this article, the operation of convolution is explored—starting with discrete rather than continuous convolution because of the relative ease of comprehension involved. With this foundation, the study is extended to continuous convolution. A proof of the convolution theorem will show that convolution and transform analysis are closely related. Of much more interest, however, is an intuitive explanation of why convolution and transform analysis techniques lead to exactly the same solution of a given problem. Perhaps the two most important applications of convolution deal with the analysis of linear systems and the sums of independent random variables—the latter problem being used to introduce discrete convolution.

Adding random variables

Transformers of a certain type are delivered by two companies, which will be called A and B, in lots of three and four, respectively. In a lot from company A, there will be zero, one, two, or three defective transformers. The probability that each of these numbers of defects, represented by a ($a = 0, 1, 2, 3$), will occur is given as

a	$P(a)$
0	0.4
1	0.3
2	0.2
3	0.1

Similarly, the probability of b ($b = 0, 1, 2, 3, 4$) occurrences is

b	$P(b)$
0	0.3
1	0.2
2	0.2
3	0.2
4	0.1

in a shipment from company B. The problem is to find the probabilities of the total number of defects in two shipments, one from each company. This sum or total number of defects will be indicated as c ($c = 0, 1, 2, 3, 4, 5, 6, 7$).

The probability that $c = 0$ is the probability that $a = 0$ and $b = 0$; that is, that there are no defects in either shipment. This is written as

$$P(c = 0) = P(a = 0 \text{ and } b = 0)$$

If the events $a = 0$ and $b = 0$ are independent of each other, which we assume here and which is necessary if the solution is to be a convolution, then this probability reduces to the product

$$\begin{aligned} P(c = 0) &= P(a = 0) \times P(b = 0) \\ &= 0.4 \times 0.3 \\ &= 0.12 \end{aligned}$$

The probability that we have a total of one defect is the probability that the shipment from A has one defect and the shipment from B none, or vice versa. That is,

$$P(c = 1) = P[(a = 1 \text{ and } b = 0) \text{ or } (a = 0 \text{ and } b = 1)]$$

Using an axiom of probability theory, the probability of the "or" statement within the brackets is changed to the sum of two probabilities.

$$\begin{aligned} P(c = 1) &= P(a = 1 \text{ and } b = 0) + P(a = 0 \text{ and } b = 1) \\ &= P(a = 1)P(b = 0) + P(a = 0)P(b = 1) \\ &= 0.3 \times 0.3 + 0.4 \times 0.2 \\ &= 0.17 \end{aligned}$$

Similarly,

$$P(c = 2) = P[(a = 2 \text{ and } b = 0) \text{ or } (a = 1 \text{ and } b = 1) \text{ or } (a = 0 \text{ and } b = 2)]$$

$$\begin{aligned} &= P(a=2)\,P(b=0) \\ &\qquad + P(a=1)\,P(b=1) \\ &\qquad + P(a=0)\,P(b=2) \\ &= 0.2 \times 0.3 + 0.3 \times 0.2 + 0.4 \times 0.2 \\ &= 0.20 \end{aligned}$$

Continuing in this way, one can obtain the probabilities for all eight possible values of c. The result is given as

c	$P(c)$
0	0.12
1	0.17
2	0.20
3	0.21
4	0.16
5	0.09
6	0.04
7	0.01

Discrete convolution

Although it is possible to solve this problem in the manner just described, it is highly desirable to find a shortcut to determine the probabilities of the values of c. Notice how the entries for a and b are used to find those for c. The first entry for c is the product of the first entries for a and b. The second entry for c is entry 1 of b times entry 2 of a plus entry 2 of a times entry 1 of b. The third entry of c is the sum of cross terms 1 and 3, 2 and 2, 3 and 1 of a and b.

We can systematize the process of finding the entries for c in the following way. Write the probabilities for a and b as sequences A and B:

$$A = [0.4\ \ 0.3\ \ 0.2\ \ 0.1]$$

$$B = [0.3\ \ 0.2\ \ 0.2\ \ 0.2\ \ 0.1]$$

Reverse the order of one of the sequences, say B. Call the reversed sequence B_{inv}

$$B_{\text{inv}} = [0.1\ \ 0.2\ \ 0.2\ \ 0.2\ \ 0.3]$$

Position sequence A and the inverted sequence B_{inv} so that the first right-hand term of B_{inv} is under the first left-hand term of A:

$$\begin{array}{ccccccccc} &&&& 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.2 & 0.2 & 0.3 &&&& \end{array}$$

The probability of 0 defects is the product of the overlapping numbers 0.4 and 0.3. Now shift the inverted sequence *one* position to the right:

$$\begin{array}{cccccccc} &&& 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0.2 & 0.2 & 0.3 &&& \end{array}$$

The probability of *one* defect is the sum of the overlapping products 0.2 × 0.4 and 0.3 × 0.3. The remaining terms in c are obtained by shifting the inverted sequence one step at a time to the right, and for each step summing the overlap products.

The process of inverting a sequence, sliding it one step at a time to the right, and summing the overlap products is called discrete convolution. It is sometimes called serial multiplication. (An excellent discussion of serial multiplication or discrete convolution is given by Bracewell.[1])

The asterisk (*) is generally used to indicate discrete as well as continuous convolution. Thus, we write

$$C = A * B \tag{1}$$

$$= [0.4\ \ 0.3\ \ 0.2\ \ 0.1] * [0.3\ \ 0.2\ \ 0.2\ \ 0.2\ \ 0.1]$$

$$= [0.12\ \ 0.17\ \ 0.20\ \ 0.21\ \ 0.16\ \ 0.09\ \ 0.04\ \ 0.01]$$

Before proceeding with the theoretical development, the reader who is not familiar with discrete convolution should spend some time practicing the technique. The objective of this practice is to get a feeling

This box contains a number of exercises intended to acquaint the reader with both the techniques and results of discrete convolution. The reader is encouraged to look for interesting forms or combinations and perhaps deduce some of the properties of convolution that have not been discussed here.

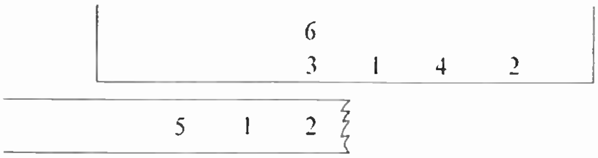
(1) [3 1 5] * [4 4 7] = [12 16 45 27 35]
(2) [1 4 4 1] * [2 6 2] = [2 14 34 34 14 2]
(3) [1 3 3 1] * [1 2 1] = [1 5 10 10 5 1]
(4) [9 5 1] * [5 3 1] = [45 52 29 8 1]
(5) [4 1 7 8] * [···0 0 1 0 0 0···] = [···0 0 4 1 7 8 0 0 0···]
(6) [1 1 1 1] * [1 1 1 1] = [1 2 3 4 3 2 1]
(7) [8 1 6 3 2] * [4 1 5] = [32 12 65 23 41 17 10]

for the operation. This will facilitate understanding of further developments.

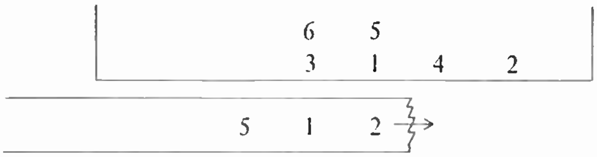
There are a number of ways one may carry out the mechanics of the operation. It is well to start with a procedure that is a little tedious but very instructive. Consider the convolution of two sequences, A and B, defined as

$$A = [3\ 1\ 4\ 2] \qquad B = [2\ 1\ 5]$$

Write the sequence A on one piece of paper. On the edge of a second piece of paper, write the inverted sequence B_{inv} with the same spacing as A.



Slide the B_{inv} sequence paper to the right until the 2 is under the 3. Obtain the product 6 and write it above the 3. Shift the B_{inv} sequence paper one step to the right:



Add the paired products 2 × 1 and 1 × 3 to obtain 5: write 5 above the 1. Repeat this operation until the paired products are finished; one should obtain

$$[3\ 1\ 4\ 2] * [2\ 1\ 5] = [6\ 5\ 24\ 13\ 22\ 10]$$

A number of examples are given in the tinted box on page 88. The reader should work one or two with the method just outlined; once he has a feeling for what is happening he will want to find a simpler method. The following approach seems to be as simple as possible. Write out the sequences in a conventional multiplication format. Do not invert either sequence. Use the conventional multiplication procedure with the exception of carrying tens—do not carry tens.

$$\begin{array}{rrrrrr}
 & & 3 & 1 & 4 & 2 \\
 & & & 2 & 1 & 5 \\
\cline{3-6}
 & & 15 & 5 & 20 & 10 \\
 & 3 & 1 & 4 & 2 & \\
6 & 2 & 8 & 4 & & \\
\hline
6 & 5 & 24 & 13 & 22 & 10
\end{array}$$

If one uses this technique to obtain some of the results in numerical examples that are given, it will be easier to understand, and to check, the following properties or facts. No rigorous proofs will be given here.

Consider the convolution of two sequences A and B to obtain a third sequence C.

$$A = [a_0\ a_1 \cdots a_i \cdots]$$

$$B = [b_0\ b_1 \cdots b_i \cdots]$$

$$C = [c_0\ c_1 \cdots c_i \cdots]$$

1. The elements of sequence C can be expressed as

$$c_i = \sum_{j=0}^{i} a_j b_{i-j} \tag{2}$$

2. If the sequence A has n_A terms and the sequence B has n_B terms, the number of terms in sequence C is

$$n_C = n_A + n_B - 1 \tag{3}$$

3. Convolution is commutative:

$$A * B = B * A \tag{4}$$

This means that either sequence can be inverted and shifted.

4. The product of the sum of the elements in sequence A and the sum of the elements in sequence B is equal to the sum of the elements in sequence C. That is,

$$\left(\sum_{i=0}^{n_A-1} a_i\right)\left(\sum_{j=0}^{n_B-1} b_j\right) = \sum_{k=0}^{n_A+n_B-1} c_k \tag{5}$$

This property should always be used as a check on simple discrete convolution problems. It is a particularly significant property in the sums of random variables problem since the three sequences must each add up to one. This is true because the probabilities in an experiment always total one.

We will see the continuous convolution analogs to these four properties in the next section.

Before moving on to a discussion of continuous convolution, the problem of inverting the discrete convolution process will be described. Consider the following numerical example:

$$[3\ 1\ 4\ 2] * [b_0\ b_1\ b_2] = [6\ 5\ 24\ 13\ 22\ 10]$$

In the inversion problem, the task is to find the sequence B. An algorithm for generating the elements of B is easily established. Consider the following:

$$6\ 5\ 24\ 13\ 22\ 10$$

$$3\ 1\ 4\ 2$$

$$b_2\ b_1\ b_0$$

Since $3 \times b_0 = 6$, b_0 is obviously 2. Now shift B to the right one step. We find b_1 from the expression

$$3 \times b_1 + 1 \times b_0 = 5$$

But b_0 is known, so we can write

$$b_1 = \frac{5 - 1 \times b_0}{3} = 1$$

The general expression for b_i is

$$b_i = \frac{c_i - \sum_{j=0}^{i-1} b_j a_{i-j}}{a_0} \tag{6}$$

Although this expression is quite simple and straightforward in principle, there are serious problems concerned with applying it. If the measured values of C are noisy, that is, if their measured values do not equal their true values, then the values of b_i obtained from Eq. (6) will be increasingly inaccurate as i increases. This is true because the errors accumulate in the summation of $b_j a_{i-j}$ in the equation. An interesting example of this accumulation of error is given by Landgrebe.[2]

At this point, our discussion of discrete convolution will end in order to review the more familiar concepts of continuous convolution. We will return to discrete convolution after borrowing an idea or two from continuous convolution.

Continuous convolution

This section will be essentially a review for many readers. However, it would be well for anyone who does not have a clear understanding of the concept of continuous convolution to follow the ideas closely. He should particularly try to relate each concept in continuous convolution to its discrete-convolution analog.

The convolution of two continuous functions $f(x)$ and $g(x)$ is written as

$$h(x) = \int_{-\infty}^{\infty} f(\lambda)g(x - \lambda)\,d\lambda \qquad (7)$$

(The reader may be used to seeing limits of 0 to x on the integral. We shall see later that this is true only for functions that are zero for x less than zero.) Equation (7) is analogous to Eq. (2). Note that the dummy variable of integration λ in (7) plays a role similar to the dummy variable of summation j in (2). If $f(x)$ and $g(x)$ are probability density functions of two continuous random variables, then $h(x)$ is the probability density function of their sum.

If the variable substitution $x - \lambda = \beta$ in Eq. (7), we obtain

$$h(x) = \int_{-\infty}^{\infty} f(x - \beta)g(\beta)\,d\beta \qquad (8)$$

which tells us that

$$f(x) * g(x) = g(x) * f(x) \qquad (9)$$

This result is of course analogous to Eq. (4).

Property 4 of the previous section also has an analogy in continuous convolution—the product of the area under the $f(x)$ and $g(x)$ curves equals the area under the $h(x)$ curve. This result is obtained by integrating both sides of Eq. (7) over all x:

$$\begin{aligned} \int_{-\infty}^{\infty} h(x)\,dx &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\lambda)g(x - \lambda)\,d\lambda\,dx \\ &= \int_{\lambda=-\infty}^{\infty} f(\lambda)\int_{-\infty}^{\infty} g(x - \lambda)\,dx\,d\lambda \\ &= \int_{-\infty}^{\infty} f(x)\,dx \int_{-\infty}^{\infty} g(x)\,dx \qquad (10) \end{aligned}$$

The last step is possible because the integral over all x of $g(x - \lambda)$ is equal to the integral of $g(x)$.

We turn now to a graphical interpretation of the process of convolution. Consider the functions $f(x)$ and $g(x)$ in Fig. 1.

The minus sign in the term $g(x - \lambda)$ in Eq. (7) represents a reversal of the order of the values of $g(\lambda)$; that is, $g(-\lambda)$ is the mirror image of $g(\lambda)$ with respect to the g axis. This represents a folding or convolving of $g(x)$. The x in $g(x - \lambda)$ represents shifting of the folded function. This process is illustrated in Fig. 2. For this case, where $g(x)$ has zero value for negative x, x is easily identified as the position of the forward edge of the function $g(x - \lambda)$ as it is shifted to the right.

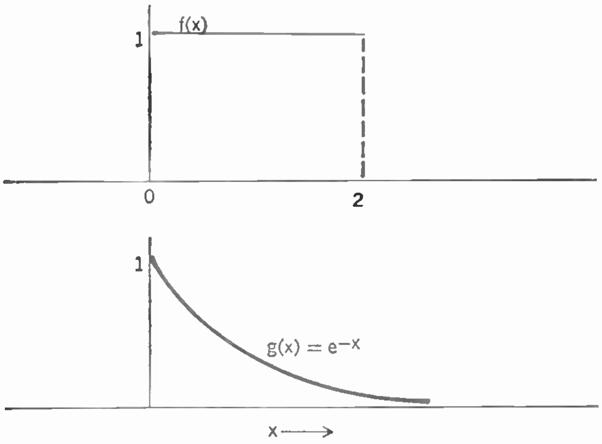
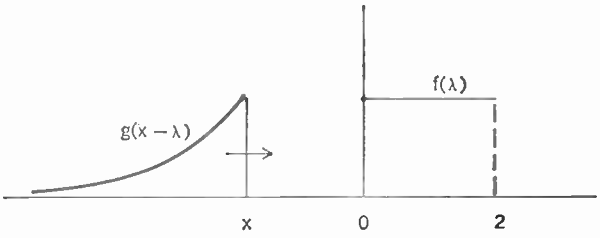


FIGURE 1. Graphical representation of the functions described by f(x) and g(x).

FIGURE 2. Description of the process of convolution for continuous functions.



The graphical interpretation of continuous convolution can now be summarized. The reader should compare this with the discrete convolution operation as previously described. Reverse the order of (or fold) one function and, starting from the far left, shift it to the right one step at a time. (The only difference with continuous convolution is that there are an infinite number of infinitesimal steps. Each position of the shift represents a value of the continous variable x.) For each position x, find the integral of the product of the functions $f(\lambda)$ and $g(x - \lambda)$ for all λ. This last step is of course analogous to finding the sum of the paired terms in discrete convolution.

An integral closely related to the convolution integral is the correlation integral:

$$\int_{-\infty}^{\infty} f(\lambda)g(x + \lambda)\,d\lambda$$

This integral is used in describing signals where $f(\cdot)$ and $g(\cdot)$ are the same function, and describing the interrelation of signals when they are different. It also gives the probability density function of the *differenc* of two continuous independent random variables.

One of the most challenging problems relating to

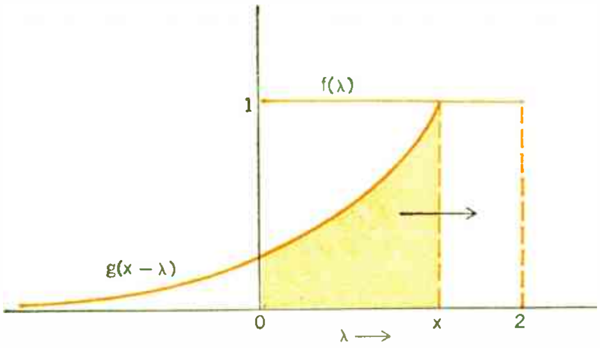
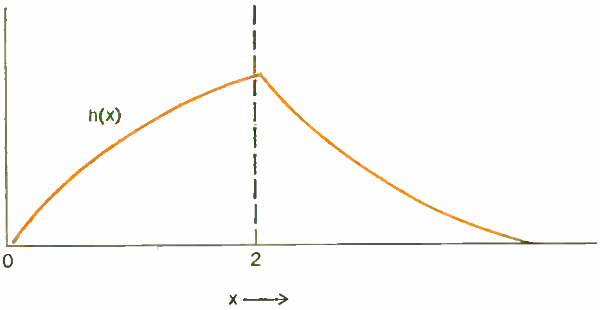


FIGURE 3. Determination of the limits of integration for the convolution integral.

FIGURE 4. Exponential impulse response to a square wave of a circuit.



application of the convolution integral is that of setting the limits on the integral. The $-\infty$ to $+\infty$ limits of Eqs. (7) and (8) apply only if the functions $f(x)$ and $g(x)$ are nonzero for all x. If we wish to integrate a function such as

$$\begin{array}{ll} g(x) = e^{-x} & 0 \leq x \leq \infty \\ 0 & -\infty \leq x \overset{<}{\cancel{>}} 0 \end{array}$$

over all x, we do so by restricting the limits from 0 to ∞; hence, a similar technique is used in the convolution problem.

Consider (Fig. 3) the convolution of the functions shown in Fig. 1, for x between 0 and 2. The area under the product of $f(\lambda)$ and $g(x - \lambda)$ is shown in color. This is the area whose value is found by Eq. (7). It is apparent in this case that the integration must be carried out from 0 to x, rather than from $-\infty$ to ∞.

We would like to have a relatively simple means of finding the limits in the general case. For the example given, the lower limit of the function $g(x - \lambda)$ is $-\infty$ and the lower limit on $f(\lambda)$ is 0. When we integrated, we chose the largest of the two as our lower limit. The upper limit on $g(x - \lambda)$ is x; the upper limit of $f(\lambda)$ is 2. We, in turn, chose the smallest of these for our upper limit of integration. Therefore, the general rule is: Given two functions with lower limits m_1 and m_2 and upper limits M_1 and M_2, use $\max(m_1, m_2)$ to $\min(M_1, M_2)$ as the range of integration.

This rule is of course quite logical. It simply specifies the range of overlap of the two functions. The next task is to find the limits m and M for the two functions. The limits for the fixed function $f(\lambda)$ do not change. They are simply the limits on the original function $f(x)$; that is,

$$m_x \leq \lambda \leq M_x \qquad (11)$$

The limits of the sliding function $g(x - \lambda)$ change as x changes. If the original argument of $g(\cdot)$ is called y, and we are given the limits on y as m_y and M_y, then

$$m_y \leq y \leq M_y$$

But $y = x - \lambda$ under the convolution transformation, so

$$m_y \leq x - \lambda \leq M_y$$

$$\therefore \quad x - M_y \leq \lambda \leq x - m_y \qquad (12)$$

Since the limits on the fixed function are given by (11), and the limits on the sliding function by (12), the integration or overlap range is

$$\max(m_x, x - M_y) \leq \lambda \leq \min(M_x, x - m_y) \quad (13)$$

It is apparent that the range of integration depends on x, the variable of the resultant of the convolution operation. Using (13) to establish the limits for the convolutions of the functions shown in Fig. 1, one has

$$h(x) = \int_{\max(0,x-\infty)}^{\min(2,x-0)} 1 \times e^{-(x-\lambda)}\, d\lambda \qquad (14)$$

Note that the upper limit in (14) is different for $0 \leq x \leq 2$ and $x > 2$. To see why this is reasonable, refer to Fig. 3 and visualize the change in the problem when the sliding function slides past $x = 2$.

Thus, the integral of Eq. (14) has different limits for different ranges of x:

$$\begin{aligned} h(x) &= \begin{cases} \int_0^x e^{-x} e^{\lambda}\, d\lambda & 0 \leq x \leq 2 \\ \int_0^2 e^{-x} e^{\lambda}\, d\lambda & x > 2 \end{cases} \\ &= \begin{cases} 1 - e^{-x} & 0 \leq x \leq 2 \\ e^{-x}(e^2 - 1) & x > 2 \end{cases} \end{aligned} \qquad (15)$$

The result, plotted in Fig. 4, is the response to a square wave of a circuit characterized by an exponential impulse response.

As a second example, consider the general problem of causal time signals applied to linear time-invariant networks, in which a signal has zero value for $t < 0$. In general, the range of both is zero to infinity. Substitution into Eq. (13) yields

$$0 \leq \lambda \leq t \qquad (16)$$

The variable t is the time at which the network output is observed and it is, in general, a function of the input applied over the entire time period from zero to t. The output at the instant t is the *effect;* the input signal applied after $t = 0$ is the *cause.* This approach to establishing the limits of integration is,

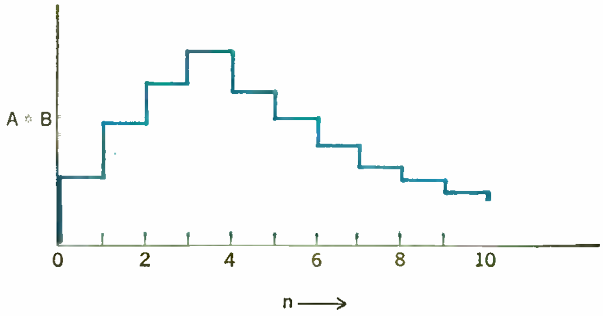


FIGURE 5. Graphical representation of the discrete convolution A * B.

in the author's opinion, the most satisfactory in the long run. It implies and demands an understanding of the problem.

An alternative method, using step functions and a table of special cases, has been proposed by Ross.[3] In order to establish the concept of convolution more clearly, it is instructive at this point to solve a discrete convolution problem that is the analog of the continuous problem that led to Eq. (15). As a parallel to the function $f(\lambda)$, consider a sequence of four identical elements:

$$A = [1\ 1\ 1\ 1]$$

The discrete analogy to a decaying exponential is a sequence in which each element is a constant fraction of the preceding element. That is,

$$b_i = kb_{i-1} \qquad 0 \leq k \leq 1$$

If we let $b_0 = 1$ and $k = 0.8$, we obtain

$$B = [1.00\ 0.80\ 0.64\ 0.51\ 0.41\ 0.33\ 0.26\ 0.21\ 0.17\ 0.13\ \cdots]$$

(Note that in this analogy no attempt has been made to make the continuous and discrete problems numerically analogous. Only the shapes are analogous.) Hence,

$$A * B = [1.00\ 1.80\ 2.44\ 2.95\ 2.36\ 1.89\ 1.51\ 1.21\ 0.98\ 0.77\ \cdots]$$

A bar graph of A * B is shown in Fig. 5 (compare Fig. 5 with Fig. 4).

More on discrete convolution

In discrete convolution, a problem is encountered in establishing the limits of summation, which is similar to the problem of establishing the limits of integration in continuous convolution. The independent variable or index of $f(k)$ is of course k and is assumed to have a range:

$$r_k \leq k \leq R_k$$

The index of $g(l)$ has a range:

$$r_l \leq l \leq R_l$$

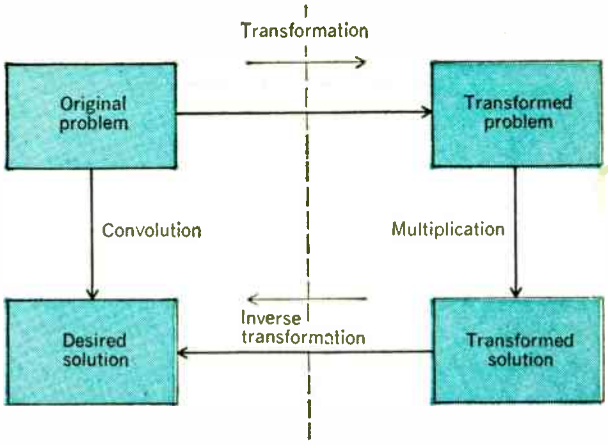


FIGURE 6. Block diagram of the alternate paths of solution that are offered by the techniques involved in convolution and transformation.

In a manner similar to that of the previous section, we obtain for the limits of summation:

$$\max (r_k, n - R_l) \leq k \leq \min (R_k, n - r_l) \tag{17}$$

The two forms of convolution can now be written for comparison:

$$h(x) = \int_{\max(m_x, x-M_y)}^{\min(M_x, x-m_y)} f(\lambda)g(x - \lambda)\, d\lambda \tag{18}$$

$$h(n) = \sum_{k=\max(r_k, n-R_l)}^{\min(R_k, n-r_l)} f(k)g(n - k) \tag{19}$$

If our discrete functions are "causal" ($k = 0, 1, 2, \ldots$ and $l = 0, 1, 2, \ldots$), then

$$h(n) = \sum_{k=0}^{n} f(k)g(n - k) \tag{20}$$

We proceed now to the subject of transform analysis and its relation to convolution.

Transform analysis

The types of problems that can be solved using convolution techniques can also be solved using various transform techniques (e.g., Laplace, Fourier, Z, etc.). Figure 6 shows a simple block diagram of alternate solution paths for a given problem. This diagram applies to the convolution problem, and is applicable to the transform solution of problems in differential equations. In probability theory, the transform—or a very close relative of it—is called the characteristic function.

There are two basic reasons for using transform techniques. The first is to simplify the mathematics required to reach the desired solution. The second is to obtain an insight or understanding of a problem that is unavailable without transform techniques. Both of these reasons apply in this study.

Let us start by proving that we obtain the same solution by following either the transform or convolution path in Fig. 6. This involves the proof of the convolution theorem. We will prove the theorem for the Fourier transform; a similar proof holds for other transforms. The Fourier transform of the function $f(x)$ is

$$F(s) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi xs}\,dx \tag{21}$$

Using this definition, we take the transform of both sides of Eq. (7).

$$\begin{aligned} H(s) &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(\lambda)g(x-\lambda)\,d\lambda\right]e^{-i2\pi xs}\,dx \\ &= \int_{-\infty}^{\infty} f(\lambda)\left[\int_{-\infty}^{\infty} g(x-\lambda)e^{-i2\pi xs}\,dx\right]d\lambda \end{aligned}$$

Letting $x - \lambda = y$.

$$\begin{aligned} H(s) &= \int_{-\infty}^{\infty} f(\lambda)e^{-i2\pi\lambda s}\left[\int_{-\infty}^{\infty} g(y)e^{-i2\pi ys}\,dy\right]d\lambda \\ &= \int_{-\infty}^{\infty} f(\lambda)e^{-i2\pi\lambda s}[G(s)]\,d\lambda \\ &= F(s)\cdot G(s) \end{aligned} \tag{22}$$

This result establishes that convolution in the plane of the original variable is equivalent to multiplication in the plane of the transform variable.

If the variable has discrete values, the Fourier transform is

$$F(s) = \sum_{\text{all } k} f(k)e^{-i2\pi ks} \tag{23}$$

We now repeat the proof of the convolution theorem for discrete variables.

$$\begin{aligned} H(s) &= \sum_{n}\Big[\sum_{\lambda} f(\lambda)g(n-\lambda)\Big]e^{-i2\pi ns} \\ &= \sum_{\lambda} f(\lambda)\Big[\sum_{n} g(n-\lambda)e^{-i2\pi ns}\Big] \end{aligned}$$

Letting $n - \lambda = m$

$$\begin{aligned} H(s) &= \sum_{\lambda} f(\lambda)e^{-i2\pi\lambda s}\Big[\sum_{m} g(m)e^{-i2\pi ms}\Big] \\ &= \sum_{\lambda} f(\lambda)e^{-i2\pi\lambda s}[G(s)] \\ &= F(s)\cdot G(s) \end{aligned} \tag{24}$$

These two parallel proofs are quite important, since they establish the validity of the use of transforms to solve convolution problems. They are not, however, particularly satisfying since they tend not to answer the question of *why* convolution and transform techniques lead to the same solution. We shall try to provide the reader with an answer to that question in the following development.

Discrete functions will be employed since the point to be made is easily seen in that case. Assume that $f(k)$ and $g(k)$ are discrete functions that can be written in the form of sequences, as

$$f(k) = [f_0\ f_1\ f_2 \cdots] \qquad k = 0, 1, 2, \cdots$$

$$g(k) = [g_0\ g_1\ g_2 \cdots] \qquad k = 0, 1, 2, \cdots$$

From Eq. (23), the transforms of $f(k)$ and $g(k)$ are

$$F(s) = f_0e^{-i2\pi 0s} + f_1e^{-i2\pi 1s} + f_2e^{-i2\pi 2s} + \cdots$$

$$G(s) = g_0e^{-i2\pi 0s} + g_1e^{-i2\pi 1s} + g_2e^{-i2\pi 2s} + \cdots$$

To find $f(k) * g(k)$, we can use the convolution theorem [Eq. (24)] just established. Multiply $F(s)$ and $G(s)$, and collect terms with the same exponent:

$$\begin{aligned} H(s) = {}& f_0g_0e^{-i2\pi 0s} + (f_0g_1 + f_1g_0)e^{-i2\pi 1s} \\ & + (f_0g_2 + f_1g_1 + f_2g_0)e^{-i2\pi 2s} + \cdots \end{aligned} \tag{25}$$

The convolution of $f(k)$ and $g(k)$ is obtained by taking the inverse transform of $H(s)$; that is, by finding the function $h(k)$ that when transformed yields $H(s)$. Inspection of Eqs. (23) and (25) suggests that this function is

$$h(k) = [(f_0g_0)\ (f_0g_1 + f_1g_0)\ (f_0g_2 + f_1g_1 + f_2g_0) \cdots]$$

We obtain this result by noting the power of the exponent with which each term is associated. What is happening here should now be apparent. The transform operation is simply performing a bookkeeping function. The transform process associates the correct coefficients with the correct power terms. Multiplication of the transforms leads to the same association of coefficients as does convolution. Thus, convolution and the transform process are seen as two essentially parallel methods of keeping track of sets of coefficients.

This argument is essentially applicable to the continuous case, though it is by no means as easily demonstrated.

REFERENCES

1. Bracewell, R., *The Fourier Transform and Its Applications*. New York: McGraw-Hill, 1965.

2. Landgrebe, D., "Undergraduate experiment in data handling and linear systems," *IEEE Trans. Education*, vol. E-10, pp. 131–138, Sept. 1967.

3. Ross, G., "Choosing the limits on the convolution integral," *IEEE Trans. Education*, vol. E-10, pp. 45–48, Mar. 1967.

Timothy J. Healy (M) was born in Bellingham, Wash., and received the B.S.E.E. degree from Seattle University in 1958, the M.S.E.E. degree from Stanford University in 1959, and the Ph.D. degree in electrical engineering from the University of Colorado in 1966. Employed by the Sperry Gyroscope Company for one year, he was responsible for developing electronic support equipment for an electrostatic gyroscope. Later work for the Boeing Company included a Minuteman missile communication system. From 1960 to 1964, he was an assistant professor of electrical engineering at Seattle University. Dr. Healy is presently an assistant professor of electrical engineering and applied mathematics at the University of Santa Clara in California, teaching courses in probability and the theory of communications. His research interest is in the application of probability theory to communication system engineering, and he is currently conducting an investigation of the statistics of signals generated by shift registers under a NASA grant. Dr. Healy is a member of Tau Beta Pi and Alpha Sigma Nu.

Scanning the issues

Advance tables of contents
Translated journals
Special publications

Conscience of the City. In an essay titled "Like It Is in the Alley," a Harvard research psychiatrist, who for many years has been working with urban and rural poor people, transcribes for us what a young boy named Peter has to say about his situation. Peter lives, Robert Coles writes, in the heart of what we in contemporary America have chosen (ironically, so far as history goes) to call an "urban ghetto."

"*In the alley it's mostly dark, even if the sun is out. But if you look around, you can find things. I know how to get into every building, except that it's like night once you're inside them, because they don't have lights. So, I stay here. You're better off. It's no good on the street. You can get hurt all the time, one way or the other. And in buildings, like I told you, it's bad in them, too. But here it's o.k. You can find your own corner, and if someone tries to move in you fight him off. We meet here all the time, and figure out what we'll do next. It might be a game, or over for some pool, or a coke or something. You need to have a place to start out from, and that's like it is in the alley; you can always know your buddy will be there, provided it's the right time. So you go there, and you're on your way, man.*"

Peter's plight—his vivid "you're on your way, man" tells us unmistakably what he needs is a way out—is shared by millions of others, both black and white. The modern city, the traditional place for a man to shake his fetters, has been and continues to be, for many millions, nothing more than a cage. In a zoo, the big cats pace back and forth, back and forth, along the bars, or they lie somnolent. For many, the city is only a zoo for our kind of animal.

What is to be done? For the fact is that the problem of the poor man in the city ghetto is but part of a much larger constellation of problems that have multiplied increasingly in recent years. Whether or not the city is a viable structure in our time has come seriously into question. A measure of just how seriously—beyond the accounts we all share through our newspapers—and a measure of the diversity of questions that can be asked about what may well be unsolvable problems comes to us in a volume of papers, of which Coles' "Like It Is in the Alley" is one.

The volume is a special issue of *Dædalus*, the journal of the American Academy of Arts and Sciences, and is an outgrowth of an AAAS study of the future of urbanism. At a time when private industries—and especially those organizations that are built around modern engineering—are being beseeched to contribute to the solution of urban problems, there is no question that engineers must take a much wider compass on such problems than their own specialized literature (including the best of their vaunted systems engineering) has thus far provided. This volume of *Dædalus*, entitled "The Conscience of the City," is earnestly recommended for your attention, whether you are working directly on urban problems or not. These essays reorient, they provoke, they stimulate, they temper one's hopes, they inform, they unsettle, and they startle, and, as in Coles' case—which weighs like the center of gravity through the direct individual human plight—they break your heart.

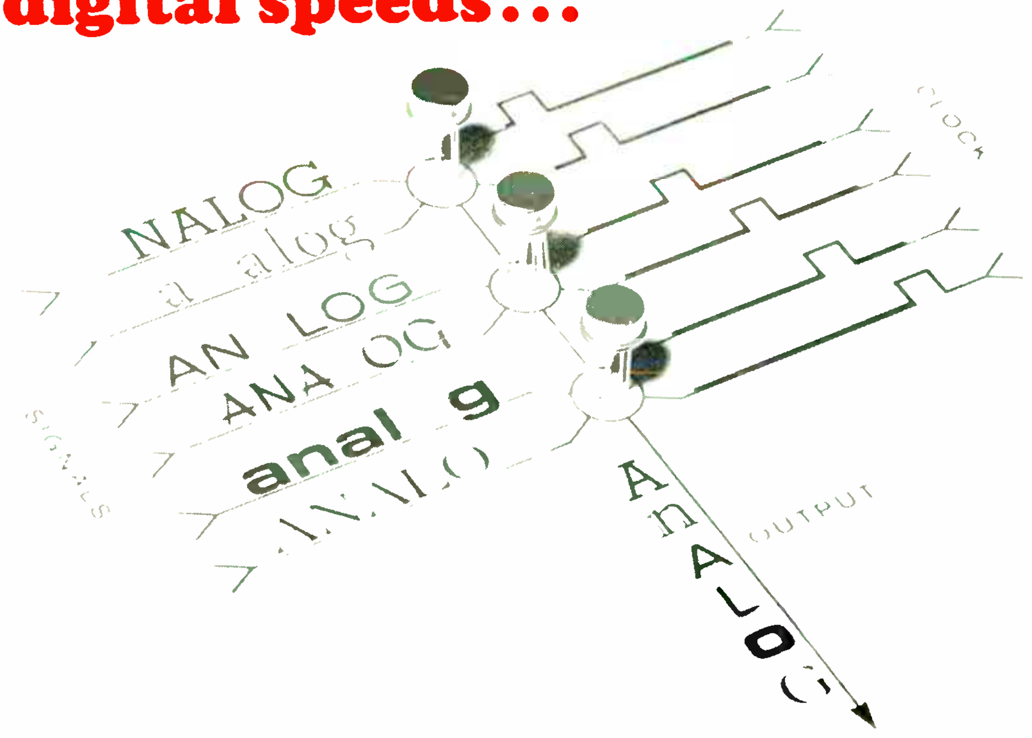
They reorient almost immediately in Melvin M. Webber's paper, "The Post-City Age," which opens the first part of the volume called "Traditional City in Transition." Webber's thesis is that current discussions about the "crisis of the cities" have been clouded by the misconception that the geographically bounded city is still the relevant unit for discussion. He sees this confusion stemming "from the anachronistic thoughtways we have carried over from the passing era. We still have no adequate descriptive terms for the emerging social order," he continues, "and so we use, perforce, old labels that are no longer fitting. Because we have named them so, we suppose that the problems manifested *inside* [our italics] cities are, therefore and somehow, 'city problems.' Because societies in the past had been spatially and locally structured, and because urban societies used to be exclusively city-based, we seem still to assume that territoriality is a necessary attribute of social systems."

But this conceptual error, Webber points out, is a serious one, "leading us to seek local solutions to problems whose causes are not of local origin and hence not susceptible to municipal treatment. We have been tempted to apply city-building instruments to correct social disorders, and we have then been surprised to find that they do not work."

The theme of our lack of understanding of the real nature of urban problems persists, in one form or an other, through all the papers. Edmund N. Bacon in "Urban Process," in "Part II: Processes and Goals for Change," opens baldly with the statement: "The failure of our cities is an intellectual one. It is brought about by the failure of the intellectuals to generate a viable concept of a modern city and a modern region." To attack the ignorance, Franklin A. Lindsay in "Managerial Innovation and the Cities" calls for, among other goals, greatly increased research on urban problems: technical, managerial, social, and economic. What is lacking now, he argues, "is a sense of direction and urgency. The most critical 'missing factors' are an understanding and acceptance of what is required to make headway on the complex interrelated social and economic problems, and the will to mobilize the necessary human and physical resources with the urgency that the problems demand." Lindsay holds forth a minimum goal to be reached ten years from now as a "level of R&D funding equal to one percent of the operating budgets of the nation's citie

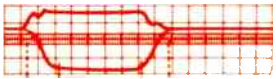
The inadequacy of current spending on research is pointed out by a recent report of the Arden House Conference

Now switch analog at digital speeds...

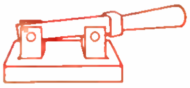


...and leave the driving to us

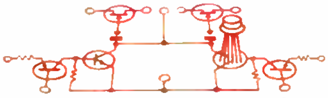
"Logic really turns us on", say our Hybrid FET Analog Gates with built-in driving circuitry. And turn on they do, in as little as 20 nanoseconds with a 10^8 ON/OFF ratio.



If you need a gate that operates in a nanowink and has an ON/OFF ratio like a knife switch,



we have stock Hybrid FET Analog Gates that fit this description to a T Five types are available, with R_{ON} as low as 6 ohms and speeds to 20 nanoseconds. [Shown above is our CAG13 Dual FET Analog Gate.]



All types include drive circuits for direct operation from DTL or TTL. They're fully described in our 24-page "Condensed Compendium Containing Comprehensively Characterized Hybrid ICs". . . . .

It's yours for the asking.

CRYSTALONICS

A TELEDYNE COMPANY

147 Sherman St. • Cambridge, Mass. • 02140

CALL YOUR LOCAL CRYSTALONICS OFFICE FOR DETAILS

(201) 233-4562
(201) 835-0332
(206) 323-5100
(213) 466-6201
(213) 469-5871
(214) 348-4283
(301) 668-4646
(303) 798-8439
(305) 647-4000
(305) 531-0200
(312) 547-6200
(313) 474-0661
(315) 457-1188
(317) 289-6322
(319) 363-6425
(415) 349-6316
(416) 751-5980
(505) 268-0928
(513) 299-4144
(514) 331-4884
(516) 593-2628
(602) 277-9739
(609) 784-5000
(612) 488-6634
(617) 491-1670
(617) 891-4582
(713) 526-7431
(716) 342-1413
(816) 471-2850

Circle No. 2 on Reader Service Card