IEEE spectrum

features

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

Bench or System — the HP 3450A gives you maximum performance in a minimum space.

A quick look at the unfolding dodecahedron shows each of the 12 functions the Incredible Dodecameter performs. What it doesn't show is just how well this 5-digit multifunction meter performs each function.

For instance, you not only get true rms capability—you also get value-plus features like true 4-terminal ac ratio testing and 4-terminal ohms measurements.

And, accurate, fast measurements in each of these twelve categories is only the start. You can add digital output and directly control external equipment like a printer. Or, add remote control and get full programmability for system use.

No matter what the application, you get more for your money with the HP 3450A.

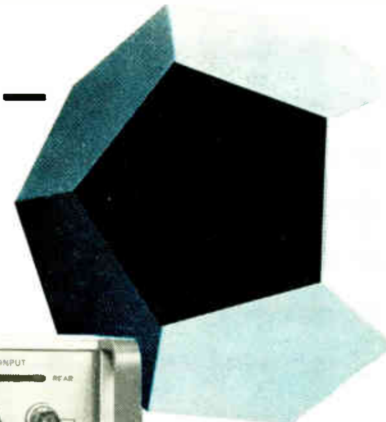
This Incredible Dodecameter lets you start with the basic dc meter and add the capability that best fits your requirements. If your needs change, any of the options (except the rear input terminals) can be easily installed in the field.

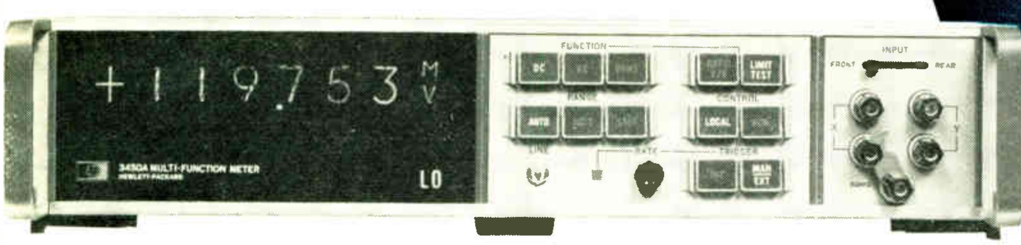
For more information on this outstanding 12 in 1 bargain, just call your local HP field engineer. Or, write Hewlett-Packard, Palo Alto, California 94304. Europe: 1217 Meyrin-Geneva, Switzerland.

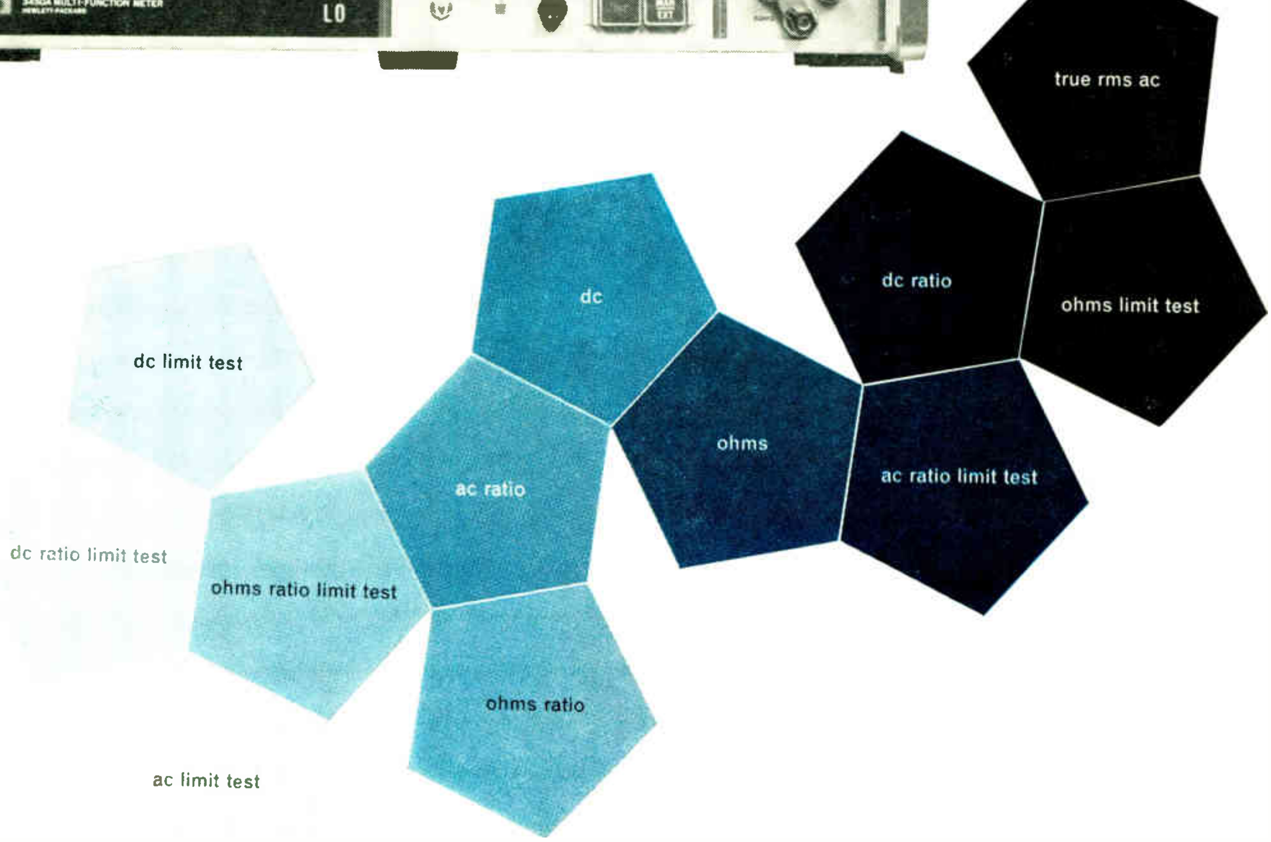
Price Basic HP 3450A, $3150; AC Option 001, $1250; Ohms Option 002, $400; Limit Test Option 003, $350; Digital Output Option 004, $175; Remote Control Option 005, $225; Rear Input Terminal Option 006, $50.

090/8

The Incredible Dodecameter — A 12-in-1 way to better measurements!







HEWLETT hp PACKARD

DIGITAL VOLTMETERS

Circle No. 5 on Reader Service Card.

the cover

The subminiature designs of integrated circuitry ought to be familiar to everyone by now. The particular labyrinth gracing this month's cover is a section from a general-purpose, high-performance monolithic regulator (MC1560) that is described in the article beginning on page 24

departments

IEEE SPECTRUM is published monthly by The Institute of Electrical and Electronics Engineers, Inc. Headquarters address: 345 East 47 Street, New York, N.Y. 10017. Cable address: ITRIPLEE. Telephone: 212—752-6800. Published at 20th and Northampton Sts., Easton, Pa. 18042. 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 $1.50. Nonmember subscriptions and additional member subscriptions available in either microfiche or printed form. Prices obtainable 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.

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 17 of the January 1970 issue. When in doubt, send the manuscript to E. K. Gannett, Editorial Services, at IEEE Headquarters, for forwarding to the correct party.

The technology behind electronically steerable phased arrays is the most rapidly advancing in the entire field of antennas. The growing number of applications has been providing the impetus for technical improvement and, conversely, advances in the technology have been broadening the scope of potential applications. Bob Sherman, Group Leader in AIL's Antenna Systems Division, proposes the use of phased arrays as a solution to the increasing number of difficult antenna problems.

An Octave Band Electronically Steerable Phased Array Antenna

AIL's Antenna Systems Division has been engaged in a program to advance the state of the art in electronically steerable array techniques. The effort has concentrated on the principal limitation of the present day phased arrays, namely bandwidth. An antenna was built to operate over a full octave in X-band, with nine switchable beam positions, five each in azimuth and elevation planes, with a common boresight. The intent was to move the beam to each of the nine positions, without altering its characteristics, and to have the nine positions remain invariant with frequency. The effort was extremely successful.

Three refinements are necessary to convert a narrowband array into a broadband one. The first, and least difficult to achieve, is to use antenna elements which operate satisfactorily over the full bandwidth. Secondly, the beam steering phase shifters must create differential phase shifts with the same phase versus frequency relationship as that existing in free space—that is energy in the medium must propagate in the TEM mode. Finally, true time delays, rather than phase delays, must be utilized. The wave emerging at the aperture is required to be coherent in time as well as in phase. The latter two requirements are discussed in more detail.

An array will point in that direction for which the phase contributions from each element in the array add coherently. Figure 1 illustrates this for the simple case of a linear array of N + 1 elements. The interelement spacing is D, and the point P is at an angle θ off boresight. The differential path length ϕ, is given by:

$$\phi = \frac{2\pi D}{\lambda} \sin\theta \qquad (1)$$

The beam steering phase shift for the kth element is ϕ_k'. The total phase at P from the kth element, ψ_k, is:

$$\psi_k = k\phi + \phi_k' = \frac{2\pi k D}{\lambda} \sin\theta + \phi_k' \qquad (2)$$

For phase coherence at P, ψ_k must be a constant, ndependent of k. For phase coherence over a broad band of frequencies, ϕ_k must vary directly with frequency (or inversely with λ) for ψ_k to be independent of k. This $1/\lambda$ dependence implies that the phase shifts must be made in a TEM medium such as free space, coaxial cable, or one of the several forms of stripline. Phase shifters mounted in any other environment, such as waveguide, make broadband operation impossible. In the AIL unit, lengths of microstrip line were used. If the line length in the kth element is $(N - k)D \sin\theta$, the total phase ψ_k is:

$$\psi_k = \frac{2\pi N D}{\lambda} \sin\theta \qquad (3)$$

independent of k for all frequencies—that is, the beam peaks θ degrees off boresight at all frequencies.

While the total phase ψ_k is independent of k, it is not independent of frequency. For single frequency operation, phase coherence at the angle θ is maintained for a phase delay system. For example, if the required ϕ' is 410 degrees, 50 degrees will do equally well. However, at half the frequency, the required phase is 205 degrees. The original 50-degree line length will look like 25 degrees, and the system will not work. The AIL time delay system uses the full required phase shift, or time delay, 410 degrees in the present example, and thereby satisfies all of the requirements for frequency independent beam pointing.

The AIL antenna utilizes a reflect-array feed technique, with microstrip mounted PIN diode phase shifters. The breadboard versions of the antenna and a typical phase shifter are shown in Figures 2 and 3, respectively. Figure 4 illustrates the measured radiation patterns from the five azimuth beam positions, displayed simultaneously on a single page. While the beamwidth varied with frequency, the beam position remained fixed over the full octave.

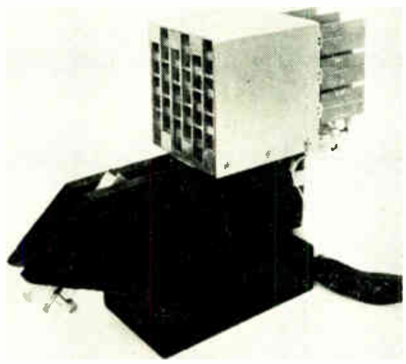


FIG. 2. ELECTRONICALLY STEERABLE PHASED ARRAY

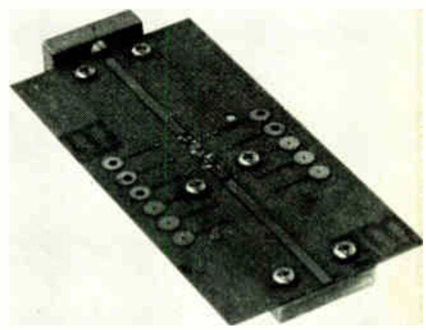
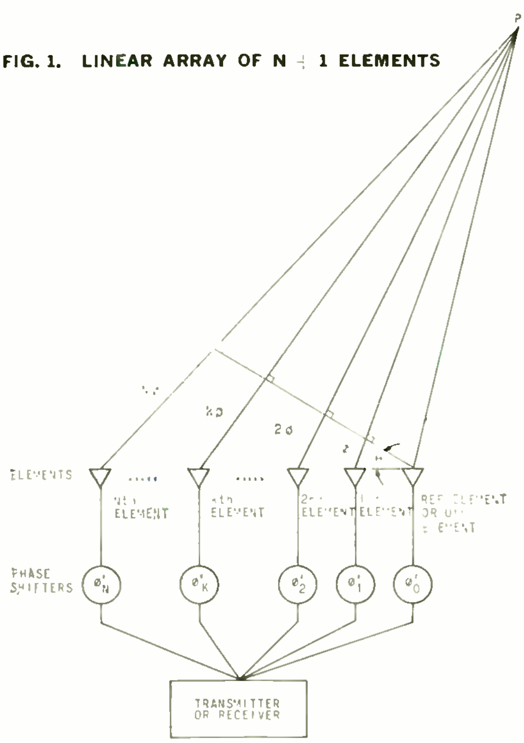


FIG. 3. PHASED ARRAY PHASE SHIFTER

FIG. 1. LINEAR ARRAY OF N + 1 ELEMENTS



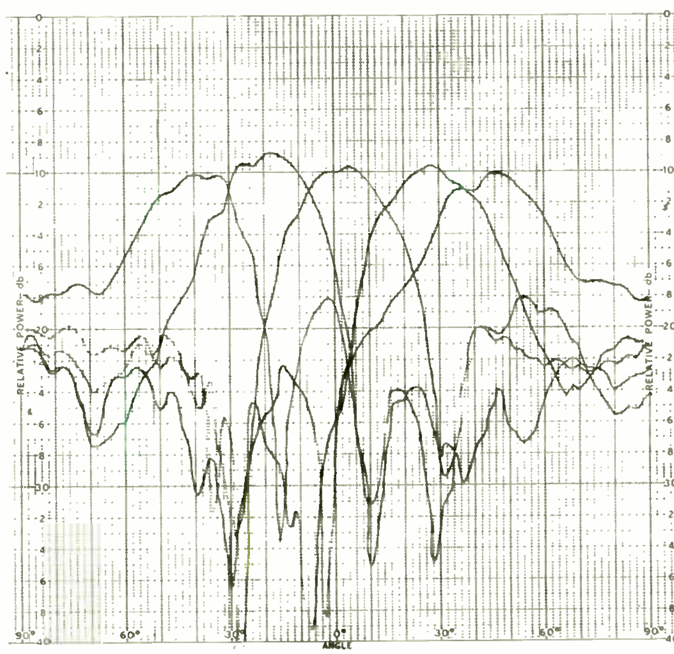


FIG. 4. MEASURED RADIATION PATTERNS

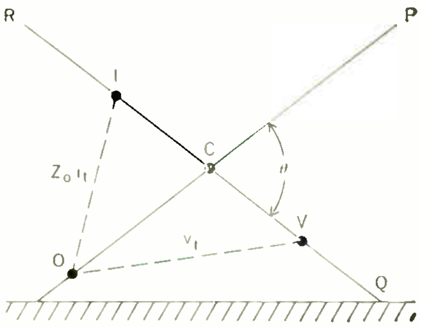
Forum

Readers are invited to comment in this department on material previously published in IEEE SPECTRUM; on the policies and operations of the IEEE; and on technical, economic, or social matters of interest to the electrical and electronics engineering profession.

Transmission-line model

The Esclangon diagram discussed by Michel Polouiadoff in the February *IEEE Spectrum* reminded me of a simple model used as a teaching aid by Prof. V. A. Bailey,* of the University of Sydney, as early as 1942.

The model could be quickly made with two bars *OP* and *RQ* from an Erector set, with the bars pivoted at point *C* and having an elastic band stretched around pegs at points *O*, *I*, and *V*, as shown in the accompanying figure.



With pegs *I* and *C* set equidistant from C, and O set to give $OC \neq rCI$, where *r* satisfies

$$\frac{Z_t - Z_0}{Z_t + Z_0} = r \angle\theta$$

it can be shown that, if *OI* represents Z_0, then *OV* represents Z_i. Setting the angle *PCQ* to θ provides an illustration of the terminal conditions; by pushing *P* and *R* together so that *OP* rotates through an angle $2\pi l/\lambda$ (and QR rotates through the angle $-2\pi l/\lambda$) conditions at any point *l* short of the termination are illustrated.

Appended is a list of some of the uses of the model as made by Professor Bailey.

A. G. Benedict
Minnetonka, Minn.

* Professor Bailey is possibly best known for his work on the Luxemburg effect; see reference F. E. Terman, *Radio Engineers' Handbook*, 1st ed., p. 732. He died circa 1966.

Use of transmission line model† (as developed by Prof. V. A. Bailey, 1942 or earlier)

As a preliminary to use of the model it should be noted:

(a) When Z_t is purely resistive, the initial position of the rods is either both horizontal ($Z_t > Z_0$) or both vertical ($Z_t < Z_0$).

(b) When $Z_t = Z_0$, *I* and *V* coincide with C.

(c) When Z_t is purely reactive, $OC = IC = IV$, with *O* on the left if Z_t is capacitive.

(d) When the output end is open, then in the initial position *I* coincides with *O* and both rods are horizontal.

(e) When the output end is short-circuited, then, in the initial position, *V* coincides with *O* and the rods are both vertical.

(f) A length $\lambda/4$ of the line corresponds to rotation of *OP* through 90°.

The model also demonstrates the following:

(g) In general, amplitude maximums and minimums of v appear alternately at sections $\lambda/4$ apart. Similarly, minimums and maximums of *i* appear at the same points.

(h) In general, the phase difference between v and *i* varies along the line, and is zero where v and *i* are at a maximum or minimum.

(i) When Z_t is reactive, or when the output end is short-circuited or open, v and *i* are in quadrature at all sections.

(j) When $Z_t = Z_0$ the amplitudes of v and *i* are both constant along the line.

(k) When Z_t is resistive but differs from Z_0,

$$\frac{i_{\min}}{i_{\max}} = \frac{Z_0}{Z_t}$$

also, at the output end, when $Z_t > Z_0$, v is a maximum, and when $Z_t < Z_0$, v is a minimum.

(l) If the impedance of the line is Z_s when short-circuited, and Z_f when open, then $Z_0 = \sqrt{Z_s Z_f}$.

(m) When a low-impedance load is placed across the line, it will draw maximum current from points $N\lambda/2$ apart.

† I have a copy of Professor Bailey's derivations if they are not obvious.

Ultrasonics in dentistry

Nilo Lindgren's review of "Ultrasonics in Medicine" in the November 1969 issue of *Spectrum* discussed medical ultrasonic scanning and ultrasonograms in diagnostics at length. The review of ultrasonics in dentistry was brief and omitted several recent contributions.

Ultrasonic drilling of teeth is a technique that has been almost completely replaced by high-speed turbine drilling. However, ultrasonic equipment is in widespread use in dental offices for dental plaque and tartar removal (Ewen, S. J., and Tascher, P. J., "Clinical uses of ultrasonic root scaler," *J. Periodontal.*, vol. 29, pp. 45–49, 1958). Also, ultrasound is in development as a tool for dental diagnosis by creating ultrasonograms of individual teeth (Smirnow, R., and Wolfe, M., "Preliminary illumination of oral structure by high frequency pulsed ultrasound," *J. Dent. Res.*, vol. 46 abstr. 439, IADR, 1967).

The use of ultrasonic techniques for bonding to dental enamel won for Dr. Hoffman the Albert Joachim Prize of the World Dental Congress in 1967. With this technique, metal is bonded directly to tooth tissue (Hoffman, R., and Gross, L., "Bonding aluminum to enamel with ultrasonic energy," *J. Dent. Res.*, vol. 44, pp. 366–373, 1965; Hoffman, R., Gross, L., and Ioppolo, A., "Aluminum bonded to dental enamel with ultrasonic energy: thermal stability," *J. Dent. Res.*, vol. 46, pp. 1048–1050, 1967; Hoffman, R., "Effect of surface treatment on the reliability of ultrasonic bonding to dental enamel," *Amer. J. Orthodont.*, vol. 52, pp. 721–731, 1966; Hoffman, R., "Ultrasonic bonding of aluminum to human dental enamel: bond strength and biological stability," *Int. Dent. J.*, vol. 17, no. 3, pp. 619–653, 1967).

Dental practice often requires the replacement of lost tooth structure with cast prostheses. An ultrasonic method of producing dental castings has been developed with the aim of improving precision and fit of dental castings (Hoffman, R., Gross, L., and Jefson, R., *N.Y. State Dent. J.*, vol. 24, pp. 15–21, 1968; Gross, L., Hoffman, R., and Jefson, R., "Ultrasonic precision casting," *Ultrasonics*, October 1969).

We note that ultrasonics has been used therapeutically in eye research. Recently, work was reported on refastening of detached retinas by means of low-frequency ultrasound (Karlen, D., "Choreoretinal lesions in rabbits by low frequency ultrasound," *Am. J. Ophthalmol.*, vol. 68, pp. 84–91, 1969).

Leo Gross, Ph.D.
Waldemar Medical Research Foundation, Inc.
Woodbury, N.Y.

A dim view of status

To improve the status of engineering, it will be necessary to make the colleges of engineering graduate schools like those of law and medicine. I doubt that anything will be done. It is reported that there are more than one million engineers in the U.S., which is at least ten times as many as can be employed in real engineering work. Most engineers work as technicians, repair men, or messenger boys. How can the status of engineers be improved under such conditions? It cannot.

Furthermore, state registration is a waste of time. How many state registered technicians, repair men, and messenger boys do we need? What sense does it make to graduate from a state-accredited college of engineering and then to take a state test? Doesn't the state believe in its own accreditation system? Worst of all, the best engineers are the most stupid about everything else, and are too dumb to realize when they are being taken advantage of. My advice is to forget it and go into some other kind of work. Engineering as a profession is dead. Let it rest in peace.

Russell A. Pettis
New Brighton, Minn.

Biological effects of microwaves

The report (November 1969) by Paul Hersch on the Symposium on "Biological Effects and Health Implications of Microwave Radiation" was a good attempt to cover a conference with some detail as well as commentary. The many inaccuracies included in the report, however, should be corrected.

First of all, Fig. 1, showing the allowable microwave exposure levels for the United States and Russia is incorrect. In the U.S., for durations less than 0.1 hour the maximum allowable exposure is permitted to go above 10 mW/cm^2 but only gradually and not discontinuously as drawn in the figure. The Russian standard as promulgated in 1968 gives 0.01 mW/cm^2 as the exposure level permitted for a full working day. No mention of the lower level of 1 μW/cm^2 shown by Hersch is made.

The suggestion that the situation in the United States is one of "statelessness" (whatever that may mean) is not tenable. Although there is a large amount of misinformation and misinterpretation, it is only fair to say that in this country there is a rather large amount of data which does have substance and has been subjected to critical analysis.

Hersch makes the statement, "Microwaves, of course, have been a potential health hazard for more than a quarter century." To give a proper perspective on the safety of microwave ovens, at least, we feel that Hersch should have placed more emphasis on the word "potential" rather than "hazard," since to date there is no evidence whatsoever of injury due to microwave exposure near these ovens. Because of the limited time of exposure, the rapid decrease of field levels with increasing distance from an oven, and the CW nature of the microwave energy, there are inherent safety factors that would indicate that an emission level of as high as 100 mW/cm^2 at 5 cm from an oven leakage point cannot be considered dangerous. By comparison, the proposed (by HEW) level of 1 mW/cm^2 at 5 cm has an enormous safety factor that hardly deserves Hersch's evaluation of "a figure supposedly with a built-in safety factor." (The probable exposure levels at a distance of two or three feet are generally two or three orders of magnitude lower than the figure quoted for 5 cm.)

Furthermore, if Hersch were to explore further HEW's efforts on the oven leakage standard, he would see that much of the drive for lower levels is based less on need for safety as determined by science, than on arguments based on technical feasibility coupled with a philosophic desire to protect the public against potentially hazardous effects not yet known to exist.

The reference to the low impedance of biological tissue would suggest a basis for reduced hazard rather than for hot spots. The reason for the possibility of the latter is the irregular curved nature of biological boundaries, which permits nonuniform field distributions through refraction and diffraction effects, but these are countered by the shunting property of low impedance, which effects a reduction in the *E*-field.

The comments on inefficiency of ovens and the magnetrons that power the ovens are in glaring contradiction to the generally acknowledged efficiency of ~50 percent of the magnetrons actually used.

The statement that there was testimony at the conference to suggest biologic responses at levels of the order of 10 $\mu W/cm^2$ is inaccurate, since there were no data of biologic effects presented at this low level of microwave exposure.

The experiments on orientation of microscopic organisms were performed at frequencies below 50 MHz. Although the authors did not report power levels carefully, it has generally been concluded that the effective power density levels, as related to the *E*-field by far-field expressions, were at least three orders of magnitude higher than 10 mW/cm^2.

This discussion of 2450-MHz oven leakage is misleading. Contrary to Hersch, this energy *does penetrate* the skin. This should be obvious since this property of penetration is one of the bases for rapid microwave cooking.

Sol M. Michaelson
University of Rochester, N.Y.

John M. Osepchuk
Raytheon Research Division
Waltham, Mass.

Insurance

From time to time, letters have appeared in *Spectrum* presenting gripes and grumbling of one sort or another, concerning the IEEE group insurance program, its advertising, premium rates, and what-not.

I recognize that IEEE does not prevail actively on the membership to join the insurance program, but nevertheless it should be worth noting that the insurance carrier has found it possible to give us half a year's life insurance for *nothing* this time around, because of favorable experience in the program.

As they say in the investment business, "past performance is no guarantee of future results"; however, this kind of performance should suffice to stifle some of the grousing for a while.

Personally, I think the insurance program is a worthwhile institution, and that most of us should be pleased that the Institute can and does provide this sort of benefit to the membership on a voluntary basis.

Willard A. Dodge, Jr.
Palos Verdes Estates, Calif.

Problems of an Asian member

I simply cannot but congratulate you and your team for making *Spectrum* very readable by including a series of thought-provoking papers on a large variety of topics. The views expressed about the social implications and the role of engineers in the making of correct decisions have a universal appeal, and similar articles have a perfectly justified place in *Spectrum*. Even to list the articles that made a lasting impression on me would be a big task, and I do not wish to do injustice to other contributors by mentioning only a few of them. I am sure that the present standard will be maintained and even improved.

I wish to point out some of my difficulties in properly utilizing the benefits of membership in the IEEE. I hope that they are representative ones for the majority of members from the developing nations. It is difficult to get the release of necessary foreign exchange for the payment of dues and considerable delay occurs in the process. Moreover, we can benefit only by receiving the publications of the Institute, since the other membership benefits are only theoretically available. I would like to know about the comparative economics of being a member and being a mere subscriber to various publications. I wish to emphasize the need for more incentive for continuation of membership. I am not in a position to join the professional Groups of my interest for obvious reasons.

The other difficulty is about the delayed receipt of *Spectrum* and the *Proceedings*. The December issue of *Spectrum* was received only three days back (February 3) and November *Proceedings* some ten days back. The issues are received in jumbled sequence and it becomes difficult to keep track of them to ensure that none are lost in transit. The readers' service cards also lose their usefulness because of expiration dates.

Shrinivas N. Sardesai
Bhilai (M.P), India

Annual subscription prices for nonmembers are:

Quarterly *Transactions*	$18
Bimonthly *Transactions*	$24
Monthly *Transactions*	$36
Spectrum	$24
Proceedings	$36

As a partial offset to the unavailability of some of the other services, members outside the U.S. and Canada receive a discount of 50 percent on all IEEE conference publications.

Spectrum and the *Proceedings*, as well as all *Transactions*, are now issued in microfiche editions; these travel by air mail to all users outside of the U.S. and Canada.

The possibility of printing in Asia for Asian members was considered some time ago; at that time it appeared uneconomical, but the question is at present being reexamined.

The Editor

An optical illusion

The apparent glorification of the goods-producing engineer by Dr. McCue ("Spectral Lines," December 1969) appears to be rather short-sighted in view of the drawbacks of a production- and GNP-oriented society —pollution, pursuit of the fast buck, and boredom among the members with their "assembly line" jobs. It seems likely that the present technology may give way to something slightly less mechanical, possibly including more study of the "useless" liberal arts. Will McCue's ideal engineer feel comfortable after such a change?

John Schmelzer
South Dakota State University
Brookings, S.Dak.

The "apparent glorification of the goods-producing engineer" is not short-sighted; it is pure illusion. The editorial dealt with a goods-producing society as a whole. It made no mention of an ideal engineer.

The role of the engineer in an affluent and effluent society received attention in "Spectral Lines" for March.

The Editor

Spectral lines

Can industry help? It was recently made known that the number of proposals in the U.S. to the National Science Foundation Instructional Scientific Equipment Program in electrical engineering was up about 20 percent this year over last. At the same time, funds were reduced because of budget cuts, and it was announced that the program is likely to be discontinued next year. The equipment programs of other agencies seem certain to experience similar cutbacks.

The need for modern laboratory equipment for electrical-engineering education is being felt more keenly now than it has been for at least a decade. There seem to be at least two reasons for this growing need. First, the pendulum is beginning to swing from the trend to stress concepts and logical development toward increased attention to laboratory skills and techniques, especially as directed toward design. Second, the digital revolution has reached a point where major changes can be made in the nature of the laboratories themselves. It is unfortunate that the means for obtaining equipment should begin to vanish at the same time that the new trend becomes pronounced.

The desire to have the products of a rapidly developing technology in the laboratory has associated with it a high price tag. Every student should have access to a computer-systems laboratory, now made practical by the development of the minicomputer. There should also be experience available in solid-state devices and integrated circuits, and perhaps a solid-state-materials laboratory. A computer suitable for hands-on experience will cost roughly $25 000, depending on peripherals. A well-equipped solid-state fabrication laboratory will cost more than $100 000 and may run to $1 million. Supplies for maintaining these laboratories are expensive also. At the same time, we must retain the more conventional laboratories in electronic circuits, electromagnetic fields, microwaves, power systems, etc. The situation is further aggravated by the high cost of computer time on a general-purpose computer, which is now becoming an important part of many classes.

University funds for the purchase of new equipment typically are small, perhaps of the order of a few thousand dollars each year. Sometimes it has been possible to borrow equipment from graduate research laboratories. But many universities remain underequipped with little prospect of securing enough apparatus so that each student can have as much laboratory experience as he needs.

There are at least two alternatives: (1) the contact with laboratory equipment will be restricted by stressing fewer kinds of equipment; or (2) a new source of assistance must be found.

It is difficult for any university to make the decision not to keep up with the Joneses. But with limited resources, some laboratory experience may have to be postponed until after graduation. A solid-state fabrication laboratory, for example, is costly and can accomodate only a few undergraduate students. The use of the computer in the "batch process" mode is relatively inexpensive and within the grasp of all students. Each university may find it necessary to rethink priorities in the face of very restricted budgets.

A source of help in the past has been the electrical and electronics industry. Some companies have established gift programs, making available equipment or supplies—often valuable and modern. Such programs have produced a trickle when a torrent is really needed. Can industry and education work together to make this the new source of assistance?

An enterprising, newly appointed chairman at a midwestern university was able to secure gifts totaling $80 000 in value from local industry by personally explaining his problem at each company. He asked for the specific equipment he needed to build two new laboratories, and he was able to obtain it. Needless to say, it was accomplished at the price of considerable time and energy.

Instruments no longer adequate for precision work may well be perfectly suited to laboratory instruction. Devices that fail to meet high-tolerance specifications can be useful in an instructional laboratory. Computers, both the analog and digital varieties, can be valuable if not completely out of date and not too specialized in their intended use.

Unused equipment and supplies in industry are likely to be kept in storage for possible later use, or sold to second-hand dealers for disposal. Can industry be convinced that this equipment will be of greater value if used in the training of future engineers? If gifts are not possible, then the university should be given the same opportunity as the junk dealer to purchase equipment and supplies at low prices.

If university–industry cooperation is possible, then coordination is essential. A clearinghouse or information bank is needed to which can be sent lists of equipment available and lists of equipment desired. This may be a function that IEEE can serve.

M. E. Van Valkenburg
Chairman, Publications Board

Monolithic voltage

The small size and high speed of monolithic regulators have opened the way to on-the-spot regulation; moreover, the low cost and high performance of these off-the-shelf devices have armed the engineer with a technique that drastically reduces design time

Don Kesner Motorola Inc.

Voltage regulation has been a specialty of semiconductors since the introduction of the Zener diode well over a decade ago. Now, there is a new dimension to the field—the monolithic voltage regulator. Available in a number of configurations that have specific advantages ranging over a variety of applications, this new component not only is giving its discrete counterpart a run for the money, but is showing a potential that easily surpasses even the most optimistic of earlier predictions.

The monolithic regulator offers the performance that enables it to replace discrete assemblies or portions of discrete assemblies in conventional central regulators. In addition, their small size, low cost, and ease of use uniquely fit monolithic units for convenient point-of-use regulation.

This idea of distributing unregulated voltage through equipment and providing local regulation has much to recommend it. Among the particular advantages are greater ease in tailoring voltage level and regulation to the requirements of individual equipment stages, as well as improved isolation and decoupling of these various stages.

Monolithic regulators can be expected to find their way into almost every kind of electronic equipment except the very-low-cost types that have traditionally excluded regulation circuitry as being too expensive. At the present time, integrated-circuit products are available that operate at anywhere from zero volts to more than 1000 volts. With appropriate external series-pass devices, the current range of such products extends to 60 amperes or more.

With these capabilities, monolithic regulators certainly deserve close inspection from all engineers who are called upon to design regulated power supplies. Probably the place to start an analysis of monolithic voltage regulators is with their basic function.

In other words, just what is it that a regulator should do?

The primary function of any voltage regulator is to hold the voltage in its output circuit at a predetermined value over the expected range of output or load currents. Working against the regulator are variations in load current, input voltage, and temperature. The degree to which a regulator can maintain a constant voltage in the face of these variations is called the basic "figure of merit."

Although there is a degree of interaction between these performance-degrading factors, especially between output current and temperature, it is most convenient to consider their effects separately.

Load regulation

The extent to which the output voltage is affected by output or load current is usually called load regulation and expressed as a percentage of the output voltage. Specifically, the general formula for load regulation is given as

$$\text{Load regulation (expressed as a percentage)} = \left(\frac{V_{ml} - V_{fl}}{V_o}\right) 100$$

where V_{ml} equals the output voltage with minimum-rated load, V_{fl} the output voltage with maximum-rated load, and V_o the nominal or reference output voltage (usually V_{ml} to minimize the numerical value of load regulation).

This relationship has been used for many years as a

regulators

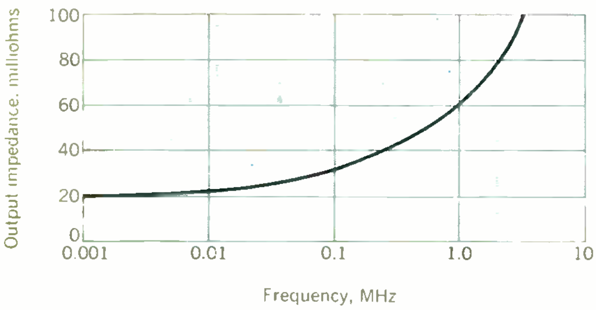


FIGURE 1. A particularly useful output specification is output impedance as a function of frequency. This particular curve applies to the MC1560 monolithic regulator, which has exceptional high-frequency performance.

standard regulator figure of merit. It is useful in this role, but has a very significant limitation. This limitation involves not what is in the formula, but rather what has been left out. Two regulators can have the same calculated load regulation, but provide substantially different regulation for transient loads.

In measuring load regulation, common practice is to adjust the load for minimum current, read the voltage, adjust for maximum current, and read the full-load voltage. In the ordinary course of the procedure, ample time is allowed for the regulator to react and stabilize the voltage. Unfortunately, many real loads are not so patient.

Regulators with limited frequency-response cannot respond to rapid fluctuations in load current. This is serious enough when a single load is being controlled by a regulator, since the voltage is uncontrolled for a period of time.

Often, however, it is even more serious when several loads are sharing a common regulated supply, since signal-frequency currents from one load can be easily coupled to other loads through the power supply.

Conventional load-regulation specifications don't cover this important aspect of regulator performance, but output impedance does. Output impedance is a small-signal ac parameter that can be measured over a wide range of load-current fluctuation rates. Output impedance can be related to load regulation by the following equation:

Load regulation (expressed as a percentage)

$$= \left(\frac{\Delta I \times Z_o}{V_o}\right) 100$$

where ΔI equals the change in load current, Z_o the output impedance of the regulator, and V_o the nominal or reference output voltage.

The output impedance should be low to minimize the necessity for effective load regulation. In a circuit, a low value of output impedance allows the regulator to function as a constant-voltage source—the basic idea of a voltage regulator.

Using output impedance as a figure of merit is an improvement only if the impedance value is appropriate for the rate of change encountered in the "real-world" application. (Using the data in Fig. 1 for calculating regulation with 1-kHz load fluctuations, an output impedance of 20 mΩ is appropriate. When 1-MHz transients are expected, however, an impedance value of 60 mΩ should be used.)

The primary value of an output-impedance specification is that output impedance is a function of the frequency of the load transients and can be measured over a wide frequency range.

Since output impedance is, in effect, a measure of the response characteristics of the regulator, the output

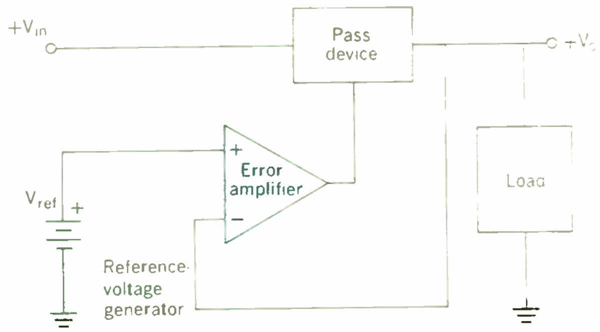


FIGURE 2. In a generalized series regulator, the output is compared with a reference signal. The output of the error amplifier is fed to the pass device, which, in turn, determines the final output.

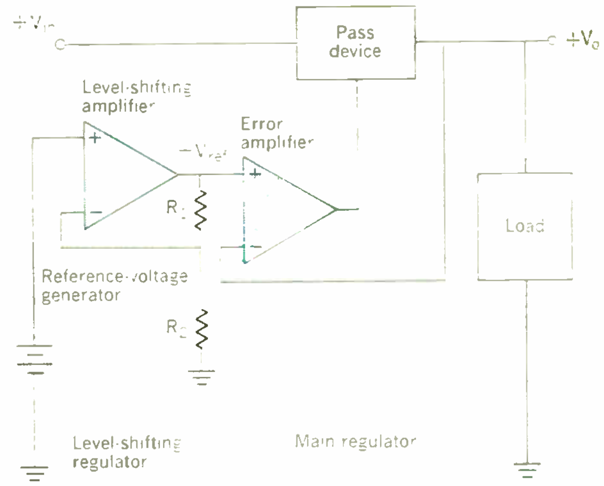


FIGURE 3. The basic approach to regulation employed in the MC1560/1561 and MC1569 involves three active stages. By separating the level-shifting function from the main regulator, a significant gain in performance is achieved.

impedance generally degrades (increases) with increasing signal frequency. A particularly meaningful specification consists of the maximum output impedance over a specified frequency range or a plot of impedance vs. frequency (see Fig. 1).

Input regulation

Input regulation is a measure of the effect of changes in the input voltage upon the output voltage. Comparable traditional specifications are line regulation and ripple rejection. Both of these terms suggest more limited meanings than input regulation and are not as descriptive.

Input regulation is given by

Input regulation (percentage of input voltage)

$$= \left(\frac{\Delta V_o}{\Delta V_{in} \times V_o}\right) 100$$

where ΔV_o = change in output voltage V_o for V_{in}, ΔV_{in} = change in input voltage, and V_o = nominal output voltage.

The V_o and V_{in} terms are frequently replaced with rms ac voltages, since sinusoidal input variations are common (power-supply ripple). Here again, a low value of input regulation is desirable.

Thermal stability

A primary thermal figure of merit for voltage regulators is temperature coefficient of output voltage. This is usually expressed in percent per degree variation from a reference voltage and temperature.

TC_{V_o} (percent per degree C)

$$= \frac{\pm (V_{o(max)} - V_{o(min)})}{V_{o(ref)}(T_{max} - T_{min})} 100$$

where $V_{o(max)}$ equals the output voltage at maximum-rated temperature (T_{max}), $V_{o(min)}$ the output voltage at minimum-rated temperature (T_{min}), $V_{o(ref)}$ the nominal output voltage (usually at 25°C), T_{max} the maximum operating temperature, and T_{min} is equal to the minimum operating temperature.

Since the temperature coefficient represents a change in output voltage for a change in temperature, a low value is desirable.

Another thermal consideration is a function of the regulator construction, and takes the form of a thermal feedback that ties regulator temperature to regulator power dissipation. Thermal feedback is most important in monolithic regulators and least important in large, very-well-cooled discrete assemblies. High-density discrete modules and hybrid regulators fall between these two extremes.

The power dissipation in a regulator is essentially proportional to the difference between input and output voltage multiplied by the load current. Thus, as the load current increases, the regulator power dissipation increases, which in turn increases the temperature of the regulator.

The effect of thermal feedback may be calculated by determining the temperature rise due to power dissipation and treating the result as an ambient temperature increase.

A monolithic regulator

Granted that the specifications just discussed indicate what a regulator must do, two questions remain: How does a monolithic regulator operate, and why does it offer the superior performance claimed?

The first point to keep in mind is that, although a monolithic regulator performs the same functions as its discrete counterpart, the circuit design may be quite different. Figure 2 shows the basic functional elements of a series voltage regulator.

These functions can be implemented with a few components or with many. In general, the higher the level of performance needed, the more complex the circuit and the greater the number of components required.

Monolithic regulators employ substantially more complex circuitry than all but the highest-performance, most-sophisticated discrete configurations. This is primarily a result of the economics of integrated-circuit fabrication. It is relatively inexpensive to add devices and low-level circuitry to an integrated circuit. Each additional transistor or diode on a monolithic chip adds

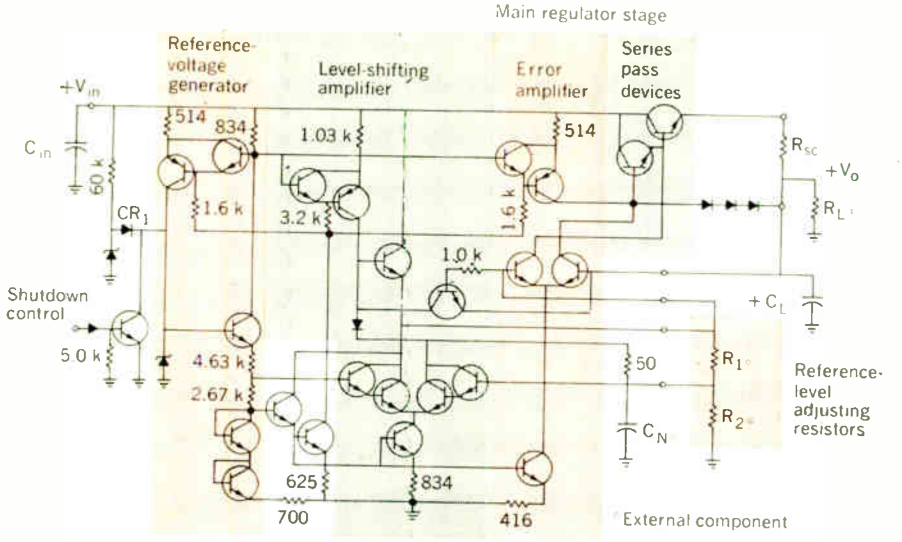


FIGURE 4. The MC1560/1561 is a relatively complex circuit. Because of the economics of IC production, however, the cost is moderate, with the complex circuitry providing very high performance.

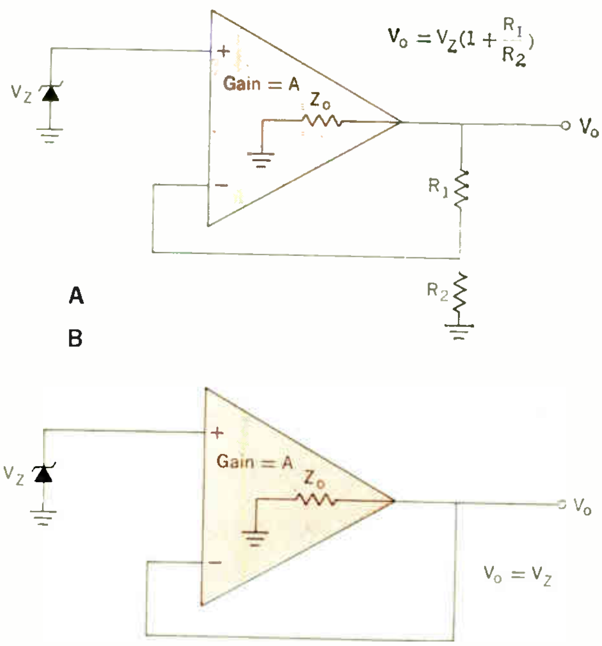


FIGURE 5. The basic advantage of separating the level-shifting and output functions is inherent in feedback operation. For this example, it is assumed that the amplifiers of (A) and (B) are identical.

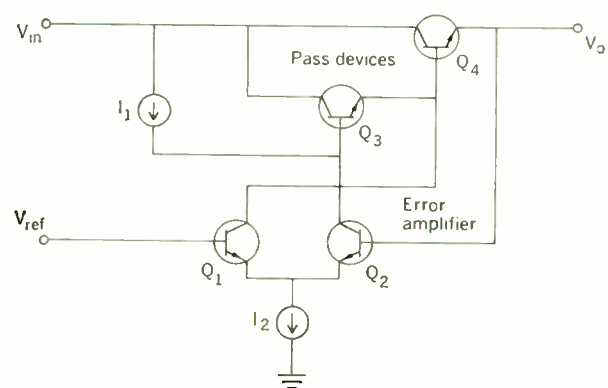


FIGURE 6. The unity-gain, main-regulator stage of the MC1560/1561 is shown in simplified form. In the actual monolithic device, the output transistor Q_4 can handle currents to 500 mA.

a much smaller cost increment than a comparable addition to a discrete circuit.

Monolithic regulators generally employ several stages and, in keeping with the usual linear-IC design practice, a high percentage of these stages are based on the differential amplifier. This tends to cancel process variations and provide high operational accuracy. A prime example of monolithic-regulator design is the Motorola MC1560/-1561—a general-purpose, high-performance device. (A functional representation of the MC1560/1561 is shown in Fig. 3.)

The MC1560/1561 is a three-stage design the output of which is adjustable over a wide voltage range. The first stage is a temperature-compensated reference-voltage generator. The second stage is a dc level-shifting, series voltage regulator; and the third stage is a unity-gain series main regulator with a current-handling capability of 500 mA. The schematic of the MC1560/1561 is displayed in Fig. 4.

The basic reason for this three-stage design is to combine operational flexibility with minimum output impedance. In a discrete regulator, the output voltage is usually determined by the selection of an appropriately valued Zener diode. Since the Zener function is performed on the chip in a monolithic regulator and it is desirable that the regulator be capable of operating at any selected voltage within its range, an alternate method of output-voltage selection is needed.

A highly effective procedure is to multiply the internally generated reference voltage by means of an amplifier with adjustable closed-loop gain. If this is done in a two-stage regulator, however, output impedance is sacrificed. To see why, consider the two amplifier configurations in Fig. 5. These amplifiers are identical except for the feedback configuration. The circuit of Fig. 5(A) amplifies the reference voltage whereas that in Fig. 5(B) does not.

The expression for the closed-loop output impedance of the configuration in Fig. 5(A) is

$$Z_{o(CL)} = \frac{Z_o\left(1 + \frac{R_1}{R_2}\right)}{A}$$

whereas the output impedance of the circuit of Fig. 5(B) reduces to

$$Z_{o(CL)} = \frac{Z_o}{A}$$

Since load regulation is

Regulation (expressed as a percentage)

$$= \frac{Z_{o(CL)} \times \Delta I_L \times 100}{V_o}$$

any increase in closed-loop output impedance results in degraded regulation. In the design of the MC1560/1561, a unity-gain output stage was chosen for the inherently better regulation resulting from its lower output impedance.

The unity-gain main regulator, depicted in simplified schematic form in Fig. 6, provides an output impedance that is essentially constant to very high frequencies. (For example, typical output impedance of the MC1560 at 1 MHz is only 0.060 ohm.)

The dc reference-voltage, level-shifting regulator (see Fig. 7) is also a differential amplifier. There are, however, three important distinctions between it and the output stage.

First, only a single series-pass device is needed since the level-shift regulator must control only enough current to drive the output stage; a Darlington configuration is used on the differential transistors to minimize current drain and reduce sensitivity to impedance levels; and external resistors R_1 and R_2 are employed to vary the gain of the amplifier.

The ratio of external resistors R_1 and R_2 determines the output voltage of the regulator. There is no requirement for these resistors to track any diffused resistors, but their temperature coefficients should be closely matched for maximum accuracy.

The zero-temperature-coefficient, reference-voltage generator (Fig. 8) combines a Zener diode driven by a constant-current source with a number of forward-biased junctions and a diffused resistor to compensate for temperature effects. The Zener and the diffused resistors have positive temperature coefficients and the forward-biased junctions have negative coefficients. The temperature coefficients of the three forward-biased junctions (two diodes and the base–emitter junction of the transistor) just offset the positive-temperature coefficients of the Zener diode and the diffused resistors. As a result, voltage V_S is almost totally independent of temperature. The other temperature-compensated reference voltage, V_P, is used to bias the constant-current source transistors for operating the main stages.

With this design, the MC1560, its companion MC1561 extended-voltage-range unit, and the MC1460/1461 for commercial temperature ranges, offer outstanding performance over an extremely important range of output

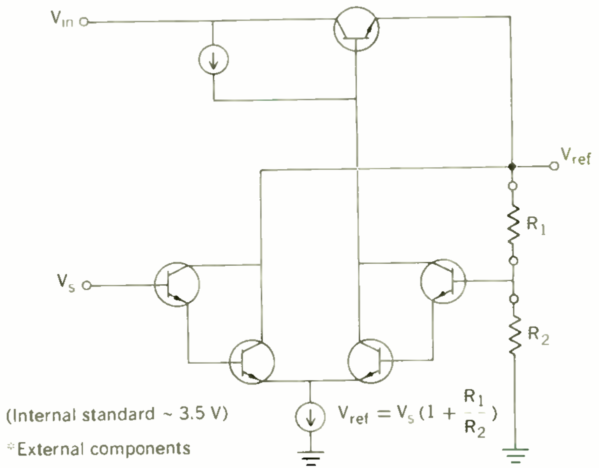


FIGURE 7. This simplified schematic shows the level-shifting regulator of the MC1560/1561. The ratio of the external resistors R_1 and R_2 determines the gain of the stage and, hence, the reference voltage for the output-regulation stage.

FIGURE 8. The reference-voltage generator stage provides a highly stable, temperature-compensated potential. The use of an active current source greatly improves ripple rejection. The V_P output biases the current source transistors.

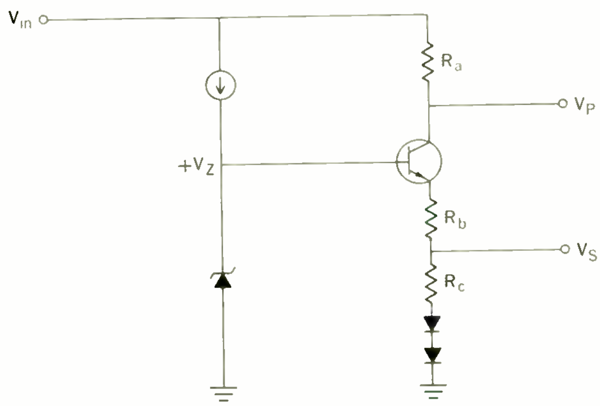
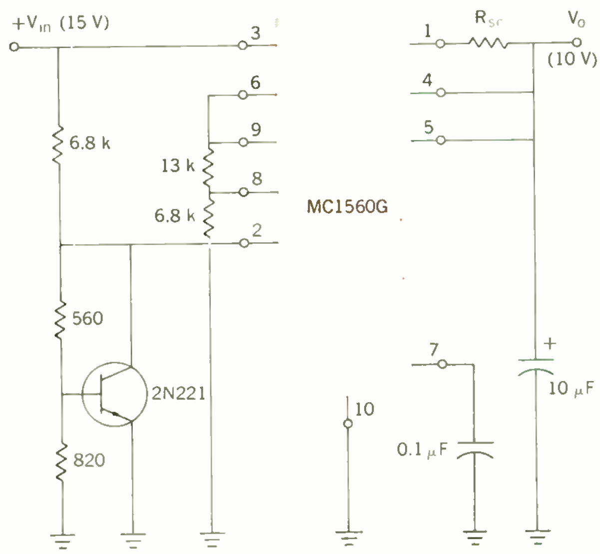


FIGURE 9. The MC1560 is shown with all external components needed for a 10-volt output from a 15-volt dc input. This configuration provides thermal shutdown if the regulator junction temperature reaches 140°C.



voltages and currents. The MC1561, for example, can provide an output voltage of from 2.5 to 37 volts. Operating load-current ranges are from 1 to 200 mA for devices mounted in the standard metal-can package; for devices mounted in the power package, the range is 1 to 500 mA and even this range can be extended by employing external series-pass transistors for higher current.

The circuits incorporate a provision for shutdown control that is compatible with the standard logic levels of RTL, DTL, and TTL. Short-circuit protection and current limiting are achieved by connecting an appropriately valued resistor between output and current-limit pins.

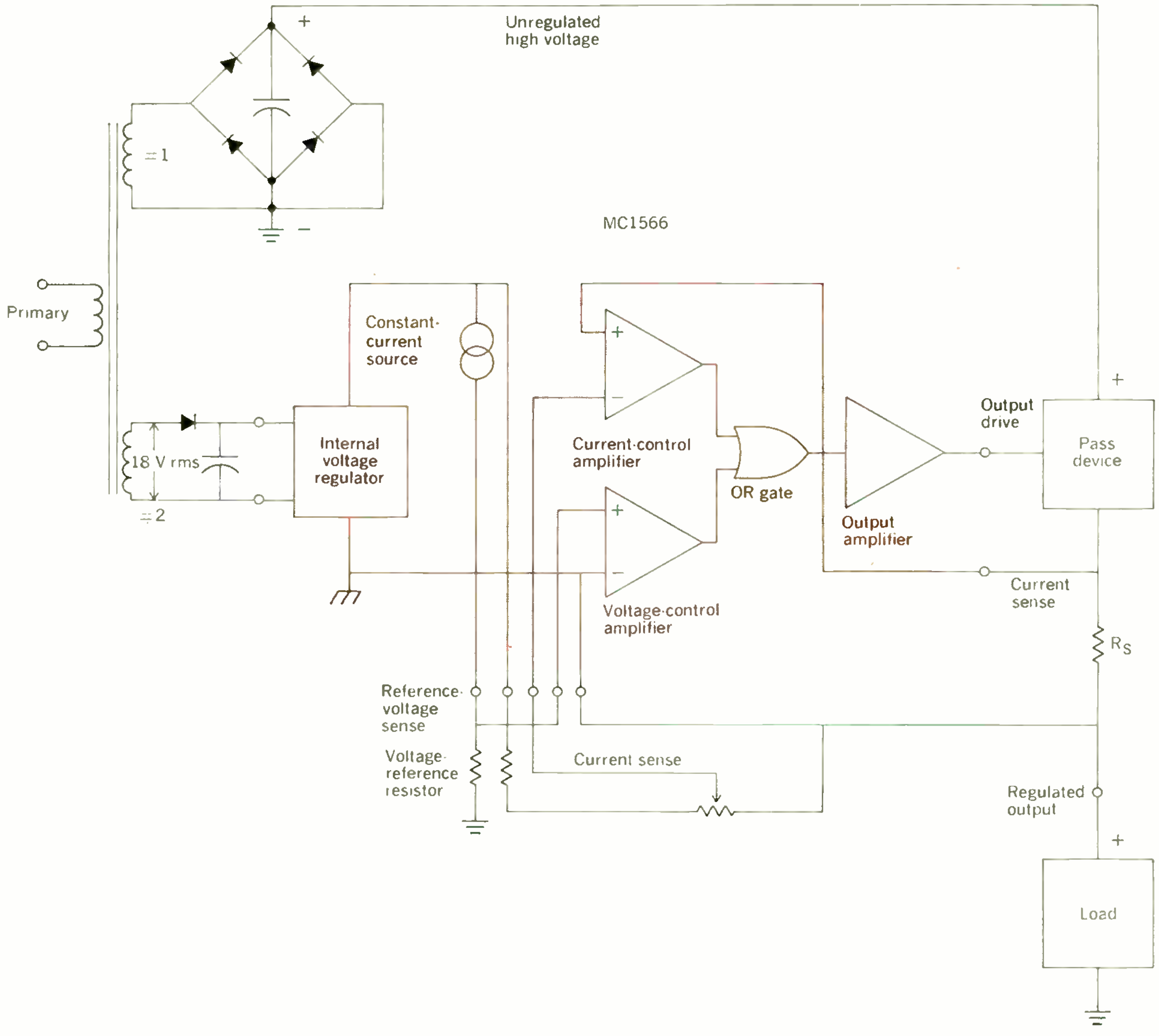
The thermal feedback inherent in a monolithic regulator can be used to provide automatic shutdown if the junction temperature exceeds a predetermined value. Such junction temperature-limiting protects the regulator in the event of load short circuits. With MC1560/1561, both the maximum short-circuit current (through external resistor R_{SC}) and maximum junction temperature can be controlled. The result is complete protection of the regulator. (Figure 9 shows a configuration employing a transistor network to bias the shutdown control. Designed for operation in a constant-temperature environment, this circuit shuts down the regulator when the junction temperature reaches approximately 140°C.)

The performance of the MC1560/1561 fully justifies the claim of outstanding performance made for monolithic regulators. For example, the output temperature coefficient is typically ±0.002%/°C. Input regulation is typically 0.002%/V_{in} with a maximum that extends to 0.015%/V_{in}.

Load regulation is typically 0.005 percent; for the extended voltage range, MC1561 has a typical absolute value of 0.4 mV. The MC1560/1561 series regulators extend load regulation to fast transients. Example: the regulation for 1-MHz load fluctuation is only degraded by a factor of three compared with the low-frequency value.

Sometimes the transient response of the MC1560/61

FIGURE 10. A method of completely freeing monolithic regulators from internal voltage restrictions is to "float" the chip on the power supply and sample current and voltage with differential amplifiers connected to external components.



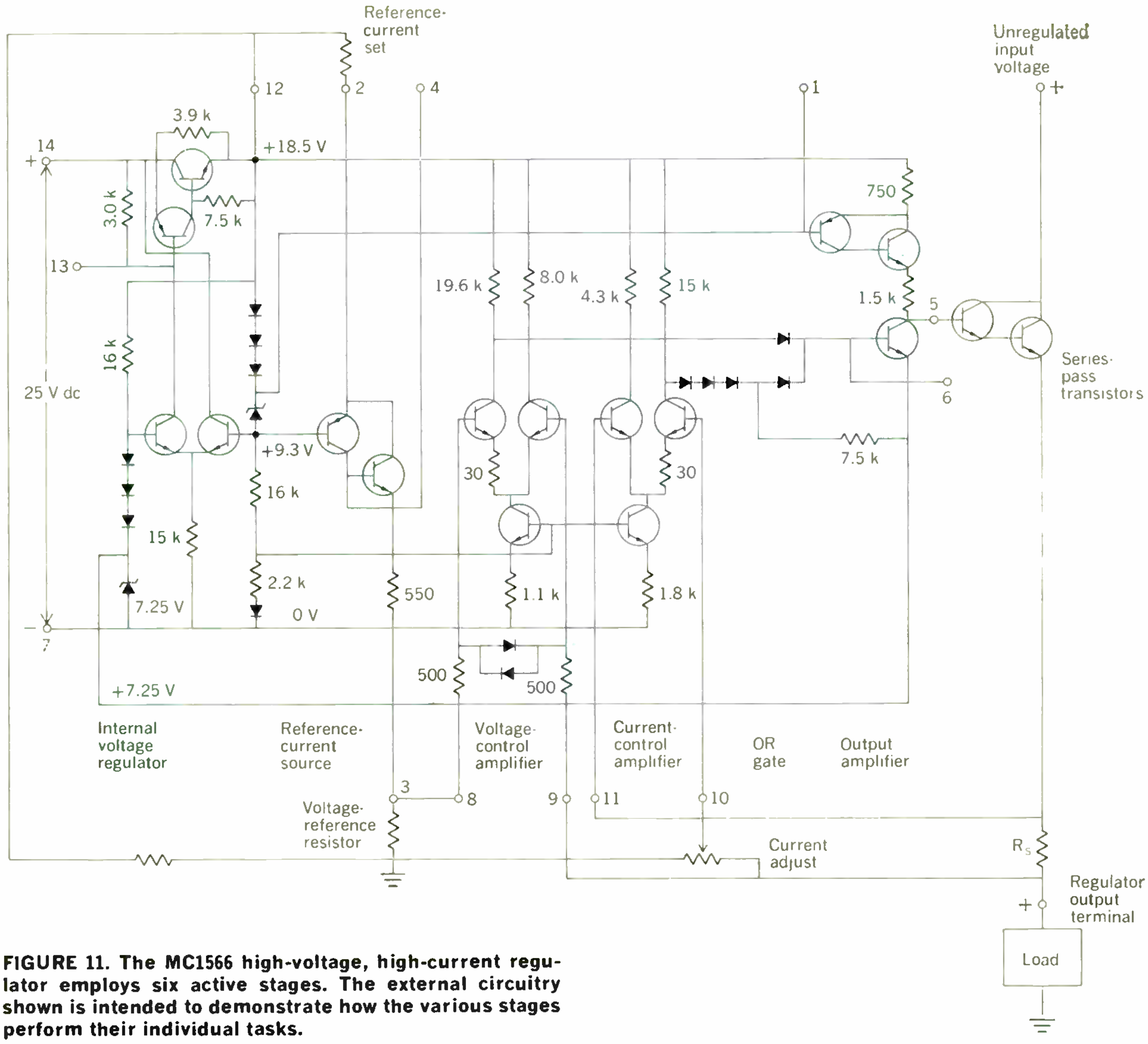
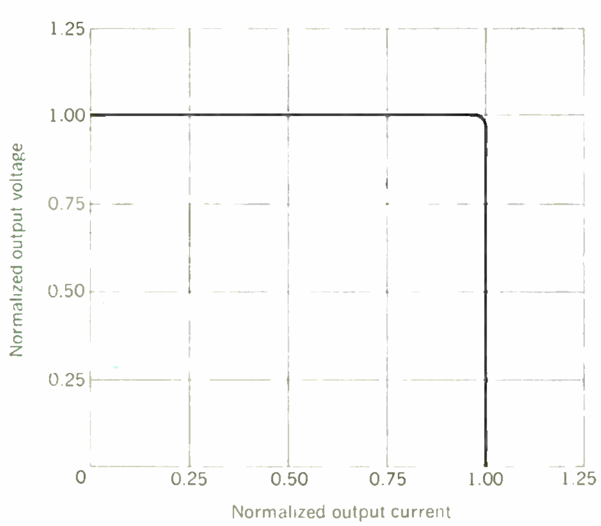


FIGURE 11. The MC1566 high-voltage, high-current regulator employs six active stages. The external circuitry shown is intended to demonstrate how the various stages perform their individual tasks.

FIGURE 12. The use of separate current-control and voltage-control amplifiers gives the MC1566 an output current-voltage characteristic with an extremely sharp knee. In normal operation, the voltage-control amplifier is active and holds the output voltage constant. If the load current exceeds the prescribed limit, the current-control amplifier takes over and the regulator operates in the constant-current mode.



voltage regulators is more than is actually needed, making it difficult to stabilize. For more general applications then, the MC1569R regulator is most suitable. It is very similar to the MC1560/61 and its dc performance characteristics are nearly the same.

The two points of difference between the MC1560/61 and MC1569 are frequency compensation (transient response) and current-limiting circuitry. With the MC1569, frequency compensation is accomplished by the addition of a capacitor from pin 4 to ground. The capacitor value range is 0.001 μF to 0.1 μF, the smaller value establishing a higher transient response.

Short-circuit-current limiting is accomplished by the addition of a transistor or diode string externally; both will produce the desired effect of drawing base current from the internal series-pass transistors. This "outside-the-package" protection has an advantage in that thermal feedback from the power device cannot degrade the "knee" region and upset regulation before current limiting is reached. If either the diode string (usually three devices in series) or the transistor had been located on the chip, threshold voltage would have been altered by approximately -2 mV/°C of junction-temperature rise.

When large amounts of load current are demanded from the chip, this junction temperature can rise quite rapidly (especially if the case is not properly mounted on an adequate heat sink) and current limiting could occur before the design value is actually reached. External limiting prevents this, providing a sharper, more consistent knee (for a constant ambient temperature). For changing ambient temperatures, the knee is definitely more predictable.

Extending the range

A medium-voltage regulator such as the MC1560 can be employed at higher voltages through the use of Zener or resistive voltage-divider networks. The regulation in such cases is a function of the portion of the total voltage that appears across the regulator output. The smaller the fraction of the total voltage across the regulator, the poorer the regulation.

Other possible routes to higher voltage capacity include special processing to enable the integrated circuit to withstand higher voltages. Such techniques increase costs disproportionately and provide only modest increases in voltage capability. An approach that frees the monolithic regulator completely from voltage restrictions, however, is reflected in a unique Motorola regulator—the MC1566.

The MC1566 is not designed to be operated as an independent regulator, but as the voltage and current control element in a composite monolithic-discrete regulator. This splitting of the regulation function allows the IC to be optimized for sensing and control while delegating the voltage- and current-handling problems to silicon power transistors designed for the task.

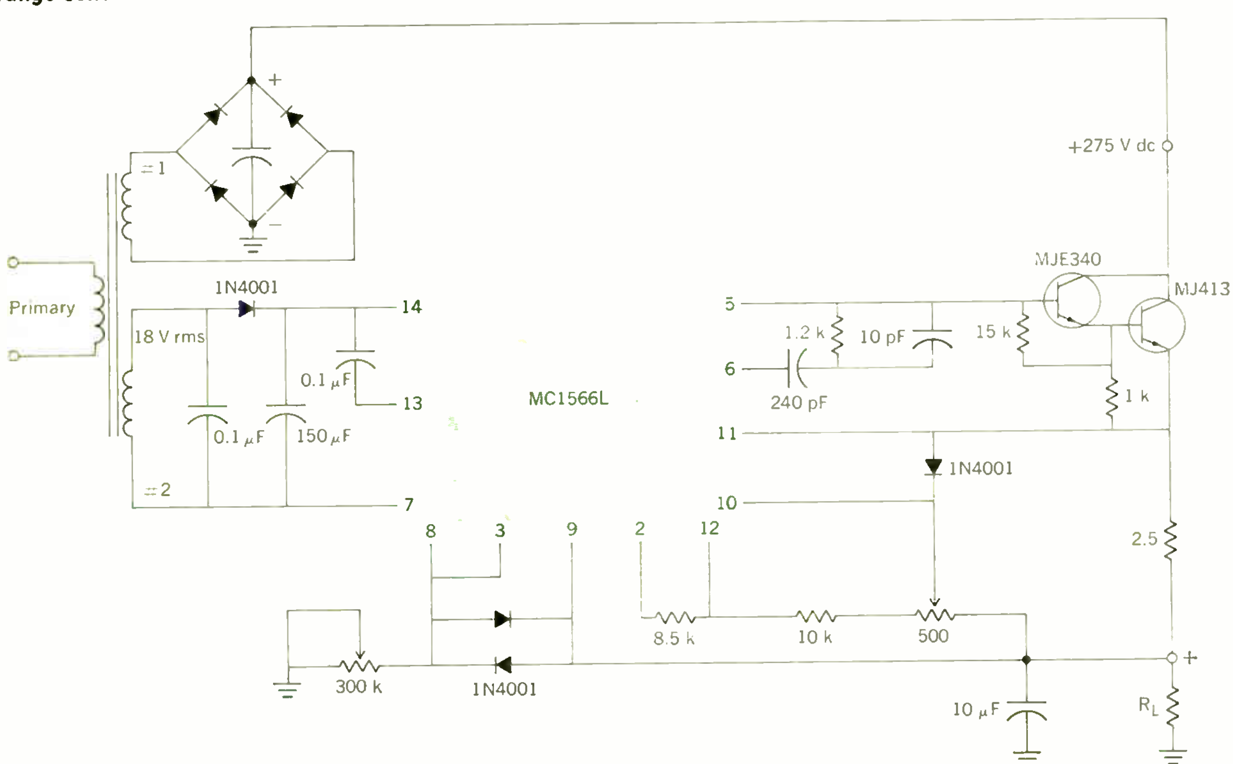
A complete high-voltage regulator, such as that described in Fig. 10, includes an MC1566, one or more series-pass transistors, rectifiers, filters, capacitors, and a transformer with two secondary windings. The transformer provides the basis for the "floating" operation that allows the MC1566 to control accurately both output currents and high voltages.

Secondary winding 1 provides the unregulated high-voltage input to the overall regulator. Winding 2 provides the 18 volts rms necessary to operate the MC1566. Because of the independent windings, the power supply for the MC1566 is isolated from both the transformer primary winding and the high-voltage circuit. The integrated circuit is then tied to the high-voltage loop "downstream" of the series-pass device at the output of the overall regulator. As a result, the potential on the MC1566 tracks the overall regulator output and "floats" at its level.

To accomplish its voltage- and current-control functions, the MC1566 incorporates six stages: an internal voltage regulator, an adjustable constant-current source, a voltage-control amplifier, a current-control amplifier, an OR gate, and an output amplifier (see Fig. 11). The functions of the first and last stages are straightforward. The internal voltage regulator provides stable voltages for operating the integrated circuit; the output amplifier provides the current drive and the proper voltage levels to control the series-pass devices.

The voltage-control amplifier functions as the error-

FIGURE 13. With the protective, frequency-compensating, and level-setting components shown, the MC1566 is used as the basis for a 0–250-volt, 0.1-ampere "laboratory quality" power supply. By varying the effective resistance of the 300-kΩ potentiometer, full voltage-range control is obtained.



sensing amplifier in a conventional regulator. The way it obtains the voltages to be compared, however, is unusual. One input to the differential-amplifier–voltage-control stage consists of the positive output of the overall regulator. This "output sense" is compared with a reference voltage obtained by passing a known current through an external resistor. The current is obtained from the reference-current-source stage of the MC1566. The value of the external voltage-adjusting resistor is selected so that the voltage drop across it is equal to the desired output voltage. The voltage-control amplifier compares the reference voltage (desired output voltage) with the actual output voltage. The output of the voltage-control amplifier is, in turn, used to control the output amplifier.

The current-control amplifier is a differential stage similar to the voltage-control amplifier. Its inputs, however, are taken from either side of an external current-sense resistor R_S. The value of R_S is selected such that, when maximum load current is flowing, the voltage drop across R_S is 250 mV. When the drop across R_S exceeds 250 mV, the current-control amplifier signals the output amplifier to reduce current drive.

When adjustable current operation is desired, a voltage divider is inserted in series with one input of the current-control amplifier. This permits in-use adjustment of the point at which current limiting occurs.

The function of the OR gate is to determine whether the voltage-control amplifier or the current-control amplifier controls the output amplifier. Normally, the voltage-control amplifier signal is dominant, but if either amplifier signals for a reduction in drive, that amplifier's signal is the one that controls.

The result of this six-stage design is a control element that offers outstanding performance at any voltage or current level. The voltage and current capabilities of MC1566-controlled regulators are not limited by the MC1566, but by the safe operating area of the external series-pass devices.

Typical input-voltage regulation and load-voltage regulation are 0.1 percent plus 1 mV; current regulation is typically within 0.1 percent plus 1 mA. Other notable features include short-circuit protection and an output voltage adjustable all the way to zero. The separate voltage-control and current-control amplifiers joined by the OR gate provide voltage and current regulation with automatic crossover. The crossover from constant-voltage to current-limited operation is extremely sharp, as is shown in Fig. 12.

An example of the capabilities of the MC1566 is given in the 0–250-volt, 0.1-ampere regulator detailed in Fig. 13. With proper selection of external components, the MC1566 can be used to provide "laboratory" power-supply performance.

Using the regulator

When applied to system requirements, monolithic regulators offer a number of advantages. One of the most significant is better regulation for a given dollar expenditure. In short, performance of monolithic regulators is extremely high whereas costs simply are not.

The small-size and low-cost features certainly qualify the monolithic regulator uniquely for local regulation. The extremely high-speed capability of these devices means that effective local regulation of communications equipment and computer boards is now practical. Hand in hand with these technical and price advantages is a significant reduction in the engineering effort necessary to design regulated power supplies.

Instead of a major engineering project, design of a power supply using the MC1560 family literally involves just a matter of minutes. The user need only determine the current and voltage requirements so as to select the most economical unit that will meet them. Once the right regulator and package are chosen, the only remaining design effort is to calculate the appropriate values for the voltage-setting resistors to provide the required output voltage, and to calculate the appropriate value for the current-limiting resistor in order to protect against overloads.

As a conservative estimate, this procedure may take as long as ten minutes the first time a designer decides to use a monolithic regulator.

When higher-current values are required or if a switching regulator is desired, the design effort will be necessarily greater in order to select the external circuitry, but is still substantially shorter than starting from scratch with a do-it-yourself regulator.

A look ahead

Monolithic regulators are now available to meet a significant fraction of regulator requirements. Existing units, however, certainly do not exhaust the potential of monolithic regulation. Among the new devices being readied to solve the designer's regulation problems is a negative-voltage regulator that offers all of the outstanding characteristics of the MC1561, but is designed for positive-ground systems. Another new regulator, in this case a recently announced product, is designed for remote-sensing and general-purpose applications. Called the MC1569R, this new monolithic device offers performance similar to that of the MC1561, but is optimized for applications where the fast transient response of the MC1561 is not required.

The monolithic voltage regulator is a relatively new component that has already come of age. Starting almost from the introduction of the earliest models, monolithic regulators have been fully competitive with discrete types in an already wide range of applications. Newer units under development will extend this range and make monolithic capabilities easier to use even in specialized applications.

It seems reasonable to state that the monolithic voltage regulator is the most significant addition to the power-supply designer's arsenal that has taken place in many, many years.

Don Kesner received the B.S.E.E. degree from Kansas State University and is scheduled to receive the M.S.E. degree from Arizona State University in 1970. He joined the Motorola Semiconductor Products Division in 1968 as a senior applications engineer in the linear integrated circuit applications section. Mr. Kesner previously worked in the radar lab at Motorola's Government Electronics Division, Scottsdale, Ariz., where his activities ranged from designing servo systems for antenna control to video and IF designs.

EHV lines in the Federal Republic of Germany

The considerations that have shaped the design of the towers, lines, and insulators used in the FRG have given that electric network an identity all its own

J. Jansen Rheinisch-Westfälisches Elektrizitätswerk

To supply the mounting demand for electricity in the Federal Republic of Germany, a network has been built that consists of 380- and 220-kV multiple bundles. Multitiered towers are used to support the conductors because available land is a limited, expensive commodity. As far back as the 1920s, 220kV, in the scope of an emerging interconnected system, became a necessity for covering great distances. Then followed the 380-kV lines in the 1950s—the first line section being energized in 1957.

Under the Federal Energy Act, utilities in the Federal Republic of Germany (F.R.G.) are obligated to deliver electricity as reliably and at as low a cost as possible. (For data on power supply in the F.R.G., see Table I and Fig. 1.[1])

By far, the largest portion of electric power generated in the F.R.G. is produced by the utilities (78 percent) and this share is expected to grow in the coming years. By comparison, of the other significant producers, industry, which generates electricity for its own needs, contributes 20 percent of the total, and the Federal Railway provides the remaining 2 percent. (Industry, incidentally, usually builds power stations only as an adjunct to its need for process steam.)

Of the energy delivered by the utilities

- 27 percent is generated by public utilities (federal, state, community)
- 70 percent is produced by public/investor-owned utilities (stock held by the public and individuals)
- 3 percent derives from private, investor-owned utilities

As in many other countries, fossil generation of electricity predominates in the F.R.G. And, as in many other countries, coal is the predominant fuel (oil being but rarely used). One cause for this situation is that prevailing law—scheduled to lapse, if not renewed, in 1971—for social reasons subsidizes the use of coal and taxes the use of oil.

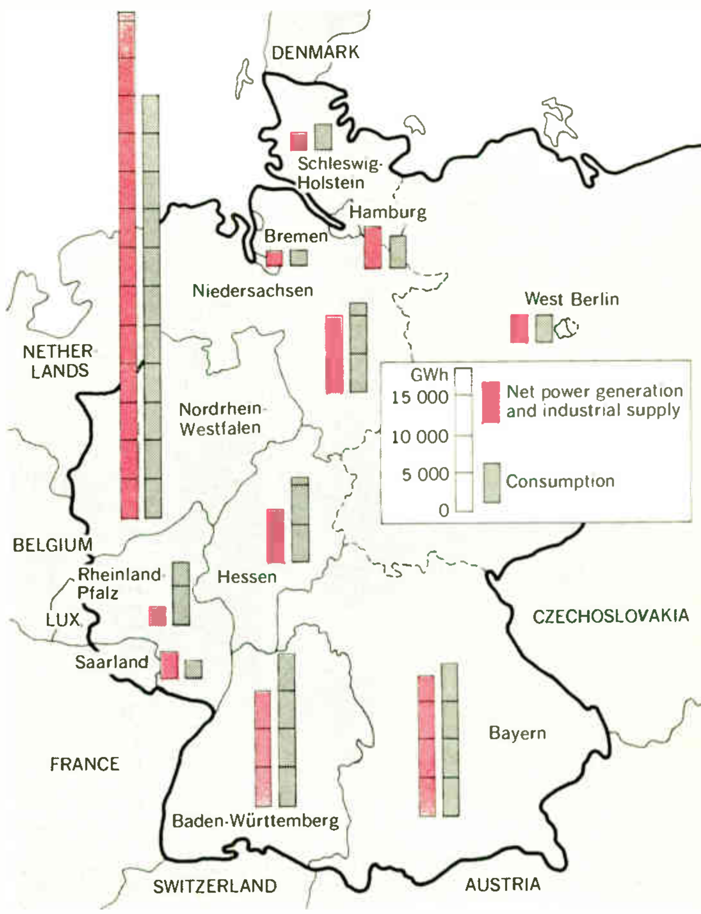
Lignite is extracted from open mines. Because of the

I. Power distribution

Characteristic	F.R.G.	U.S.A.
Total energy consumption, TWh	184	1330
Maximum load (winter), GW	30	228
Power proportion, summer/winter	0.83	1.05
Population, millions	60	198
Population density, per km²	245	25
Surface, km²	245 000	7 840 000

F.R.G. Characteristic	Average Value	Country	Large City
Consumption density, GWh/km²	0.7	0.3–0.8	4–8
Electric power density, MW/km²	120	55–320	1000–2000

FIGURE 1. Power generation and consumption within the individual states of the Federal Republic of Germany.



somewhat higher investment costs for the equipment to burn it, but substantially lower fuel costs as compared with hard coal or oil, lignite-fired stations must achieve a higher utilization factor. Such stations, therefore, are used for base-load generation.

Water-power resources in the F.R.G. have been harnessed extensively. There are, of course, two types of hydroelectric plants: the pump-reservoir variety and the small-dam class. The first pump-reservoir power stations were constructed in the Federal Republic of Germany as long ago as the 1920s and they continue to be a necessity in the scheme of power potential because of their ability to meet demands during periods of peak loads. At present, the installed capacity of such pump-reservoir stations amounts to 2 TW (2×10^{12} watts). On the other hand, hydroelectric plants with small dam (and therefore small reserve) capacity, although still erected, are constructed in connection with river-regulation projects only.

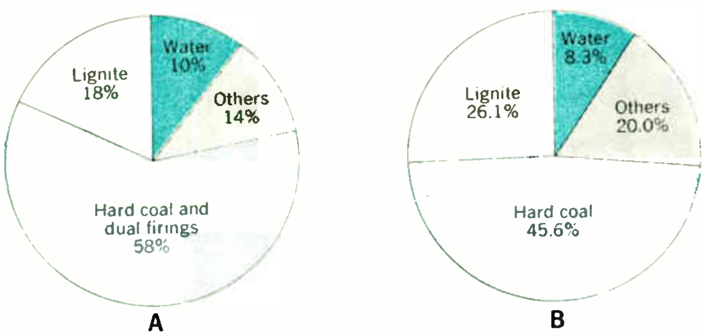


FIGURE 2. Parceling the electric-energy sources in the F.R.G. A—As a percentage of the installed capacity. B—As a percentage of the total generated in 1968.

FIGURE 3. Electricity generated (in TWh) and energy enchanged (in GWh) among UCPTE countries in 1967.

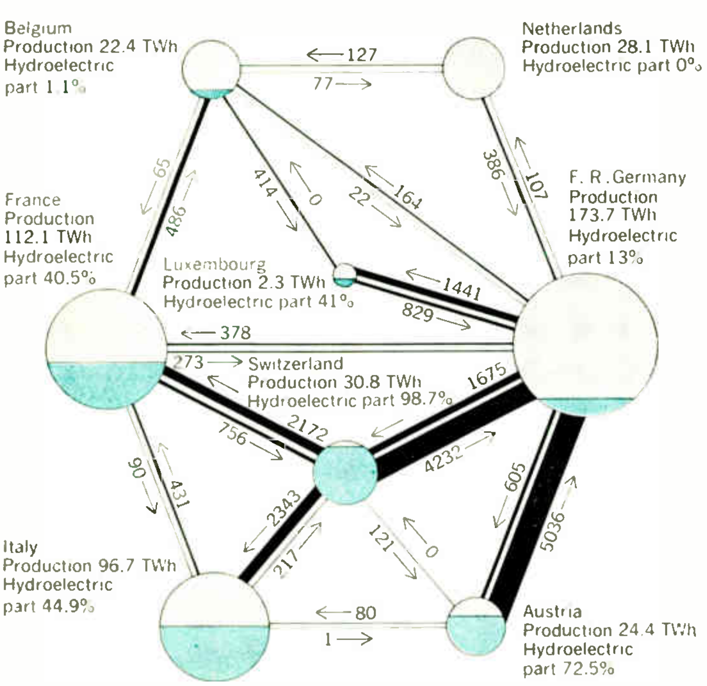


Figure 2 indicates the manner in which the various sources of energy contribute to the overall power-supply picture both as a percentage of the installed capacity within the Federal Republic of Germany and as a percentage of the total amount of electricity that was generated during 1968.

Conventional generating units now being built have ratings up to 350 MW. Nuclear power stations of 670 MW (unit rating) are under construction, and 1150-MW units are on order. The largest power station now operating is lignite-fired, has 15 generating units, and develops 2.3 TW, installed. Single-boiler units predominate; single-shaft machines are used to the exclusion of all other types.

It has been more than 40 years since interconnected operations commenced between power facilities in Germany. The first such link occurred between hydropower stations in the Alps and steam-power stations along the Rhine and Ruhr via a 220-kV line. Today, nine large utilities have joined into a cooperative called Deutsche Verbundgesellschaft (headquartered in Heidelberg). The nine companies mutually share technological developments and do the coordination work on mutual economic and operational interests. Participating companies with adjoining territories also work out power-sharing plans.

The interconnected system of the F.R.G. fulfills the the following functions:

- Interconnection of thermal- and water-power plants.
- Collection of the output of power stations that are locally bound to energy sources.
- Reduction of the need for reserve capacity within any one region.
- Power exchange within and across frontiers.

Favored by a central geographical position, the interconnected utilities have an active power exchange with those in other Western European countries—and together form the UCPTE (Union pour la Coordination de la Production et du Transport de l'Electricité). The magnitude of the systems in parallel and the power transfers across the frontiers are shown in Fig. 3. The whole interconnected group aggregates about 100 000 MW of electric power.

The transfer capability of the tie lines is high enough to ensure a high degree of reliability for the participants. The power exchange between the Western European countries is free from public regulations and approvals, partly as result of bilateral agreements and partly owing to the efforts of OECD (Organisation Européenne de Coopération Economique).

At present, the interconnected system within the F.R.G. consists of about 13 500 circuit-kilometers, composed of 380- and 220-kV lines (Fig. 4).

Conductors

For economic reasons, the hollow copper conductors (25- and 42-mm diameter) initially used for 220-kV lines—and which are partly still in service—were replaced by aluminum-cable-steel-reinforced (ACSR) conductors (although at first only for single conductors). In the Federal Republic of Germany, the only type of line now installed is ACSR (the most commonly used types being standardized in DIN 48 204).

Now, too (and for the past 15 or so years), only twin bundles are used for 220-kV lines—this as a by-product of work on 380-kV cabling. Until quite recently, subconductors almost exclusively consisted of 240/40 ACSR conductor having a diameter of 21.9 mm and a thermal current rating of I_D = 645 amperes.

With the trend toward a higher aluminum-to-steel ratio, some modifications were made, however. The changes consisted of switching to 230/30 conductor (21.0-mm diameter) with I_D = 630 amperes and 265/35 conductor (22.4-mm diameter) with I_D = 680 amperes for subconductors. The original reason for choosing a subconductor with a diameter of 22 mm was, among other factors, determined by a resolve to standardize—since such cable had already been in use for several years in the 110-kV network as a single conductor. As a matter of fact, experience has proved that the transmission capacity of 120–130 MVA per circuit obtained with these conductors is sufficient for almost all supply contingencies.

The selection of a twin bundle has proved to be successful for 220-kV lines. The transmission capacity (about 500 MVA) rendered by this cable is a necessity in many places. If this capacity were to be transmitted over single conductors, an aluminum cross section of approximately 900 mm² would have to be selected. This would on the basis of Verband Deutscher Elektrotechniker (VDE) Specifications 0210 entail substantially higher cost for erecting the towers in addition to the higher price of the cabling itself. Moreover, the twin bundle also solves the problem of corona discharge that occurs increasingly in densely populated areas. (Note that it is not so much the effect upon high-frequency devices but rather the discharge noise in the vicinity of existing single-conductor lines that brings about complaints.)

Technical discussions concerning 380-kV lines took place as recently as the late 1940s. Both a triple and a quad bundle were considered but the decision ultimately was made in favor of the quad bundle since it provides the most economical way of achieving a low surface-voltage gradient.

The decision was also prompted by another factor: The towers designed ultimately to carry 380-kV cable were to operate economically at 220 kV for an initial period of several years using the twin bundle as initial equipment. At the time of conversion to 380 kV, two subconductors of the same type could be added and the insulation level increased. The good performance of this cabling during 12 years of operation at 380 kV has completely fulfilled expectations.

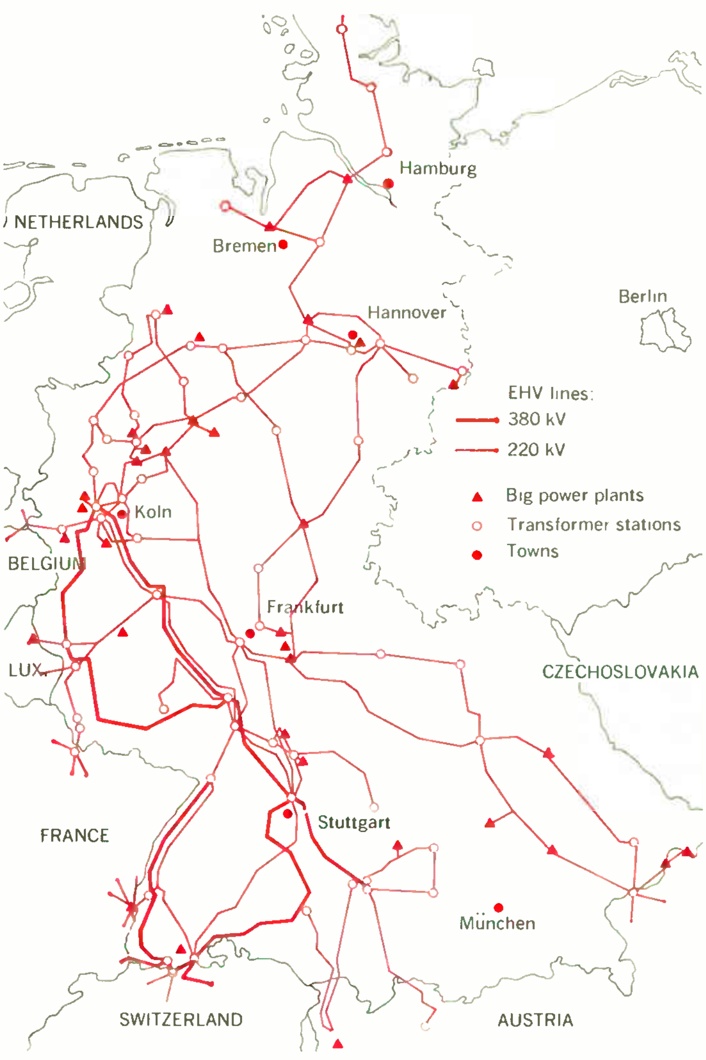
In spite of these excellent results, there were endeavors to decrease the costs for erecting 380-kV lines by designing them as twin bundles in some regions of the country where such lines rarely approach, cross, or span residential areas and where higher interference can be tolerated.

The 560/50 ACSR conductor preferred for twin-bundle lines has a diameter of 32.3 mm and a thermal-current rating of I_D = 1040 amperes. In view of the lower current rating, as compared with a quad bundle, much will depend upon the distribution and size of future power plants to determine whether the transmission capacity technically possible with this twin bundle is sufficient—the thermal capacity for 380 kV amounts to 1400 MVA as against 1800 MVA for the quad bundle—and can also meet economic requirements. (Apart from this, some use is still made of a triple bundle with 380/50 ACSR subconductors having a diameter of 27 mm and an I_D = 820 amperes for 380 kV.)

In conformity with the different cable cross sections and because of economic considerations, tensile stresses on conductors have also been revised. On the basis of a span of 400 meters, the design for everyday stress (EDS) amounts to approximately 16 percent of the ultimate stress for a quad bundle with 240/40 or 265/35 ACSR cables as subconductors, and approximately 22 percent for twin and triple bundles with 560/50 and 380/50 ACSR cables as subconductors.

The relatively low EDS of 16 percent is not, by any means (or standard), required. As a matter of fact, the new VDE Specifications 0210/5.69 permit more liberal stresses. Thus, bundled conductors whose subconductors have a diameter of up to 25 mm and an aluminum-to-steel ratio of 7.7, strung without dampers, are allowed an EDS value of 5.5 kg/mm². (This corresponds to an EDS of 20.7 percent for a 265/35 conductor.) It is rather an economic consideration with regard to tower-erection

FIGURE 4. EHV system of the F.R.G.—220- and 380-kV lines. There are approximately 13 500 circuit-kilometers of these high-voltage lines now in the country.



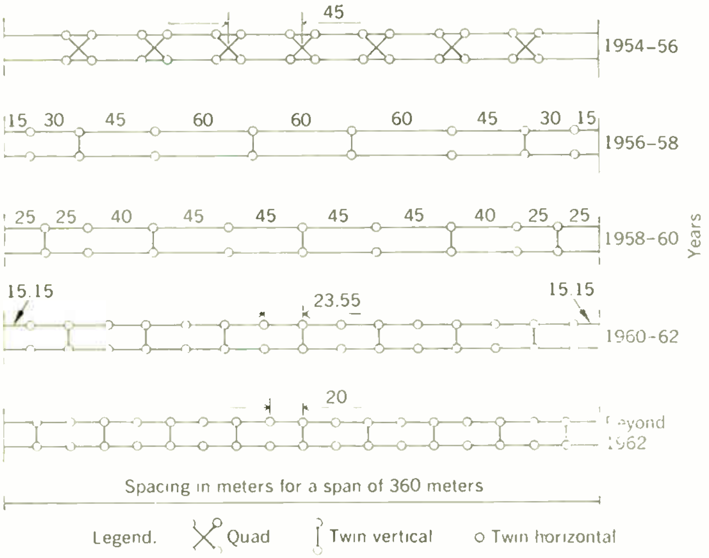
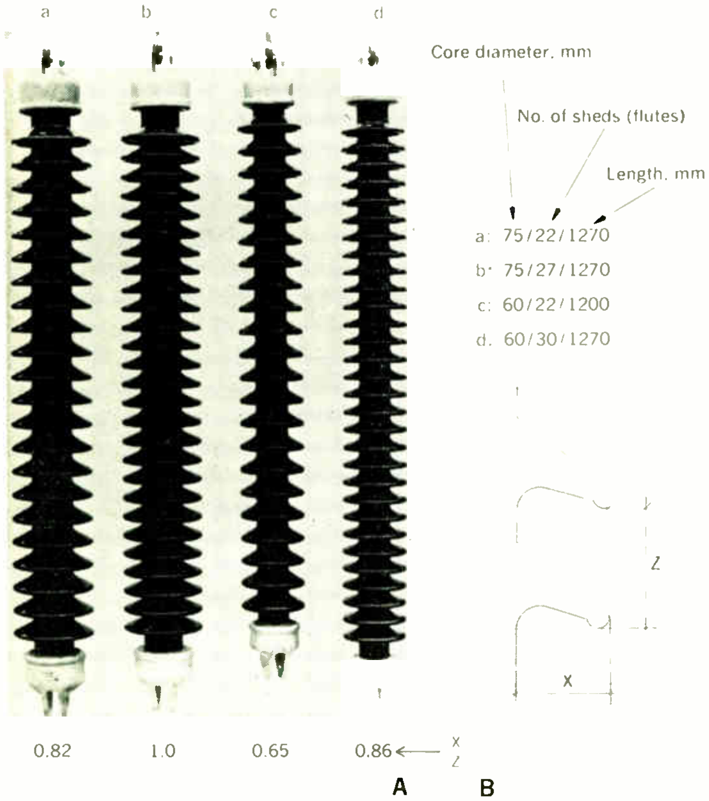


FIGURE 5. Revised intervals (in meters) for spacing arrangements for the 380-kV quad bundle. Since the spacing has been set, few problems have been encountered. The same spacer arrangement is now also used for the twin bundle.

FIGURE 6. The most commonly used types of long-rod insulators employed in the F.R.G. The various designs are optimal for different ambients. X/Z shed ratio is set above 0.85 whenever the air is heavily polluted. Long-rod design is preferred in West Germany because it has an exceptional service life.



costs that leads to low-tensile-stress designs.

In more densely populated areas in which the quad bundle is also generally used, the ratio between the number of suspension towers and angle (tension) towers is so low that an increase of the tensile stress would bring about a cost increase for the structurally more substantial towers. (Additional discussion appears in the section "Towers.")

As a pleasant side benefit, with a low EDS there are no problems with conductor vibrations. On the other hand, more recent measurements have indicated that, even under high cable tensions, vibration dampers become unnecessary when the suspension clamps are built in the form of a three-point device.[2]

It is not only necessary to design a proper clamp for tying the lines to the tower but, in the case of the line bundle, it is also important to keep the lines suitably separated.

The early spacers for the quad bundle were X-shaped and designed to provide considerable joint articulation. However, it soon became apparent that these separators were experiencing a high rate of failure. Mainly, these failures occurred at the articulated joint, which was designed in the form of a ball-joint socket arrangement. One of the primary reasons for this failure was the movement between the steel bolt and the aluminum clamp lid, and the aluminum part eventually disintegrated through erosion. As a result, the X-shaped spacers were replaced by horizontal and vertical twin spacers alternately installed along the line. To get more movability, these spacers were connected with spiral springs rather than rigid rods. The horizontal connecting springs, however, failed.

The latest solution provides for spacers with clamps fixed on a rigid strap by a strong pin joint. This joint permits movement only in the direction of the longitudinal axis of the conductor.

The satisfactory performance of these spacers is primarily due to the fact that the distance between spacers has been reduced over the years. Figure 5 shows the change in this spacer arrangement as it was made, for example, on the network of the Rheinisch-Westfälisches Elektrizitätswerk AG between 1956 and 1962. Since that time, there generally has been no problem with separator failure in the Federal Republic of Germany. For this reason, the same separators are also used with two-conductor bundles at 40-mm intervals.

Insulation

The good experience gathered with short, solid-core insulators—introduced in Germany during the 1930s after a rather large number of failures of cap-and-pin-type insulators—contributed to their continued use after World War II. Also, improved manufacture enabled the length of these insulators, which had been about 30 cm, to be increased so that by the early 1950s, rod insulators with an overall length of approximately 1.25 meters could be standardized.

The mechanical performance of these long-rod insulators underwent a continuous increase as manufacturing processes were updated and procelain compounds were improved. For example, the average values for mechanical failing load listed in the specifications—12 tonnes for the insulator with a core diameter of 75 mm, and 15 tonnes for the insulator with a core diameter of 85 mm—are considerably exceeded. Quality scatter is mainly marked by minor variations of the values of the mechanical failing load.

Reliable manufacturing methods have recently permitted the production of a 1.25-meter long-rod insulator with a core diameter of only 60 mm. (The standards are presently under revision and will take this manufacturing technology into account.)

It is now conceivable that this insulator might supply most of the demand for long-rod insulators in Germany although, no doubt, those with core diameters of 75 mm and 85 mm will continue to be used. (The 75-mm core will be needed to withstand the mechanical stresses within the 380-kV twin bundle and some 220-kV cabling. However, when two insulators are used on a dead-end tower supporting a 380-kV quad bundle—triple dead-ending increasingly is being abandoned—a core diameter of 85 mm must be retained.)

Several types of long-rod insulators have been developed to work in different ambients. The types most commonly used today are shown in Fig. 6. As a rule they are used either singly or coupled in chains of two and three for use at voltages of 110, 220, and 380 kV, respectively.

In contaminated-air areas the creepage (total flute length) distance is either increased by using a combination of several long-rod insulators having a smaller overall length than those shown in Fig. 6 or, as is sometimes done, by enlarging the overhang.[3]

An excellent performance record has been compiled by long-rod insulators. Their behavior in a dirty atmosphere has indicated that failures due to flashovers can be reduced considerably with their use.

There are other reasons for the almost complete dominance of the long-rod insulator. The installation of these insulators has nearly eliminated the maintenance that is necessary for cap-and-pin types. (However, pin-type in sulators, for economic reasons, have not yet been totally replaced with the long-rod type.) In order to ensure the electrical integrity of the cap-and-pin-type porcelain insulator, it is customary in Germany to make yearly inspections of the energized line during good weather, as shown in Fig. 7. The voltage at the individual insulator is sampled and the quality of insulation is evaluated according to the intensity of the discharge.

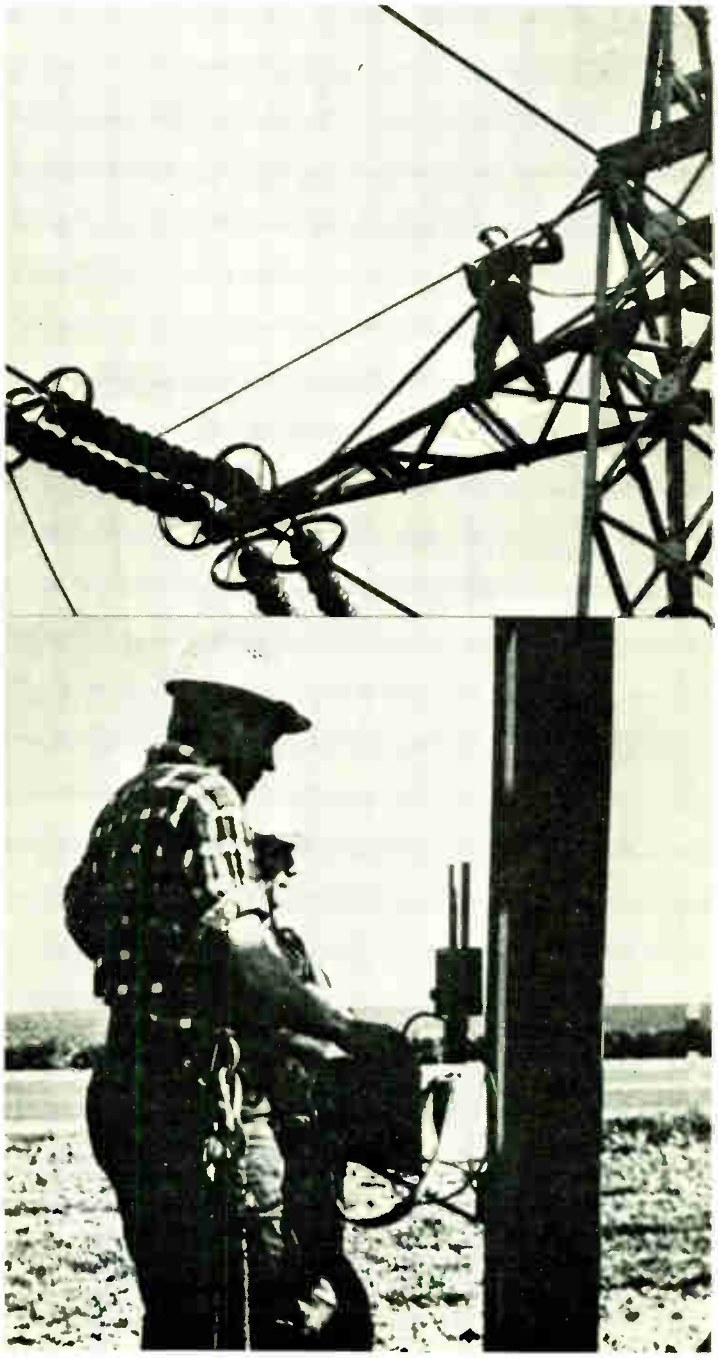
If necessary, the insulator can be replaced during periods permitting disconnection. (Note that work on energized lines, as done in the United States, is not allowed in Germany for lines above 1 kV. This may be attributed to the fact that in Germany, starting with 110 kV, when working on towers with two or four circuits, the danger of contact is considerably greater than for the single-circuit, one-crossarm, towers that are prevalent in the U.S.A.)

To this point, no mention has been made of glass—with reason: Glass insulators are rarely used in Germany. It is true that a cap-and-pin-type glass insulator eliminates the need, and thereby reduces the cost, for electrical checks (any electrical breakdown being physically observable). But experience has shown that glass insulators undergo too high a rate of shed breakage so that the work involved in replacing them proves more costly than the savings made on the purchase price of the glass insulators.

In addition, the long-rod insulator is as satisfactory as the cap-and-pin type from a mechanical-safety point of view. Experience gathered during the past years has shown that, with a total number of about 1.8 million long-rod insulators now installed, not more than one insulator out of 300 000 annually fails due to breakage. And qualitative improvements of the porcelain compound during recent years as well as tightened restrictions in testing for defective pieces are expected to reduce this fault ratio.

The long-rod insulators also are electrically nonpunc-

FIGURE 7. Hot-stick testing of cap-and-pin-type insulators. Whereas in the U.S.A., work on lines is accomplished without cutting the power through them, this practice is forbidden in Germany because there is more danger that lines will be inadvertently contacted, due to the way they are strung from the towers.



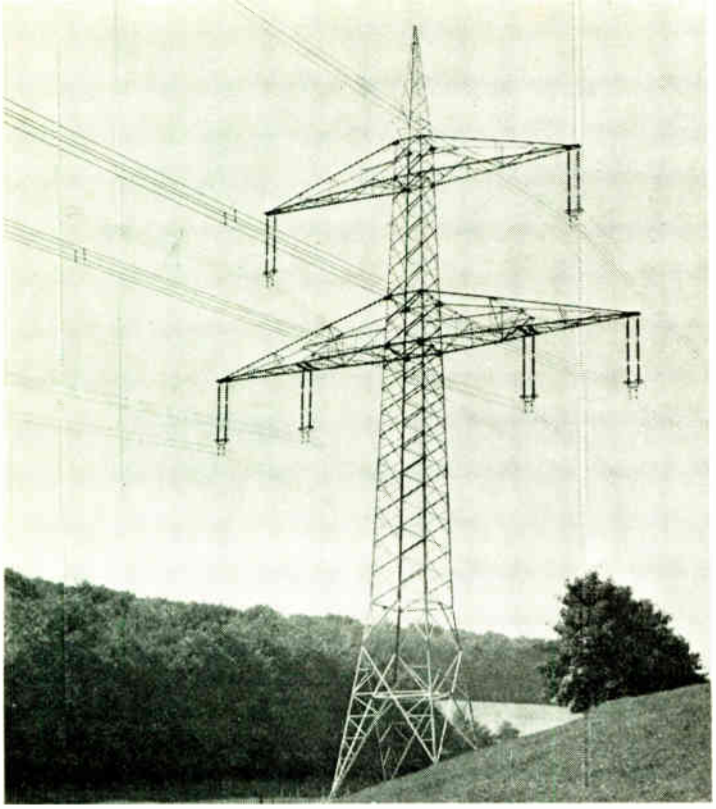
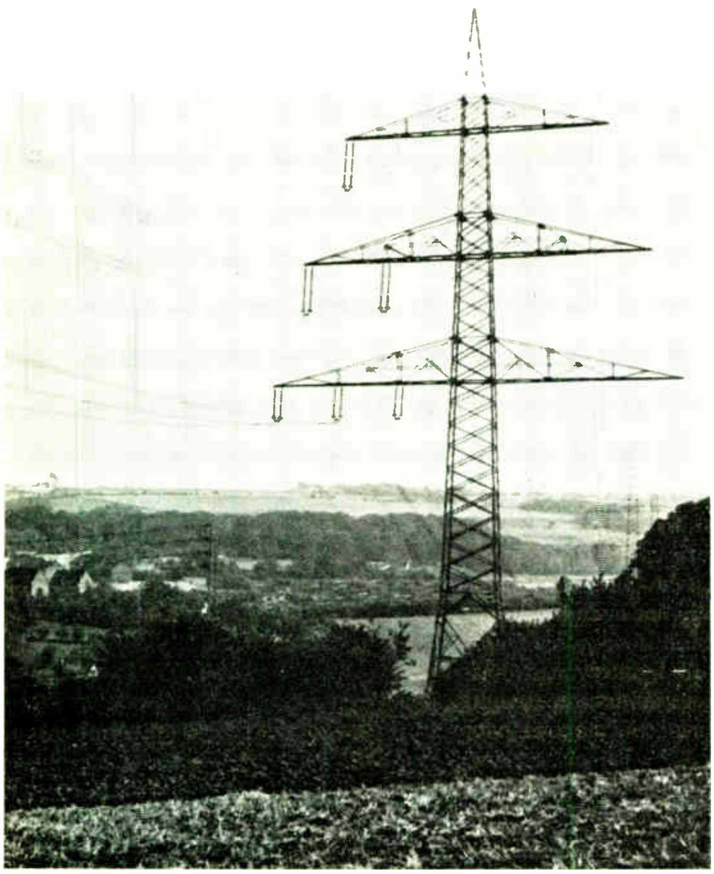


FIGURE 8. A 380-kV double-circuit line.

FIGURE 9. A 220/380-kV four-circuit line.



turable and eliminate all sources of radio and television interference, even temporary ones.

Although porcelain is satisfactory, it is probable that in the near future plastics will be used extensively as insulating materials for new EHV lines. The experience gathered in test plants in the U.S.A.—and, lately, also in Europe—using long-rod insulators made of glass-reinforced cycloaliphatic epoxy resin with overhang (shed) made of silicone-based plastics has advanced the technology to a stage where plastics are ready for trial under operating conditions.

With regard to the economical aspects of the use of plastic, it appears that, even now, it can be used for voltages of 110 kV, and more, provided that production is sufficient. In this connection, it is important to note that plastics make it possible to use a single, long insulator rather than the two (for 220-kV lines) or the three porcelain units (for 380-kV lines) now used—eliminating the need for, and cost of, connecting fittings.

Towers

The dense population in West Germany has limited the electric line routes, and two-circuit towers have always been used for EHV lines. The most commonly used type of tower has two crossarms with double triangular arrangement of the two circuits. This type—shown in Fig. 8—is generally called the "Donau" tower. Apart from the optimum symmetry of the electric field provided by this tower, the weight of this tower, on the basis of the VDE Specifications 0210, is lower than that for a comparable vertical-double tower. Moreover, observations of 220-kV lines have revealed less tripping with the Donau as opposed to the vertical-double tower from touching conductors when there is a tendency for conductors to "gallop."

On the other hand, it is not always possible—despite these clear advantages of the Donau tower—to avoid using the vertical-double tower when there is a lack of line routes. In such cases, compensation for premises crossed can be relatively high and the higher costs for erecting the vertical-double variety are offset by the economies provided by reduced right-of-way width.

Moreover, in view of the shrinking land area, the demand by authorities and property owners to make optimum use of the right-of-way puts pressure on the utilities to erect towers that are designed to carry several circuits—as a rule four. The electric utilities now must build towers for four 220-kV, two 220 plus two 380-kV, and even four 380-kV lines in addition to the more established towers that suspend four 110-kV or two 110 plus two 220-kV lines. As an example, Fig. 9 shows a 220/380-kV four-circuit line, Fig. 10a shows a 380-kV, four-circuit line (presently equipped, however, with two 220-kV and two 380-kV circuits), and Fig. 10b is a drawing of a suspension tower. In the Rheinisch-Westfälisches Elektrizitätswerk AG system, quadruple towers designed for circuits of 220 kV and more, carried 11 percent of the total length of the lines carrying more than 220 kV by the end of 1964 and 17 percent by the end of 1968. As a matter of fact, 80 percent of the 400 km of 220-kV and 380-kV lines built between 1964 and 1968 were suspended from quad towers.

It should be noted that the power availability of such four-circuit lines does not correspond to that of two double-circuit lines for system operation since mainten-

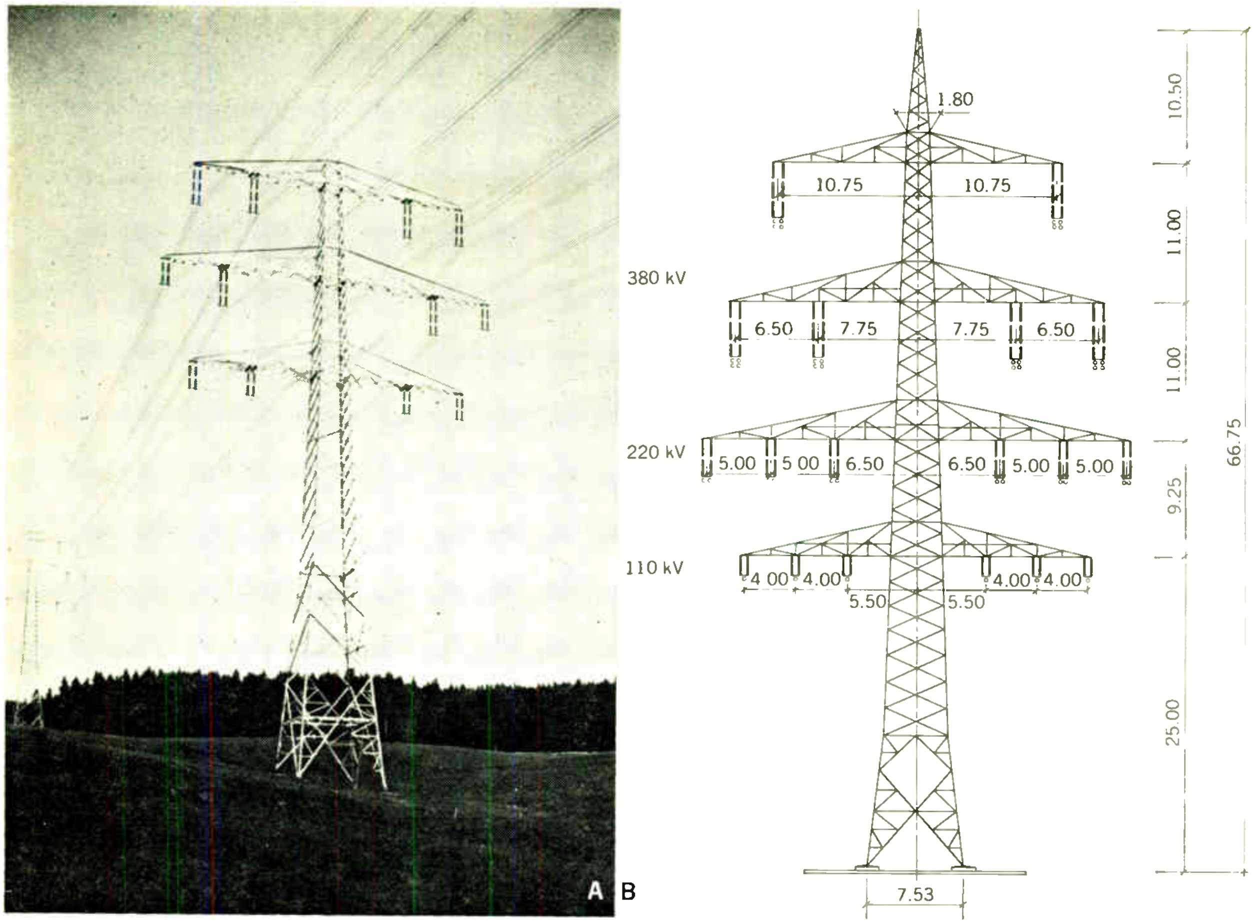


FIGURE 10. Four-circuit lines. A—380-kV four-circuit line. B—Schematic indicating dimensions in meters. Of the type of towers that are used in West Germany, the so-called Donau tower is said to provide several favorable features. Apart from the optimum electric field, the weight of the tower is appreciably less than that of a comparable vertical-double tower. In the Federal Republic of Germany, for the period of 1964 through 1968, 400 km of 220- and 380-kV lines were installed; and of this total, 80 percent were suspended from quad towers.

ance on one circuit generally requires the second circuit on the same side to be deenergized.

In addition, four-circuit lines, although they must be paid for at the time they are installed, often are not used to design (hence, efficient economic) capacity for several years. (The ultimate total cost for two double circuits would not be considerably higher than a four-circuit-line tower installation.) As a rule, the initial capital outlay for tower construction overrides all other cost-saving factors if the towers are not to be used at design capacity within the first four years of construction.

Whatever the disadvantages of the multiple-line towers might be, it is a fact that only by using them is it possible to postpone a considerably greater disadvantageous alternative—buried cable—until solutions are found that make such a substitution technically and economically (more or less) reasonable. The biggest obstacle, of course, is the problem of heat disposal. In addition, if lines are to be buried, more, not less, right-of-way width is required for their installation than for the equivalent lines suspended from towers. However, some short stretches (of the order of a few hundred meters) of 220-kV lines have been laid underground in connection with industrial installations.

Conventional construction and deployment of towers can be as uneconomical as the use of buried cable. For example, towers having two rows of six circuits require only about half the right-of-way land area needed for equivalent power capacity carried by conventional towers.

The towers of the 380-kV lines, on principle, are not designed for accepting higher voltage. There are several reasons for this: Even if higher voltages eventually are adopted, the 380-kV level would still have a place in the overall system. Overdesigning the towers for an eventual step-up would require considerable investment, returns from which would not be forthcoming for many, many years, and then newer technology may outmode the value of such conversion. Lastly, distribution needs can change. Therefore, although the need for higher voltages in the Federal Republic of Germany may arise in perhaps 20 years, anticipating the need at this time is not economically sound.[4]

As shown in the illustrations, the towers for the 380-

kV lines are generally equipped with one ground wire only. In the case of the quad bundle (used mainly on 380-kV installations), the ground wire is designed in the same way as the cables of the subconductors. The angle of protection (i.e., angle between ground and live wire) achieved amounts to 35 degrees for the Donau suspension tower as opposed to the 40 degree angle utilized with the so-called "dead-end" tower.

As previously mentioned (see "Conductors"), the geographical and structural conditions of the routes are decisive in the optimum choice of towers. Due to bottlenecks in densely populated areas, the distances at which routing direction must be altered are shorter than in the open country. As a rule, this implies a reduction of the span length (which in the Federal Republic of Germany is 360–380 meters on an average.) Accordingly, it is possible to calculate, for such a span length, the needed number of each type of tower. On an average, this works out to be 2.0 suspension towers and 0.7 tension tower per kilometer. Such a distribution implies a ratio of suspension towers to tension towers of 2.85:1—a not too favorable proportion.

There is the possibility of increasing this ratio, but only in certain parts of the country. If the terrain is open country, the number of tension (or angle) towers can be held relatively low—at 0.4 per kilometer—whereas the number of suspension towers can be maintained at the 2.0 per kilometer previously cited. In this way, five suspension towers generally can be installed for every tension tower erected.

To select the most economical type of foundation, test borings of the soil are carried out at all tower locations. The new VDE Specifications have determined a method for computing the dimensions of the foundation. According to these specifications, both driven- and bored-pile footings as well as the conventional separate concrete foundation are used. Building-permit restrictions allow for relatively small dimensions with regard to the footing spread. For example, the distance between the main legs of a normal 380-kV Donau suspension tower at the ground outlet is about 7 meters.

According to the new VDE Specification 0210, anchored towers are also permissible. It remains to be seen, however, to what extent a conceptual cost decrease may actually be achieved. The guy wires—because they must lie relatively far apart for small-based towers—may give rise to higher land compensation costs and thus to a shifting of the economical conditions.

St 52 steel constitutes approximately 25 percent of the weight of a 380-kV suspension tower and about 55 percent for dead-end towers; the residual proportion consists of St 37. The girder shapes used represent a ratio of leg width to leg thickness of as much as 1 to 15.

For some years, hot galvanizing of all tower components has been a common practice. Views concerning paint differ. Although some electric utilities have hitherto dispensed with paint after galvanization, others have applied either a single layer or a double layer of paint after about two or three years for this reason: Measurements of zinc-coating abrasion have been carried out on existing towers for several years.

Even in agricultural areas an abrasion of about 8μm per year is to be expected. Under such circumstances the zinc coating would be more or less totally abraded after 10–15 years. By painting the towers after about two years, an average endurance of about 12 years may be expected. After this period a relatively cheap paint can be applied, since the undamaged zinc coating requires no removal of rust. In regions having a rather corrosive atmosphere, the repainting interval is of the order of eight years.

Due to repainting costs, the economical use of Cor-Ten steel recently has been investigated. However, the number of shapes offered by the rolling mills is rather restricted, and there is, as yet, no proved economical justification for using this steel instead of galvanized material—not even in an industrial atmosphere where very severe corrosive effects are to be expected.

Approval of lines

On the basis of the Energiewirtschaftsgesetz (Federal Energy Act), before any lines can be erected, the utility must get the approval of the Department of Commerce of the individual state(s) that constitute(s) the Federal Republic of Germany. Ordinarily, a utility does not plan network expansions that affect more than one state at a time. The vote on the actual line routes is accomplished during the course of one or more hearings by the planning board(s) at the "county" level within the state.

This established, the planning board works with the "county" commissioner of expropriation along with the property owners tentatively affected by the expansion routing, and a final settlement is established. Should a property owner and utility not be able to agree upon equitable compensation, the commissioner intercedes as final arbiter. Rights granted to the utility permit it to erect, operate, and maintain the line for the duration of its existence.

The owners receive one type of compensation for the tower locations—in rural areas based on the loss of yields—whereas the right-of-way width is compensated for according to some different base. As a rule, compensations are agreed upon with farmers' associations on the basis of certain percentages of the market value (up to 20 percent) and the returns on the land, respectively. The distinction between agricultural areas and those used for building purposes is defined by the Bundesbaugesetz (Federal Building Act) that obliges the communities to establish plans for areas to be used for buildings. A special regulation governs compensation within a built-up area. It is becoming unavoidable not to span existing buildings or future buildings or to avoid future construction within the protected strip.

After 12 years' operating experience with 380-kV four-circuit lines in the F.R.G., it may be concluded that both the quad bundle and the long-rod insulator are excellent components.

REFERENCES

1. "Die Elektrizitätswirtschaft in der Bundesrepublik Deutschland im Jahre 1968," Statistischer Bericht des Referats Elektrizitätswirtschaft im Bundesministerium für Wirtschaft, Bonn.

2. Bueckner, W., Kerner, H., and Philipps, W., "Stresses in transmission line conductors near the suspension clamp," CIGRE, Rept. 23-07, Paris, 1968.

3. Baatz, H., Boll, G., Brennecke, D., Niehage, G., Reverey, G., and Vogelsange, Th., "New field experience with outdoor insulators in pollution areas and methods of assessing the performance of insulation under conditions of pollution," CIGRE, Rept. 212, Paris, 1964.

4. Lappe, F., "Entwicklungstendenzen im 380-kV-Bereich," *Energiewirtschaftliche Tagesfragen*, no. 10, pp. 392–394, 1967.

A plea for a proper balance of proprietary rights

Highly restrictive employment agreements involving ownership rights to inventions may well have had the unfortunate effect of stifling innovation in the United States

Robert H. Rines Rines and Rines

Blanket agreements between employers and employees, requiring the latter to relinquish in advance any claims to any ideas springing up during the term of employment, are considered illegal and against public policy in most highly industrialized countries—with the notable exceptions of the United States and Great Britain. In the author's opinion, these agreements are eminently unfair to the creative individual; moreover, since they tend to discourage innovation, they have a broad crippling effect on a nation's technological advancement.

Why does the United States Government demand of its development contractors many more proprietary rights than the law would deem it entitled to as a result of the actual inherent contractual relationship between contractor and government, in the absence of special written provisions abdicating those rights?

Why does a large segment of U.S. industry impose the same extralegal demands upon employed inventors and even those employees who by no stretch of the imagination are paid their salary for the purpose of inventing?

Why do many universities, supposedly dedicated to the advancement of knowledge and the free encouragement of new ideas, poll-parrot the government and industry and similarly demand such extralegal rights?

And why does Anglo-Saxon law (particularly the branch that has developed in the United States, which has historically championed the rights of the individual) sanction the pressures of employers, be they government, industry, or university, upon employees by blanket agreements forcing the employee to give up in advance any claims to ideas springing up during the term of employment—a concept that is considered absolutely illegal and against the public policy in most of the other highly industrialized countries of the world?

Why, indeed, are we undergoing such a "grab-all" policy in the greatest country known to man, with the greatest dreams of freedom for the individual and the greatest respect for the intellectual fruits of his mind, at a time when even the noncapitalistic socialist and communistic systems encourage and reward their employees who evolve inventions, even though the inventing is paid for by the state?

The answer to these questions is really quite simple; but we doubt if many persons have taken the time to reexamine how we in the United States came to this position and to requestion whether a continuation of this policy is truly in the best national interest and whether the policy is actually calculated to provide maximum incentive to invention and innovation either in contractual work for the government or in the course of university and industrial life.

First, let us examine how we arrived at this position—particularly when even U.S. Government laboratories, such as the Naval Research Laboratory, used to encourage their inventive engineers and scientists to seek commercial application for and reap financial reward from developments made in those organizations on behalf of the government. Without this freedom it is doubtful that a particular group of former professors and scientists, many of whom were not previously inventors, would have formed the nucleus of the Naval Research Laboratory. These men had earlier banded together at New London, Conn., during World War I to apply science to the solution of pressing military problems.

The legal obligations of the employed inventor

What we have to say here is equally applicable to the relationship between government and contractor, even though it is couched in the illustration of the specific problems of the employed inventor.

Although our Anglo-Saxon tradition includes a background of statutory law dealing with the master-and-servant relationship in industrial life, the rights and obligations of the employed inventor as such have never, to my knowledge, been the matter of statutory enactment, either in Great Britain or in the United States (save perhaps for modern rules and regulations, having the force and effect of statute, and related statutory provisions dealing with government employees and their obligations to the government). Instead, however, the law governing the rights and obligations of employer and employed inventor developed through court considerations based upon the inherent contractual relationship between the parties stemming from the master-and-servant legal tradition. If an employer were entitled to certain rights, these would have to be justified in terms of the actual relationship of the parties. Was the employee paid his salary to do the particular thing he did, and therefore has the employer the right to claim that particular thing as the property of the employer?

It's just that simple. If it can be said that the employee received his salary or other remuneration for doing the very thing he did—or, otherwise stated, that the employee's work product was within the scope of the employment relationship for which the employer made payment—then the fruit of the employee's work is the property of the employer; otherwise, not.

Also stemming from general principles of equity came the modification that there were circumstances wherein, though the employee's activities were not strictly within the scope of the employment, the employer made sufficient contributions to the employee's end product that the employer was entitled to a free personal right or license to use the same, even though not entitled to the ownership thereof—a so-called "shop right."

It is upon this Anglo-Saxon and American jurisprudence—which decrees that one is not entitled to what one hasn't actually paid for—that the law in the United States (in the absence of a specific contract provision vitiating the same) still holds fast to the following general maxims:

1. An employee who is not employed to invent owns all inventions that he may make, even if they are in the employer's field of operation.

2. An employee whose primary employment designation is for invention and who is specifically assigned the task of making the type of invention that results, has no claim upon that invention; it is, rather, the property of the employer.

3. The inventor-employee, however, who has no specific or implied orders to make that type of invention and who makes such an invention that is not of the nature or type actually contemplated by the employment, owns such invention.

4. In the case of items 1 and 3, if substantial facilities of the employer and/or the employer's personnel, and/or other similar significant contributions of the employer, were used in the making of these inventions, the employer may be entitled to the aforementioned free nonexclusive personal shop-right license to use the employee's invention.

5. If an inventor-employee makes an invention, unassigned as an employment duty and outside the field of activity of the employer, even though conceived on the employer's time, this is the property of the employee, again subject to a possible shop-right license for the employer in the event that the conditions of item 4 are met.

6. If an inventor-employee, following termination of his employment, makes an invention in the field of his former employer, this is the inventor's property.

It is rather interesting that when a graduating class of engineers at one of our leading engineering universities was polled, over 80 percent had been brainwashed by our society's norms of conduct to the extent that they felt that, as a matter of right, their employer owned all their inventions made during the term of employment. And, I daresay, some of the foregoing points may be startlingly new to many otherwise educated engineers who may read this article.

What has just been said applies not only to the private employer; it is equally applicable to the U.S. Government, which in our society is no more entitled to pick the brains and intellectual efforts of the individual than any other organization. Our courts have put overanxious government officials in their place when they have attempted such actions as, for example, demanding ownership of, as distinguished from a shop-right license to, a radio invention of radio engineers in the employ of the Bureau of Standards, which invention the engineers had licensed to a commercial corporation on their own behalf. The court found that although the inventors were instructed, as part of their employment, to engage in certain research projects and problems in the radio field, these had nothing to do with the particular type of radio invention in question; and that, though the inventions were developed in the laboratory of the government, the most that the government was entitled to was a shop-right license (*United States v. Dubilier Corp.* 49 F. 2d 306).

Refreshing language is found in other decisions; for example, in *Solomons v. United States* 137 U.S. 342, it was stated that

> "The government has no more power to appropriate a man's property invested in a patent than it has to take his property invested in real estate; nor does the mere fact that an inventor is at the time of his invention in the employ of the government transfer to it any title to, or interest in it."

Employment policies of the present

We witness, however, in each of governmental, industrial, and university relationships with engineering and scientific employees, a far different set of rules from those cited in the law, as enumerated in the preceding paragraphs. We choose to call these rules "extralegal," because they result in the relinquishing of rights, by companies to the government and by employees to employers, that the law holds are not actually paid for by the government or the employer, respectively, unless the contractor or employee—through coercion, ignorance, or surrender—chooses to give them away for nothing, just to have a contract or a job.

In short, the Government generally demands of its contractors all ideas and inventions conceived during the course of the contract—even those having nothing whatsoever to do with the actual subject matter of the contract and those not developed or worked on during the contract—and from all employees of the contractor, including those who are not paid their salary to conceive ideas or to invent.

And although U.S. industry may indignantly cry "foul play," it buckles under just the same in order to get a contract. Companies might do better by banding together to teach government officials a proper balance of morality and ethics, and to ensure a policy that is calculated to encourage incentive, as distinguished from the kind of hypocrisy, lack of effort, and even cheating that is today rampant in government-industry relationships, as later discussed.

Ironically, while claiming that it has been abused by government, industry largely imposes on its employees the same stifling and oppressive conditions that it objects to from the government. This is justified by a rationalization, later explained, which does not seem too convincing, particularly for the second-, third-, and perhaps fourth-generation management caretakers of once generally free-enterprise organizations—companies that are today prostituting their dignity, giving in to almost any conditions the government demands, cowering before the threats of ambitious Justice Department lawyers, and trembling with anticipation of imagined potential displeasure of the stockholders.

Perhaps most crass and unbefitting of all is the similar role played by many of our universities, which seem to think it is their function to get into the act of grabbing the advances made by their staff, because it is the procedure that their government has taught them to follow and because they have lost sight of the greater moral responsibility that should differentiate university from government and from a commercial enterprise.

A survey shows that a sizable number of the largest corporations in this country extend this grab-all policy even to a secretary, a technician, or a janitor—and even in fields outside the scope of present active development programs of those companies. Many U.S. corporations at least have the decency to restrict their contract terms to engineers—however, usually not only to engineers employed to invent but also to those not in any way associated with actual engineering and related functions. Bear in mind, moreover, that a limitation in any of these agreements to inventions in the field of interest to the company, particularly in the case of large corporations, is no limitation or restriction at all. Some of these contract provisions even attempt to extend for a period after the termination of employment. I use the word "attempt" advisedly in the light of the way in which I believe modern-day courts will treat such provisions. I merely leave with you the interesting thought that, under some of these clauses, an invention actually made for a second employer after the termination of an agreement with a first employer could allegedly be the property of the first employer.

U.S. Government employment contracts

I think we can attribute the policy of going after "extra-legal" rights to the efforts of the U.S. Army Signal Corps, during World War II, to try to prevent possible subjecting of the government to royalty payments on inventions, the making of which the government largely sponsored. As each new bushy-tailed government attorney took over and each agency and department of government developed and expanded its own patent sections, the bureaucratic pace quickened to the point of the present contractual requirements for the "privilege" of doing business with the government.

I know as a matter of fact that, in a number of cases, corporation lawyers have just blindly followed what the government made the corporations do in their contracts and adopted what I consider to be hoggish practices as corporate policy.[1] The spillover found its way into universities, which seemed (and unfortunately in many cases still seem) to be more interested in the potential dollar than in maintaining their traditional and once-respected role of intellectual freedom and encouragement of exploration of knowledge. Again, sizable government influence has played its part in the universities, which have become heavily dependent upon the government for support.

Apart from this historical trend to government, I think the management, or at least the counsel, of many industrial companies genuinely feels that the employed inventor gives up a measure of individual rights for the relative security of employment (though I must comment that this appears to be an ill-befitting argument from those companies that deal so largely with government business and that, as contracts are not renewed or are lost, release engineers in droves on a moment's notice).

In my judgment, corporate employers do, however, have worries and justifiable concern. Perhaps the common law does not go far enough in protecting their interests in cases in which the fact of employment itself and exposure to the needs and problems of industry are seized upon by an employee who, though not strictly employed to invent, can perhaps reap an unjust benefit for himself or create an embarrassing competitive conflict situation under circumstances resulting from the relationships in items 1 and 3 listed earlier. There are circumstances in which a shop-right license doesn't seem enough protection for the employer, particularly if the employee can peddle the invention to a competitor and be put in a position of somewhat conflicting interests and loyalties in serving both his employer and his own interests as an inventor with a competitor.

However, is the limit of intellectual capability and social invention among U.S. businessmen, policy makers, and lawyers for a solution to this problem restricted to the kind of grab-all policy that has been evolved? Are we really that sterile?

The West German law, for example, while involving basically the same rights in the employed inventor, has found a different solution. Instead of emasculating the inventor by taking all rights in advance as a condition of employment, the West German law (see Table I)* puts a requirement on the inventor who had made an invention for which the employer did not really pay. It requires that he must promptly submit this invention to the employer, giving the employer the first right to purchase same with protection against unreasonable demands.

Is it not ironical that Nazi Germany, in its efforts to stimulate invention for its own selfish aims, became the father of what should have been an American concept, almost literally called for by the spirit of Article I, Section 8, of the U.S. Constitution that purported to "*secure*" to "inventors" the rights to their discoveries?

Is it not further irony that the democratic United States has been upstaged by similar social inventions recognizing the dignity of the individual by socialist Sweden?

Nor should we think that these West German or Swedish solutions are the only possible social inventions. The problem is that in the United States some of us have stopped thinking of social invention and even have stopped thinking about modernizing our free-enterprise concepts to suit them to the latter part of the 20th century,

* Table I is a generalization of the law in different lands; and, as in the case of many generalizations, is fraught with possible misinterpretations. Although a twist of the facts one way or another may throw the conclusion in another direction, this information is believed helpful nonetheless.

I. A generalization of laws governing employed inventors (in the absence of special contractual arrangements to the contrary)

	U.S.A.	Sweden	West Germany	Netherlands	Great Britain	Japan	Canada
What kind of law and who is covered?	Court-interpreted law. Applies to any inventor employed by anyone	Statutory (1949). All inventors except university or other teachers	Statutory (1957). All inventors	Statutory (1957). All inventors[a]	Basically court law stemming from master–servant and contract law[b]	Part of Patent Statutory Law (April 1, 1960)	Master–servant common law[c]
If invention is primary task of employment (1) If there are specific orders to make that type of invention (2) If there are no specific orders or contemplation for that type of invention	(1) Employer ownership. No payments other than salary (2) Inventor or employer[e]	So-called "service" invention. Employer ownership. Inventor must disclose[f]	So-called "service" invention. Employer ownership. Inventor must disclose[g]	Must supplement wage[d] (1) Employer ownership (2) Inventor ownership	Employer ownership. Under (2), the courts seem to raise up a "trustee" doctrine that the employee hold the invention in trust for the employer	Employer may acquire patent for compensation additional to salary[h]	Appears slightly closer to analysis under Great Britain than U.S.A. Paucity of definitive modern court law
If invention is in field of activity of employer (1) If there is some relation to employment other than assignment to invent (such as special knowledge of need, use of facilities, etc.) (2) If there is no connection with employment	(1) Inventor ownership. Employer may have shop right (2) Inventor ownership	(1) Inventor ownership subject to shop right but must be disclosed to employer (2) Inventor ownership but must be disclosed to employer	Inventor ownership, but must submit and offer the invention to the employer	Inventor ownership	(1) Inventor ownership. Employer may have shop right (2) Inventor ownership	(1) Appears to be the same as in item above (2) Unknown; apparently undecided by court law[i]	
Invention outside field of activity of employer	Inventor ownership	Inventor ownership	Inventor ownership	Inventor ownership	Inventor ownership	Inventor ownership	
Invention made after employment, in field of former employment	Inventor ownership	Presumed employer ownership[j]	Inventor ownership	Inventor ownership	Inventor ownership	Inventor ownership	
Contract giving blanket title to all inventions in advance in return for salary alone	Legal	Illegal	Illegal	Illegal	Legal	Illegal	
Forum for settling disputes	The courts	Official state board and the courts	Patent office, arbitration board, and the courts	Patent office and the courts	The courts	The courts	

[a] Except state-administered institutions that have special regulations.
[b] Applicable to all except government servants.
[c] Although special statutes are applicable to government employees, it appears to be governed by some common-law principles.
[d] By special payment based on monetary importance of invention and circumstances under which it was invented.
[e] May be inventor's ownership subject to shop right, or employer's ownership, depending on how close it comes to being within expected duties.
[f] Must supplement salary with special compensation related to unusual commercial value of invention. Economic value, technical scope and circumstances, and significance of employment are factors in determining such compensation.
[g] Formula has been developed for compensation above salary with factors relating to economic usefulness, actual tasks and position of inventor with company, and share that company contributed to origin of invention.
[h] Reflecting profit to employer from invention and extent of employer's contribution to invention; otherwise employer has nonexclusive license as a result of the employment relationship.
[i] Belief of Japanese counsel is that this is inventor's ownership.
[j] If done within six months after termination of employment (but presumption is rebuttable by proof).

while preserving the traditions of individual dignity, individual initiative, and encouragement, and constantly guarding against the encroachment of bureaucratic and stifling governmental pressures—something that used to be as American as apple pie.

What we have spawned

We now have the opportunity to reflect upon what have been the effects of these policies.

Surveys conducted by my students in cooperation with the Academy of Applied Science[2] have gleaned the following facts:

1. About 85 percent of all employed engineers interviewed in a wide variety of companies (including companies chiefly in the electronics industry, some primarily in government work, others in commercial pursuits, and companies with mixtures of government and commercial activities) felt that the patent system held no particular meaning for them as individuals.

2. About 92 percent considered that there was no difference in reward from their employers for invention, as distinguished from good engineering.

3. About 45 percent of the engineers employed by companies doing work for the government felt that there was no sense in taking the risk of fostering radically new ideas because the government contracts would not give their employers sufficient patent advantage.

4. Some 84 percent of the engineers admitted that they were not enthusiastic and, although they had ideas that could be of benefit to their employers, there was no incentive to "fight city hall" and to embark on the risky and unpopular role and politics of fighting either to try to force adoption of significantly new concepts or to expand the scope of their employer's field of operations.

5. Those companies that had separate divisions working in somewhat similar fields—one for government and the other for commercial purposes—admitted the tendency of not developing significantly new ideas under the government contract, but either transferring them to the commercial operation or burying them. In one significant instance, the NASA communications operations of a large company produced four inventions in a five-year program; whereas the corresponding commercial communications division of that same company admitted to filing 30 to 50 patent applications on communications inventions per year in the same period.

The consistency of these figures, at least in the New England area, irrespective of whether the interviews were conducted in large, medium-sized, or small companies, or in government or university laboratories working under government sponsorship (where applicable), was startling.

The fact is that a grab-all policy, from government to industry to university, has taken much of the spark out of what could be a dynamic, challenging, and inventive technological society; and we must content ourselves with just the very few diehards who still fight the inventing battle no matter what the discouragement and the odds. I myself am shocked at the apathy in universities and some business organizations with which I am connected, on the part of an engineering community that has been subdued to the point where it just wants to do a good day's engineering for a good day's pay and not much more. I am disappointed that much of U.S. industry[3] is not amazed and bewildered that its rate and quality of innovation is so inadequate and that, despite the existence of costly research organizations, there is not the cornucopia flow of new products and ideas that managements give at least lip service to wanting to achieve.

Writers in numerous magazines today refer to the fact that the NASA search to try to prove spin-off inventions from space research has been swept under the rug. Hasn't Congress ever wondered why, with the billions spent by NASA and the military, there is such a paucity of commercial innovations stemming therefrom?

A suggested change of philosophy

I should like to suggest that government, industry, and universities in the United States stop trying to take what they haven't paid for.

Inventions come from individuals—not from governments, nor corporations, nor universities. Unless individuals are stimulated to take the inherently risky course of invention—often fighting the establishment, which almost by definition is not inventive and resists rocking the boat—there can be no invention.

Invention is a 24-hour-a-day job. It often requires a lifetime of background and experience for which neither government (in the case of contractors) nor employers (in the case of employees) has paid anything. The mind must find and get locked into the problem. Invention knows no office hours and no other bounds. It inherently is upsetting to organizations with missions and preconceived policies and programs, and it is mostly unpopular.

I seriously doubt, moreover, that significant invention, as distinguished from clever engineering, is really in the contemplation of most employers and engineers when employment contracts are signed, no matter what the contracts may say. And, to me, the requirement that, as a condition of employment, all the brainstorms that may result are the employer's property in advance, is neither a fair bargain, paid for by the salary and normal promotion procedures, nor an inducement to many to embark upon the risky and unpopular road of radical change that accompanies invention.

If those of us engaged in this innovative cycle will only be truthful with ourselves, we shall have to admit that the U.S. Government, in many instances, is not getting the creative effort it hopes to get; that in many other instances it is really getting shortchanged through normal practices and understandings between contracting officers and the contractor; and that, in still other instances, it is certainly not receiving the highest quality of development work by the most competent organizations or the most excited kind of enthusiastic development by employees.

And if we will be but true to ourselves, we must admit that most companies are not getting anything that remotely approaches the real potential of their engineering staffs.

Again, in a moment of truth, we must recognize the almost complete apathy in most universities that ought to be hotbeds of excitement for creative people.

The foregoing policies have not produced what any of us in the United States should be happy or content about; and it is my thesis that a change of attitude and of rewards could rekindle the inventive spark, spirit, and dignity of the employed inventor. Instead of being intimidated into signing away rights as a condition of employment, we should have access to the many rewards that can be keyed specifically to the inventive process. It is, indeed, a miracle that we do such magnificent things despite these discouraging attitudes; and we certainly outclass much of the rest of the world in what we do accomplish even though crippled by these unimaginative policies.

But the challenge is upon us. We are no longer beset by mere copyists, but by true innovators and significant inventions from abroad, playing key roles in a myriad of our industries ranging from steel to electronics. We no longer have a monopoly even upon the secrets of mass production. Just ask our people in the electronics industry about the threat to their inventive pace from the Far East alone.

Not only should we return the dignity of the employed inventor to him with safeguards against abuse, as discussed earlier, but we should give serious thought to the matter of compensating the employed inventor for his inventions in those circumstances where they very clearly are not compensated for by the circumstances of employment. As seen in Table I,* West Germany, Sweden, the Netherlands, and Japan have recognized that salary alone is not compensation for invention—a role that demands not merely a 9 A.M. to 5 P.M. dedication, but a spark, a persistence, and generally a fight against the establishment. Certainly in the case of inordinately important inventions, the work can hardly be said to have been within the original contemplation of the parties when the employment contract was signed. To give this away in advance seems shocking; and I am apparently in good company when able lawyers and teachers in at least the aforementioned countries have seen to it that any such technique is illegal, even though Representative Brown of California was earlier unable to muster support for similar legislation in the United States. Some of the factors that are taken into consideration in assessing additional compensation include whether this was a specific invention assigned as a project or whether it was an extra idea beyond the scope of employment, or whether it reflects significant contribution of the employer in the fulfillment of the invention—considerations calculated to be fair to both parties.

I should mention that some companies claim that it is often difficult to single out the inventor and reward him; and that there may be hard feelings among others in the company who perfect the invention and make it commercial. One also hears that invention is really rewarded by salary adjusted in response to performance.

My answer to these people is simple. First, if they are content with the level of invention and excitement among engineers today they should continue as in the past. If they are not (and I have yet to become intimate with an organization that should be satisfied, and this certainly includes the United States Government), then it is hard to accept these excuses. The patent laws define who the inventor is and lawyers have little trouble in filing patent applications in the inventor's name despite the fact that other noses may be out of joint. Even "project Hindsight" showed that the government could, in every case of the various military missile systems investigated, find the one, two, or three inventors.

* Study No. 30, 87th Congress, 2nd Session, "The Law of Employed Inventors in Europe," Study of the Subcommittee on Patents of the Senate Committee on the Judiciary.

Second, it has been pointed out that a very large percent of those inventors interviewed felt there was no distinction between the rewards they received for invention (if any) and those received for good engineering; thus, management is apparently in a dream world if it believes that it is getting its message across.

Conclusion

At the present time, Congress is considering a few changes to the patent laws. My own view is that these are insignificant and a face-saving effort in the light of the furor created by a group of "social-science innovators" known as the Presidential Commission on Patents.

Perhaps some useful purpose could be served by this bill, however, if sufficient numbers of the engineering, business, university, and legal communities found themselves in agreement with the premise of this article and were willing to appear before Congress in the impending future sessions on the patent bill, and/or to write to their Congressmen and Senators about the inequity and lack of adequate stimulation in a public policy that sanctions contracts providing, in advance, for blanket title to all inventions as a condition of employment. This might persuade the adoption of a new provision in our law that could render the patent bill a worthwhile contribution after all.

Who knows, this fight for the sanctity of individual and corporate rights might even spark Congress out of its present lethargy and rekindle some kind of an exciting and once-again leading spirit in the United States that can recapture the respect and admiration of the rest of the world.

REFERENCES

1. Rines, R. H., "Patents and the stimulation of innovation," *IEEE Trans. Engineering Management*, vol. EM-15, pp. 70–79, June 1968.
2. Rines, R. H., "Current problems in the protection of creative ideas and writings," *1963 IEEE Internat'l Conv. Record*, pt. 10, pp. 8–11.
3. Rines, R. H., "Create or perish," *Producers' Council Tech. Bull.*, vol. 3, p. 39, Mar. 1965.

Robert H. Rines (SM) received the B.S. degree in physics in 1942 from the Massachusetts Institute of Technology, and the L.L.B. and J.D. degrees in 1947 from Georgetown University Law School. He is a partner in Rines and Rines, Boston, Mass., specializing in patent, antitrust, and industrial property law. He is also a lecturer at M.I.T. on patents, inventions, and the development of innovative business organizations. He was formerly Gordon McKay Lecturer on patent law at Harvard University. He is a registered professional engineer in Massachusetts, and a member of the Bars of Washington, D.C., Massachusetts, U.S. Supreme Court, the Federal Communications Commission, the Tax Court, and the Federal Courts, and the U.S. and Canadian Patent Offices.

Mr. Rines is a member of Sigma Xi, RESA, the American and Boston Bar Associations, and the Patent Law Association. Formerly a member of the Commerce Technical Advisory Board, under Presidents Kennedy and Johnson, and of the National Inventors Council, he is now president of the Academy of Applied Sciences and a member of the National Academy of Engineering Patent Committee.

Transmission delay and echo suppression

When satellites are interposed as relay stations, long-distance telephone transmissions may carry problems as well as conversations

Richard G. Gould **Office of Telecommunications Management**
George K. Helder **Bell Telephone Laboratories**

We are all aware that radio communication is not instantaneous. With the coming of satellite communication we have fallen prey to some of the consequences of telephony with long propagation time. So it is that when one person speaks to another via satellite link, he must wait perhaps 600 ms for a rejoinder. Such a delay, of itself, may not be too disconcerting, but in combination with echo it can contribute to poorer quality transmission. This is an account of how echo is produced, how it can affect a telephone conversation, and what steps are being taken to overcome the problems.

> "*Shall we have the Thompsons over for dinner?*"
> "*Sure, dear.*"
> "*If you feel that way about it, I won't invite them!*"

This is said to be the substance of a telephone call that purportedly took place between a man and his wife during a simulated satellite-relayed conversation.

The story may be apocryphal, but it is indicative of the problems that many laboratory workers in the early 1960s felt would be caused by transmission delay. Fortunately, since that time, subjective tests and the successful operation of synchronous communications satellites for intercontinental calls have largely dispelled the fear that this problem would be widespread. Not all questions have been answered, however, and there is continuing study into the effects of delay and the associated problems of echo on satellite telephone connections.

Source of delay

"Geostationary" satellites orbit at an altitude of about 36 000 km. Consider an international telephone call from San Francisco to London via such a satellite. The speech must travel from one telephone through domestic long-distance facilities to the earth station at Andover, Me., then up to the satellite, down to the distant earth station at Goonhilly Downs in Cornwall, England, and again through long-distance facilities to the telephone of the party in London (Fig. 1). The distance between earth

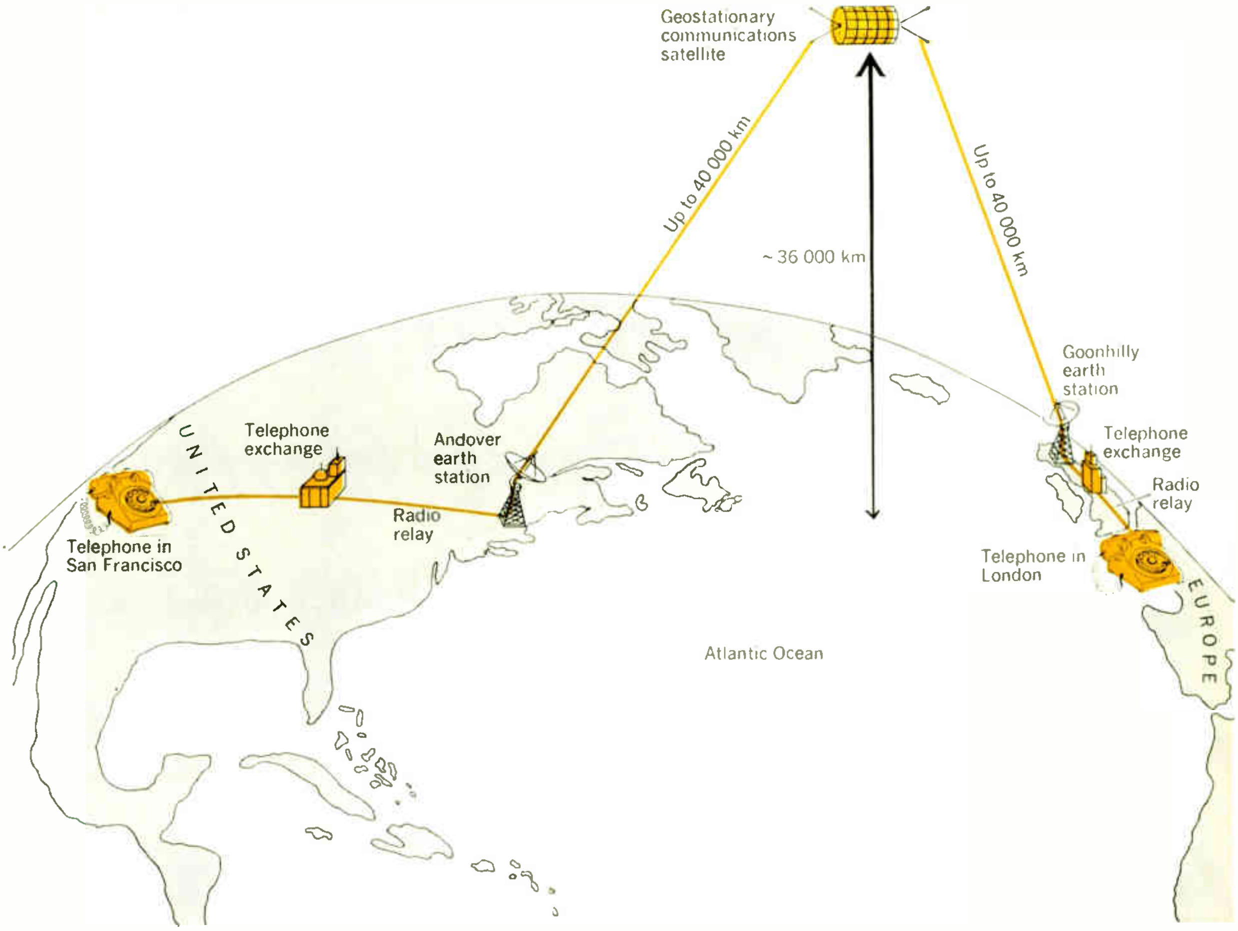


FIGURE 1. Typical satellite circuit.

stations via the satellite in this example is 80 000 km, but other circuits can be as long as 87 000 km if both earth stations are at the limit of the satellite's coverage area; and although signals travel at the speed of light, it takes 260 ms to reach the distant earth station. Delay in the domestic facilities might add another 20 ms at each end of the circuit—producing a total delay of the order of 300 ms. After finishing a statement, twice 300 ms must elapse before one party can possibly hear the response of the other. It is this 600-ms added delay that was supposed to have made the wife in the conversational exchange that introduced this article, think her husband was hesitant about inviting the Thompsons.

Echo

In addition to the delay between his remarks and a response, a caller will hear the echo of his own voice (talker echo), delayed by 600 ms, if preventive steps are not taken.

Echo is ever present. Our speech waves impinge on the walls around us. Some energy is absorbed but some returns to the ear as echo. Normally this is not noticeable because of the short time delay between the speech and echo. However, echo can sometimes be very noticeable, as in a large, enclosed area with low-absorption surfaces. A stone cathedral is a good example. Moreover, echo that is delayed sufficiently, and is loud enough, can interfere with the ability to speak. For example, many people speaking into a test-telephone handset, in which the transmitter and receiver are connected through a delay device, find it almost impossible to speak without stuttering if the delay is about 250 ms and the echo is as loud as their speech. In one experiment, the only person able to speak coherently was a professional organist, accustomed to playing at a given tempo despite the differing amounts of delay with which he heard his own output. (Not only do organists hear their output delayed, but different notes are heard with differing delays because of the disparate distances of the many pipes from the keyboard.)

Interestingly, the confusion caused by delayed auditory feedback has been used to identify malingerers who claim to be deaf, or hard of hearing, in one or both ears.[1] Any stuttering or disruption of a suspect's normal speech indicates that that person hears the echo.

Such a delayed-echo effect may in part also be responsible for our annoyance at being mimicked. It could be that during portions of the mimicking episode, the "echo" interferes with the normal physiological feedback mechanism and confuses, and hence annoys, the speaker. Such mimicking—as may sometimes be practiced by children—is not usually as effective as the speech-stopping experiments just described because the normal aural-feedback path is not blocked and the mimicking is not a replica of the original speech in time or frequency; nor is it customarily as loud as in the experimental setups we've mentioned.

Echo also can be generated on telephone circuits.[2] A simplified long-distance telephone connection is shown in Fig. 2. The transmitter and receiver in the telephone are connected to the central office by a two-wire line. However, within the telephone set itself the transmitter and the receiver each require two wires. These are connected to the line through the use of a hybrid coil (discussed later). The hybrid coil also couples some of the signal from mouthpiece to earpiece. This coupled speech is, in a sense, echo with zero delay, and it is called sidetone. Sidetone is desirable since it makes the circuit seem alive. Try speaking into a telephone handset while holding the switchhook or button down and observe how "dead" it is.

The use of only two wires in the subscriber line, for both the transmission and reception of speech energy between parties, is based on reasons of cost. (Besides less wire, there is less switching equipment since two, not four, wires need be switched.)

Two-wire transmission circuits are perfectly acceptable for local telephone calls when two subscriber lines are simply connected together in the local switching office. But for long-distance calls, it is frequently necessary to split the directions of transmission. For example, the use of multiplexing equipment—where a number of telephone circuits share the use of one radio channel—requires that the signals in one direction be sent over a different radio-frequency channel than the signals in the other direction. To perform this separation, a hybrid coil is used for each telephone circuit at the points in the connection where a two-wire circuit is changed to a four-wire one.

Basically, a hybrid coil consists of two transformers in a bridge circuit. Speech energy entering the line (two-wire) side of the hybrid coil is coupled to the transmit branch for transmission to the distant end. (It is also coupled to the receive branch, but it is blocked there by a one-way amplifier in the multiplexer.) Speech energy entering the receive branch from the distant end is coupled to the two-wire line. Unfortunately, if the bridge is not balanced, a portion of this receive-branch speech will also be coupled to the transmit branch. This energy will then be transmitted back over the circuit and appear at the distant end as echo.

The only way this echo can be prevented at the hybrid coil is for the electrical impedance of the hybrid-coil network to match exactly the impedance of the connected two-wire line. However, there are some 100 million two-wire subscriber lines in the United States—no two of which have precisely the same impedance. Thus, a hybrid coil in the telephone central office might be connected to any one of a large number of subscriber lines. Since it is not possible to select one network that will match all possible lines, a single "compromise network" (900-ohm resistor in series with a 2-μF capacitor) has been chosen. The use of such a compromise network throughout the United States results in a distribution of echo losses having an average of about 15 dB and a standard deviation of 3 dB.

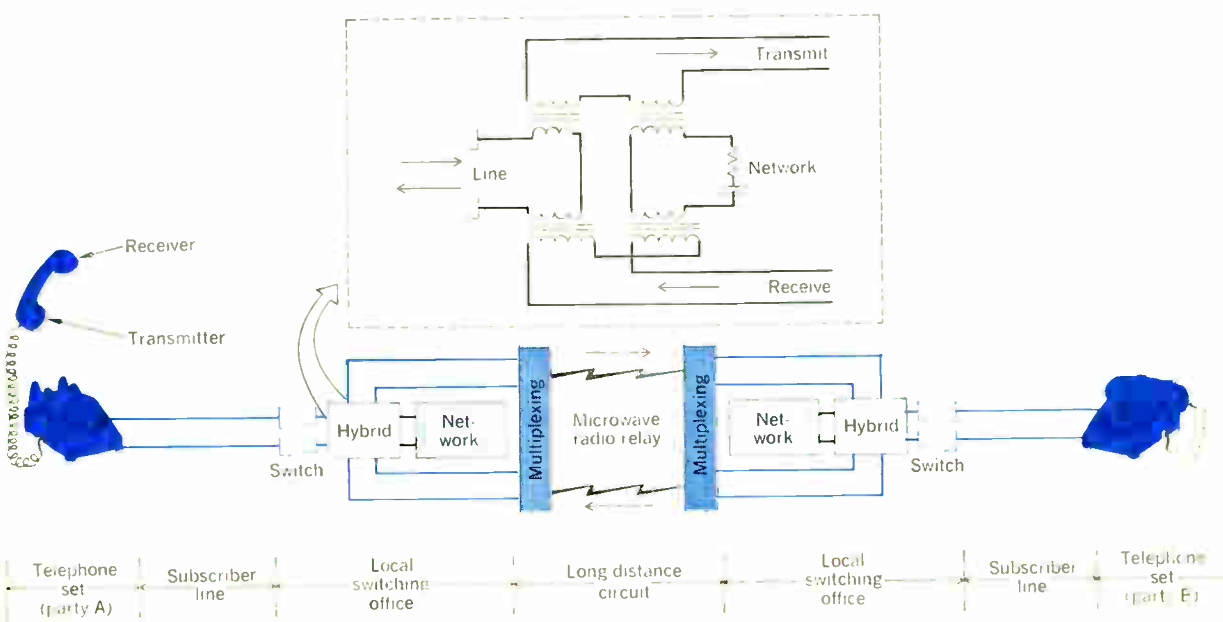
When the long-distance connection is less than about 2500 km (the distance between hybrid coils), the time delay of the echo is such that the echo loss through the hybrid and through the other portions of the connection is sufficient to make the echo almost undetectable.

For longer connections, and particularly when a geostationary satellite forms part of the connecting loop, the delayed echo can create annoyance unless echo suppressors are used.[2-4]

Echo suppressors

An echo suppressor is a voice-operated device placed in the four-wire portion of the long-distance connection that blocks the echo's return path. Figure 3 shows a simple echo-suppressor scheme. If party B speaks, the echo suppressor near party A recognizes speech energy in its receive path and a circuit blocks A's transmit path. Thus the echo that leaks around the hybrid is blocked in

FIGURE 2. Simplified long-distance telephone connection.



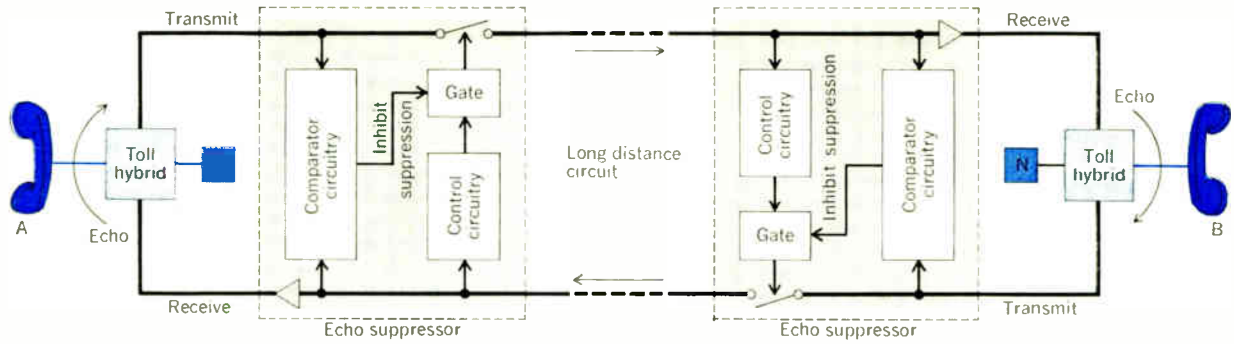


FIGURE 3. Simplified long-distance telephone connection with echo suppression added for better intelligibility.

the echo suppressor and is not heard by party *B*. The suppressor near party *B* works in the same way for the echo of party *A*.

However, interruptions during conversations are a fairly common occurrence and if *B* is speaking and *A* interrupts, this simple suppressor is not adequate. Paul T. Brady at Bell Telephone Laboratories has found that when one party is speaking, there is about a 20 percent probability that the other party will interrupt—i.e., the parties begin "double-talking."[5] If *A* interrupts when *B* is speaking, the simple echo suppressor will suppress the speech of *A* also, and *B* will not hear the interruption.

This is not desirable, so the echo suppressor must include circuitry to compare the speech energy in its transmit path with the speech energy in its receive path to decide whether the transmit speech energy is caused by *A* speaking or by the echo of *B*. If the decision is that the energy is caused by *A*, the comparison circuit will remove the suppression in the transmit path and permit *A*'s speech energy to pass. (In newer suppressors, not all of the suppression is removed during these double talking situations; a small amount of loss is inserted in the receive path. This has the effect of reducing the volume of the direct speech a little, but also of suppressing the echoes somewhat during such periods.)

It is not always possible for the echo suppressor to make correct decisions based on this comparison. For example, *B* may be a loud talker and have small echo loss through the hybrid whereas *A* may be a quiet talker. In this case, *B*'s echo energy may well be greater than *A*'s speech energy, and the decision will be to not remove suppression. Thus, *A*'s desired speech will be suppressed.

Much recent study has gone into this problem of distinguishing double-talking from echo.[6–8] However, comparison circuits of reasonable cost cannot make infallible decisions, and this results in some loss of speech (speech chopping) when both parties are talking at the same time. The number of times in a conversation that chopping, and its concommitant annoyance, occurs increases with lengthening transmission delays.

Tests of effects of transmission delay

The effects of transmission delay alone, in the absence of echo and echo suppressors, cannot be determined on regular telephone circuits because echo is always present and cannot be removed. Thus, studies of delay alone must be made in the laboratory.

In laboratory tests, a four-wire circuit is used in which delay is introduced by using a record-reproduce device such as that shown in Fig. 4. Speech from party *A* is recorded on the magnetic disk and played back to party *B* after delaying it by an amount proportional to the speed of the disk and the angle between the record and playback heads. (It is sufficient to use only one device in the connection that provides a delay equal to the total round-trip delay to be tested.)

Dennis L. Richards of the British Post Office Research Station at Dollis Hill performed some of the first investigations in the field of echo-free delay.[9] Richards was interested in the changes in the dynamics of conversations on circuits having appreciable delay. He defined the conversational states that one party (say party *A*) would encounter. First, if *A* only is speaking, he is in a *talk* state; if *B* only is speaking, *A* is in a *hear* state; if both are speaking, *A* is in a *talk and hear* state; finally, if neither is speaking, *A* is in a *neither talk nor hear* state. Transitions can take place between the states, as shown in Fig. 5. Richards defined the transitions from *talk* to *talk and hear* and from *hear* to *talk and hear* as *double-talk* transitions. He further defined these as follows: If party *A* goes from the *hear* state to the *talk and hear* state, he has begun talking voluntarily. On the other hand, if he is in the *talk* state and goes to the *talk and hear* state, it is involuntary on his part. With no delay between the parties, this transition is caused by party *B* voluntarily interrupting. However, if there is significant delay in the circuit, party *B* could have been speaking before *A* began. In this case, the transition to the *talk and hear* state is involuntary on the part of both *A* and *B*.

Richards determined that this type of involuntary double-talking is closely related to confusion in the conversation—particularly when it results in the original speaker stopping and the interrupting speaker continuing. This he called *alternating involuntary double-talking*. In tests where the test subjects (parties *A* and *B*) were working on a puzzle involving a fair degree of rapid interchange of ideas, he found that during the rapid interchange, the incidence of this type of double-talking increased from two per minute at zero delay, to four per minute at 600-ms round-trip delay, and to six per minute at 1200-ms round-trip delay.

Peter Bricker at Bell Laboratories took a different approach to the question of delay.[10] He performed a test to determine if subjects could detect added delay. For his test subjects he chose pairs of people who felt strongly

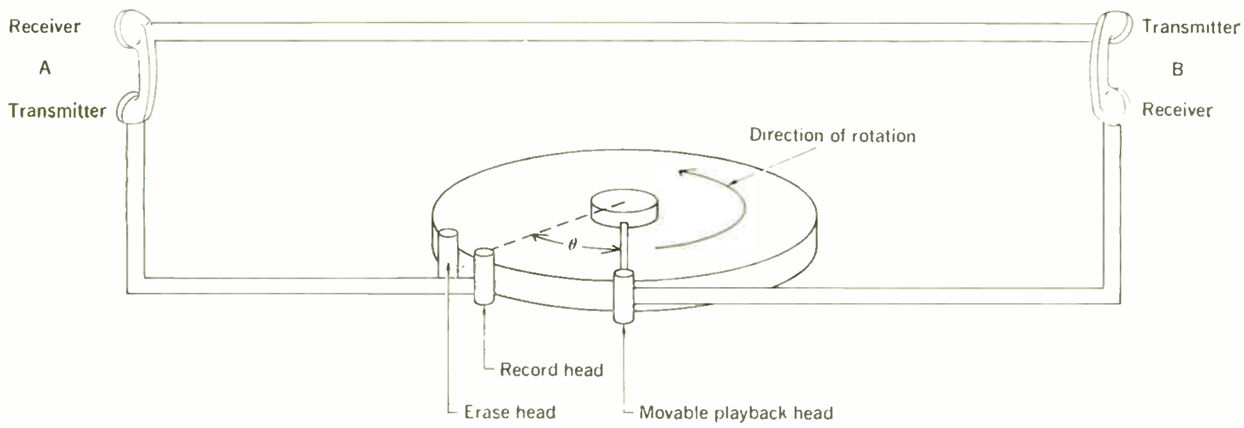


FIGURE 4. To simulate varying delay in the laboratory, a record–reproduce device is inserted in a four-wire network. Delay is equal to period of rotation times $\theta/2\pi$.

about certain political or social issues and who could engage in an animated conversation about the issues. Prior to the test the subjects were warned that a 600-ms delay would be inserted arbitrarily sometime during the conversation. They were instructed to push a button when either detected the delay. The histogram in Fig. 6 presents the results of 20 tests made with six pairs of subjects. Note that it took most subjects a minute or longer before detecting.

Robert Riesz and Edmund Klemmer at Bell Laboratories approached the problem in yet another way.[11,12] They wanted to determine the reactions of users to delay in normal telephone conversations. To do this they converted about 20 telephones in the Murray Hill Laboratory to four-wire instruments. These telephones belonged to a group of administrative personnel who made frequent calls among themselves. Special switching equipment was used to provide a full, four-wire path between telephone sets whenever any one member of this group called another member in the course of normal business. On alternate days, 600-ms and 1200-ms total delay was introduced into the four-wire path.

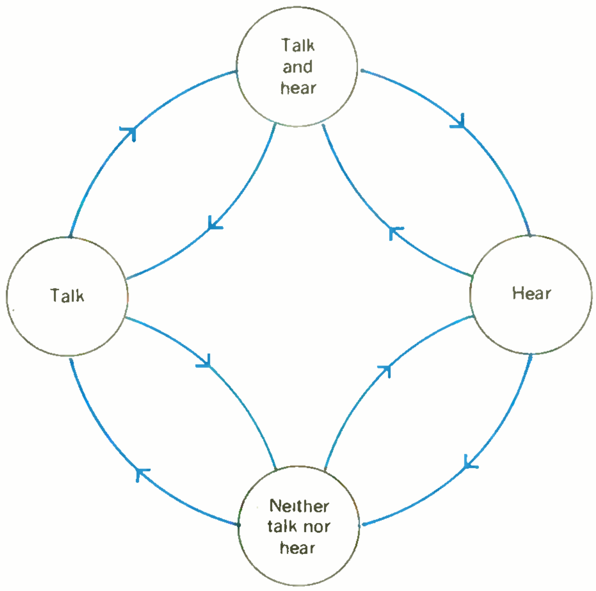
The subjects were told that some of their calls would go over special experimental circuits, but they were not told of the nature of the experimental circuits nor were they told which calls would be affected. They were told that if they found the circuit "unsatisfactory for normal telephoning" they should dial the digit 4 without hanging up. This would restore the standard circuit. The test ran for 12 weeks during which time 523 calls were made over the test circuits. Initially, a few 4s were dialed, but for the last half of the trial no one dialed a 4—obviously indicating that the subjects found the majority of the calls satisfactory. The subjects did not become sensitized to the delay and more critical as the test progressed as had been thought possible before starting the test.

Krauss and Bricker[13] performed other tests, as has one of this article's authors while at Stanford Research Institute.[14] Considering the findings of results of all tests described, there appears to be less transmission degradation on echo-free, long-delay circuits than originally feared. There are changes in the dynamics of the conversation, however, and there is yet no final answer concerning the effects of widespread use of echo-free circuits having long delay.

Effects of delay, echo, and echo suppressors

In testing for the effects of delay alone, researchers have been restricted to the laboratory; when echo is permitted, studies can also be made on regular telephone calls. In the early 1960s, studies were in progress at Lenkurt Electric Co., Bell Laboratories, and the Dollis Hill Laboratories of the British Post Office. These tests showed that the then extant echo suppressors had to be modified before they could be used for telephone calls on long-delay circuits. Moreover, these studies resulted in specifications for new echo suppressors specifically designed for long-delay circuits. Various experimental versions of these new echo suppressors were used in the subsequent studies on working telephone circuits.

FIGURE 5. Defined conversational states for a single party on a telephone connection.



Two major studies were conducted by Bell Laboratories in which actual telephone subscribers participated.[15] Transatlantic circuits were used for both studies. The first used simulated delay; the second was conducted with Early Bird. Telephone administrations in the United Kingdom, France, West Germany, and Italy participated in one or both studies. In the United States, the FCC, NASA, and Comsat also were involved in planning the studies.

In both studies, persons in the United States who had just completed a transatlantic call were called by professional telephone interviewers and asked a series of questions—almost all related to the call just completed. Two important questions were

1. Did you, or the person you called, have any difficulty talking or hearing over that connection?

2. Which of these four words comes closest to describing the quality of that connection: excellent, good, fair, or poor?

If the answer to the first question was "yes," the interviewer tried to get the customer to describe the difficulty by probing in a nondirective fashion. The customer's description was recorded verbatim for later analysis.

The first study took place in 1964 before the advent of Early Bird, the first commercial geostationary communications satellite. A delay device was inserted in the outgoing path of the four-wire cable circuit at the overseas gateway in the United States. The delay was adjusted at various times in the study such that the round-trip delay in the circuit from gateway to gateway was 800, 600, 400, 200, and 90 ms. The last-mentioned delay is close to the inherent round-trip delay of the ocean cable circuits, whereas 600 ms represents that of geostationary satellite circuits. The other delays are representative of various possible combinations of satellite and terrestrial circuits.

When the Early Bird satellite was launched in 1965, similar interview studies were performed. The best type of echo suppressor, out of a number used in the 1964 tests, was used for the Early Bird circuits. Also, during the Early Bird studies, an opportunity was available, not present in the 1964 studies, to determine whether any sensitizing would take place for users who had talked a number of times over long-delay circuits. (In the 1964 studies, most users had experienced only one call over a long-delay circuit at the time they were interviewed.) Punched-card records were kept for a few months of almost all the satellite calls made or received by customers in the New York City area. Interviews were then begun with all customers who had just completed either a cable or a satellite call. (The records indicated those who had a previous history of satellite usage.)

FIGURE 6. A histogram from a study conducted to test for time to detect the insertion of delay.

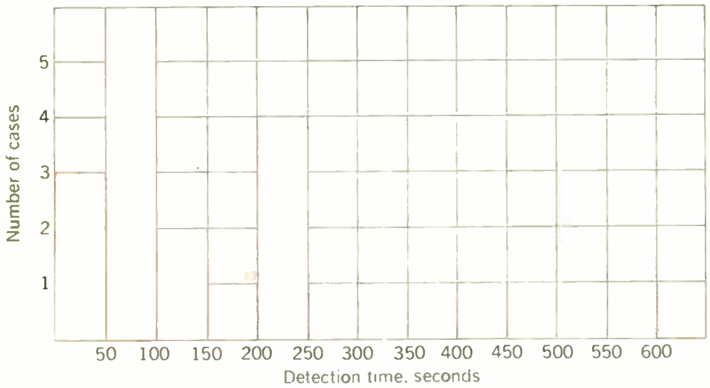


Figure 7 shows, as a function of round-trip delay, the percentage of all customers in both the 1964 and 1965 studies who answered "yes" to question 1 and who rated the connection "fair" or "poor" in answer to question 2. Included in Fig. 7 are the results of all interviews (some 7000) regardless of previous satellite-calling history.

To determine the effects of repeated satellite use in the Early Bird studies, the customers were arbitrarily divided into those who had made five or less satellite calls and those who had made six or more. There was essentially no difference in the interview results for the two groups, indicating that, at least for the limited duration of the study (about five months), there was no sensitization to the longer-delay circuits.

Another interesting result of the studies was that the average duration of delayed calls (about ten minutes) was neither longer—as some researchers had postulated on the premise of time lost in transferring information, or shorter—as predicted by others who anticipated a reaction to dissatisfaction with the poorer circuit quality.

We mentioned that the customers also tried to describe the difficulties they had had. These descriptions were later coded into nine different categories, such as *noise*, *low volume*, *cutting*, and *echo*. It was found that the largest increase in dissatisfaction occurred in categories that we believe most nearly describe the speech mutilation caused by echo suppressors—such as *cutting* or *fading*. *Delay* and *echo* complaints were rare. It appears that delay by itself did not cause difficulty and that echo suppressors do a good job of suppressing echo but a less than perfect job of permitting double-talking. Similar Early Bird studies were conducted by the British Post Office with almost identical results.[16]

Different Early Bird studies were conducted by the telephone administrations in The Netherlands, Scandinavia, and other countries.[17,18] These involved monitoring calls to ascertain (1) how delay affects the conversational dynamics and (2) the comments made by the customers during the conversation relating to the quality of the circuits—such as the difficulties they might be having ("What did you say?" "I didn't get that."). These studies also showed some differences in quality, favoring cable circuits, but they were not nearly as large as those indicated for the Early Bird studies.

Advances in echo control

There is continuing research into new techniques for echo control. A full four-wire circuit for every telephone in the world is not economically feasible in the foreseeable future. Nor is it possible to match the impedance at all the hybrids in the long-distance network to such a degree that the echo loss would be sufficient to eliminate the need for echo suppressors.

New methods for echo control have been proposed by many people throughout the world.[19–22] One proposal is for an adaptive device that injects a signal into the echo path, which cancels the echo. An analog adaptive echo canceler using a transversal filter is shown in Fig. 8; this replaces the echo suppressor of Fig. 3. In essence, this "canceler" operates as follows: The speech $x(t)$ in the incoming path is fed to a tapped delay line, having delay increments τ. The output of each tap is then sent to a cross-correlator consisting of a multiplier and an integrator. The other input to the cross-correlator is the residual echo $e(t)$. The output of each cross-correlator is a weighting or gain $g_i(t)$, which can be positive or negative. It is used to adjust the delayed speech signal. The products of $g_i(t)x(t - i\tau)$ are then summed and the sum $\hat{y}(t)$ is subtracted from the uncanceled echo $y(t)$.

Important characteristics of such a canceler are the amount of achievable cancellation and the time required for adaption. The subjective importance of these characteristics are not well known at this time.

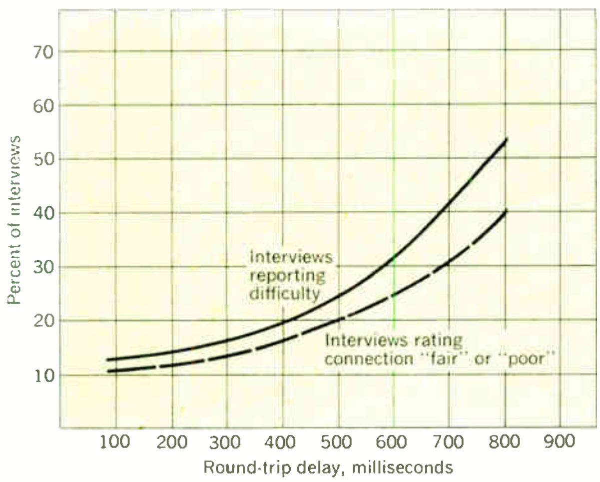


FIGURE 7. Persons interviewed after a telephone conversation on circuits with transmission delay answered according to this sort of response pattern.

Consider such a canceler, which has adapted to cancel the echo. Now if speech from the near party appears in the transmit path, the echo canceler will try to readapt to the new echo. Means must be employed to prevent this. Also, if the echo path changes, perhaps because a PBX operator transfers the call to another telephone, the echo canceler must readapt. Slow changes of gain in the echo path are also possible and these require continuous adaption.

An alternative to the adaptive echo canceler is one based on the measurement of the impulse response of the echo path.[22] Immediately after a connection is established, the echo canceler transmits an impulse in the receive branch toward the echo path. The signal appearing in the transmit branch is the impulse response. This signal is converted into digital form and stored in the echo canceler where it is multiplied with the subsequent speech in the receive branch to produce an echo-canceling signal. Essentially this canceler adapts to the echo path using the impulse response signal rather than the echo of the speech signal.

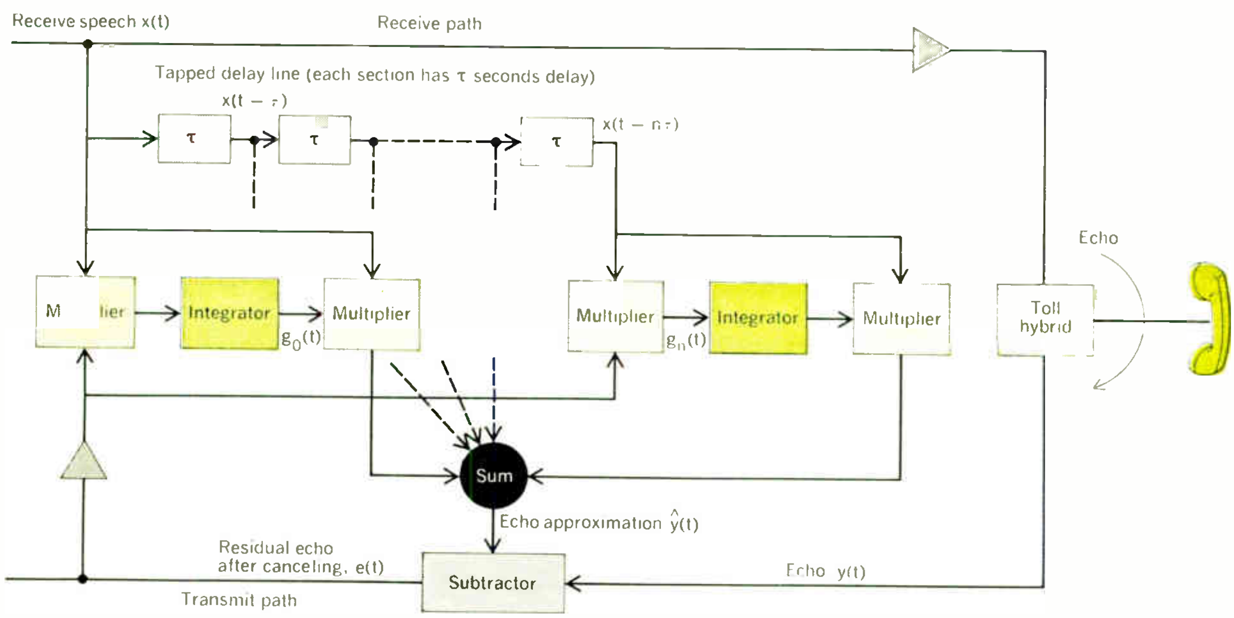
This method has problems similar to those just mentioned for the speech-adaptive echo canceler. In addition, echo paths are not always linear. In these cases the response of the path to an impulse may be different from the response due to speech, and the subsequent compensating signal may not, in fact, cancel the echo.

If echo cancelers could work perfectly, the present-day telephone circuits could be made echo-free. Laboratory simulations of the speech-adaptive echo canceler on ideal circuits have been encouraging. However, success depends on actual telephone network conditions—and these are far from ideal. The proposals to date are still in the early research stage and there is no assurance of ultimate success.

Conclusion

In all the studies conducted to date, transmission delay, typical of geostationary satellites, by itself has been shown to be undetectable by many; it is unobjectionable

FIGURE 8. Simplified echo canceler.



to most. Studies of actual telephone circuits with echo and echo suppressors in addition to delay have almost consistently shown low-delay circuits to be preferable to high delay. The subjective test results vary considerably depending on the experimental method employed.

Transmission quality is not the only factor that must be taken into account when choosing among different types of transmission facilities on a given route. Cost, availability, and protection against major catastrophe are other considerations that also apply. Factors such as these are included in the wordwide telephone network planning studies of the CCITT (International Telegraph and Telephone Consultative Committee—a body of the International Telecommunication Union, one of the specialized agencies of the United Nations). The CCITT has made the following recommendations regarding circuits having delay: (1) round-trip delay less than 300 ms—acceptable without reservation, (2) round-trip delays between 300 to 800 ms—acceptable, provided that increasing care is exercised on connections if the round-trip propagation time exceeds about 600 ms and provided that echo suppressors designed for long-delay circuits are used; (3) round-trip delays greater than 800 ms—unacceptable except under the most exceptional circumstances.[23]

The use of satellite circuits is expanding. Because the effects of widespread use of telephone circuits with long delay are not well known, and because advances in echo control are possible, the CCITT is continuing to study these items.

REFERENCES

1. Tiffany, W. R., and Hanley, C. N., "Delayed speech feedback as a test for auditory malingering," *Science*, vol. 115, pp. 59–60, Jan. 18, 1952.

2. Emling, J. W., and Mitchell, D., "The effects of time delay and echoes on telephone conversations," *Bell System Tech. J.*, vol. 42, pp. 2869–2892, 1963.

3. Phillips, G. M., "Echo and its effects on the telephone user," *Bell Lab. Record*, vol. 32, pp. 281–284, 1954.

4. Richards, D. L., and Buck, G. A., "Telephone echo tests," *Proc. IEE (London)*, vol. 107B, pp. 553–556, 1960.

5. Brady, P. T., "A statistical analysis of on-off patterns in sixteen conversations," *Bell System Tech. J.*, vol. 47, pp. 73–92, 1968.

6. Brady, P. T., and Helder, G. K., "Echo suppressor design in telephone communications," *Bell System Tech. J.*, vol. 42, pp. 2893–2918, 1963.

7. Unrue, J. E., "Echo suppressor design considerations," *IEEE Trans. Communication Technology*, vol. COM-16, pp. 616–624, Aug. 1968.

8. Richards, D. L., and Hutter, J., "Echo suppressors for telephone connexions having long propagation times," *Proc. IEE (London)*, vol. 116, pp. 955–963, 1969.

9. Richards, D. L., "Conversational performance of speech links subject to long propagation times," *Internat'l Conf. on Satellite Communications*, IEE, London, pp. 247–251, 1962.

10. Bricker, P. D., "Would time delay be a problem in using a synchronous satellite?," *Satellite Communications Physics*, Chap. 5, 1963 (high school science teaching aid prepared by Bell Tel. Labs).

11. Riesz, R. R., and Klemmer, E. T., "Subjective evaluation of delay and echo suppressors in telephone communications," *Bell System Tech. J.*, vol. 42, pp. 2919–2941, 1963.

12. Klemmer, E. T., "Subjective evaluation of transmission delay in telephone conversations," *Bell System Tech. J.*, vol. 46, pp. 1141–1148, 1967.

13. Krauss, R. M., and Bricker, P. D., "Effects of transmission delay and access delay on the efficiency of verbal communication," *J. Acoust. Soc. Am.*, vol. 41, pp. 286–292, Feb. 1967.

14. Gould, R. G., "Echo suppression and time delay, a subjective evaluation," *IEEE Trans. Communication Technology*, vol. COM-12, pp. 74–82, Sept. 1964.

15. Helder, G. K., "Customer evaluation of telephone circuits with delay," *Bell System Tech. J.*, vol. 45, pp. 1157–1191, 1966.

16. Hutter, J., "Customer response to telephone circuits routed via a synchronous-orbit satellite," *Post Office Elec. Eng. J.*, vol. 60, pp. 181–186, Oct. 1967.

17. Hansen, E., "Analysis of traffic via and operation of satellite circuits," *Teleteknik*, vol. 1, no. 1, pp. 1–10, 1967.

18. "Subscribers reactions to 'Early Bird' circuits," *Het PTT Bedrijf, Third Internat'l Symp. on Human Factors in Telephony*, vol. 15, pp. 37–38, 1967.

19. Becker, F. K., and Rudin, H. R., Application of automatic transversal filters to the problem of echo suppression," *Bell System Tech. J.*, vol. 45, pp. 1847–1850, 1966.

20. Sondhi, M. M., "An adaptive echo canceller," *Bell System Tech. J.*, vol. 46, pp. 497–512, 1969.

21. Thies, A. W., and Zmood, R. B., "New ways of echo suppression," *Australian Telecommun. Rev.*, vol. 1, pp. 14–19, 1967.

22. Muira, A., Kobayshi, S., Sato, R., and Nagata, K., "A blockless echo suppressor," *IEEE Trans. Communication Technology*, vol. COM-17, pp. 489–495, Aug. 1969.

23. *CCITT White Book*, vol. V, Recommendation P. 14, 1968.

Richard G. Gould (SM) received the B.E.E. degree from New York University in 1949 and the M.S. degree in electrical engineering from the University of Southern California, Los Angeles in 1954.

As an engineer and senior engineer in aerospace and electronic organizations, he was engaged in the manufacture of components and the design and analysis of missile guidance and radar systems. He has also served as the assistant radio officer for the Eighth U.S. Army in Korea.

In 1956 he joined the staff of Stanford Research Institute, Menlo Park, Calif. There he conducted studies of the effects of natural and man-made interference on HF systems, of ionospheric sounders, communications satellites, and transmission delay and echo suppression.

For six years during this period he was also an instructor of electrical engineering extension courses at the University of California, Berkeley. In 1964 he joined the Communication Satellite Corporation, where his first responsibilities were for studies of special-purpose systems. Later he became assistant to the technical vice president. Mr. Gould joined the Office of Telecommunications Management, Executive Office of the President, as a consultant in 1968. He has been working on advanced technology on the design and management of an extensive propagation experiment, and on CCIR and CCITT matters. He has been a member of several U.S. delegations to the CCIR.

George K. Helder (M) received the B.S. degree in business from the University of Colorado, Boulder, before switching academic interests. Then, in 1958, he earned the B.S.E.E.—also from U.C.—and two years later was graduated from New York University, with the M.E.E. degree. He has been with the Bell Telephone Laboratories since 1958, first engaged in exchange-area transmission (including methods of subscriber loop testing), and later concerned with the problems of echo control on telephone connections and with studies to determine subjective effects of telephony over long-delay circuits. Currently, he is chairman of the CCITT Working Party on Echo Suppressors, which is studying various methods to improve echo control in worldwide telephony. At Bell Labs, Mr. Helder heads a group that provides long-range plans for the evolution of the telephone network from the predominant, present-day use of analog methods in transmission and switching to future use of digital methods.

Conversion and the import problem: a confluence of opportunities

Finding new markets for the defense electronics industry and facing up to an increasing import problem are two principal tasks for industry planners in the 1970s. It may be possible to use surplus defense resources for a major quality and production improvement program to meet these two challenges

John E. Ullmann Hofstra University

The defense electronics industry is again faced with the prospect of cutbacks, which brings the problem of its conversion once more to the fore. At the same time, the growth of imports presages sharp competition for U.S. industry for the products it can devise to take up the slack. Further, the electronics industry has a quality problem, in its military, industrial, and consumer sectors. In this article, the writer suggests that the industry's displaced technical talent should be organized to tackle systematically its scientific, engineering, and economic problems in order to put it ahead once more. As a vehicle of change, he proposes a National Technology Foundation to oversee the program and to undertake tasks now beyond the scope of individual companies.

As we enter the 1970s, the U.S. electronics industry can look back on two decades in each of which its sales volume more than tripled. It might seem ungracious in view of blessings received, or else highly alarmist, to call attention here to two problems, but these promise to loom large in the industry's future and so must be taken into account by its business planners.

One important problem is conversion—the task of finding commercial or other nonmilitary tasks for the large sector of the industry that now does defense work. A second problem is competition from other countries. Although it has not been extensively noted, the two actually are quite closely linked in at least three ways. First, it will not be easy to find new nonmilitary markets for electronics, and those that do seem promising will prove illusory if they are taken over by imported items. Second, the ability of U.S. industry to respond to international competition depends on its cost structure and the quality and variety of research talent it brings to bear on these new areas; both of these factors are critically affected by the extent of its defense concentration. Third, the changes necessary for success in the coming years must be made against a background of a crisis of confidence in product quality, of which the monumental cost overruns and performance deficiencies of military systems on one hand, and inferior electronic consumer products on the other, are prominent symptoms.

These problems, of course, are not peculiar to U.S. industry; the question of quality in consumer items, for example, is obviously as international as the manufacturers themselves. However, the simultaneous occurrence of all three factors is perhaps more pronounced in the United States than anywhere else. No discussion of the future of U.S. industry, therefore, can properly fail to take them into account.

Conversion, competition, and quality control are extensive topics in themselves and the following discussion of their interconnections necessarily must be somewhat selective. Still, the proposals made here attempt to resolve some difficulties that have beset all U.S. electronics firms at one time or other. This is not meant to suggest that specific firms actually can and will take advantage of the opportunities described. Rather, since each firm ultimately is responsible for the crucial decisions that determine its future, it is only possible, as it were, to create more waters to which horses may be led. Emphasis on new products and new methods is essential, because in nonmilitary items it would be exceedingly difficult for defense electronics firms to buck established competition.

The conversion problem

The existence of a conversion problem is generally conceded, at least in the sense that manufacturers of military equipment have to find other work when their contracts expire. However, as a practical matter, so far there has not been much need for planning and implementation of this kind, mainly because no sustained cuts in defense spending have been made in at least the last eight years. Market planning by defense contractors could thus look forward in general to an expanding overall market, or in any case to another contract once a first one was completed or irretrievably bungled. Contracts were "slots," and surely the ratchet mechanism that defense budgeting had exhibited would assure that nothing drastic would happen.

The stability in spending at around the $35 billion level of the last Eisenhower years was succeeded by a sharp jump to some $50 billion as the Kennedy Administration came in and decided on a quantum jump in warmaking capability. Things changed later, but the cuts made by President Johnson in 1964 soon disappeared in the escalation of the Vietnam war, which boosted spending to $80 billion. These past periods of stability coincided with considerable interest on the part of management in conversion, but those were also times when company funds were limited. Management tended to act like the proverbial farmer who can't paint his barn when times are good because he is too busy, and can't do it when times are bad because he then has no money.

Developments over the past year suggest that strong concern with the conversion problem is again warranted. For the first time since 1961, there has been a fundamental change in the nation's strategic planning, involving a cut in combat readiness from "two and a half wars" to "one and a half wars."[1] The implications are that the coming years will include a 40 percent manpower cut in the armed forces[2] and a cut in expenditures for fiscal 1971 of some $5 billion, which in turn will mean the loss of some 212 000 defense-related jobs. Three hundred DOD installations will also be closed, with the loss of a further 70 000 jobs, some of them technical.[3] Recent estimates have put the total effect at close to 1.25 megajobs. Moreover, these reductions are coming at a time when increased unemployment is being expected as a side effect of anti-inflation medicine. Finally, the space program is suffering from what might be called a *postpartum* depression; many of its more talented executives and engineers have left and various facilities have been closed.[4] Cape Kennedy has become a distressed area.

It has been argued, of course, that all these cutbacks are only temporary, and that in a few years the defense budget will be back at the $80 billion level even without a Vietnam-type enterprise. It is pointed out that weapons inventories have been depleted because of the war, that some projects have been deferred, and that, because of inflation, everything will cost more. Again, it is said, there is not really much to worry about. Disarmament seems more remote than ever, since the SALT talks appear designed more to keep Russo-American weapons preeminent than to lead to any real reduction.

It is not impossible that future events will evolve along these lines. At the same time, however, the nonmilitary needs of the nation—in such areas as housing, the environment, transportation—are being brought before the public as never before; the "priorities" issue is now being presented as a matter of real choice, so that the expectation of a blank check for the military from here to eternity does not appear to be a sound basis for business planning. In any event, even the possibly short-term problems identified in the foregoing merit action; furthermore, the considerations of international competition in what follows make a reorientation of the industry's talents and resources imperative.

It is not the purpose of this article to discuss conversion in terms of a "wishing list." Alternative products for the defense industry have been extensively reviewed in the literature, including works pertaining particularly to the defense electronics industry.[5] In general, new products can be classified as (1) those that derive from what might be called "natural" markets—i.e., the array of actual and potential consumer products that the industry might produce and would be bought by customers without government intervention or subsidy (computer systems, industrial controls, such new consumer products as television tape units); and (2) those that would require government stimulation for the markets to become realities.

It is in the latter area that much of the public discussion has taken place. There have been extensive proposals, for example, for new approaches to air transportation. Thanks to ground and air congestion, the chant of the witches of *Macbeth*, "Fair is foul and foul is fair. Hover through the fog and filthy air," accurately describes the situation at our major airports. In many of the areas subsumed under this second category—transportation, housing, education, pollution control, oceanography—the role of the electronics industry, though important, depends on satisfactory performance by others. In transportation, for instance, pollution abatement is not specifically an electronics problem, although its measurement is, and certain new forms of electric power transmission may be. Other areas require technical and cost breakthroughs, as in structural design; or new knowledge, as in ecology; or extensive software, as in educational systems.

Unfortunately, there is doubt within the industry that conversion actually is the best solution to its future business problems. Many major defense electronics firms are parts of conglomerates, or are divisions of old, established nondefense firms that could close the plants concerned, or at least consolidate them in the area with the lowest labor costs, taxes, etc., and spend their resources where more immediate profits can be realized. As the former head of a defense plant once told the writer," At the end of a war we put a battleship in mothballs. Why should a defense plant be any different?" His actions, alas, matched his words. The firm sold its main plant, moving into an outhouse behind it. This occurred at about the same time as another major local defense contractor cut back by more than 75 percent. These situations are now being duplicated as more work reductions are being carried out elsewhere.

The implications are plain that little will be done to find alternative work within the present firms and localities, which is, of course, the only nonconvulsive way of solving the problem. Such efforts as appear to be in hand would not be nearly enough, even for the limited cutbacks now decided upon. Finally, whatever new plans the industry chooses to make can no longer be carried out independently of the non-U.S. producers of comparable equipment. The days when the industry could expect to solve its problems *in vacuo* are over.

The objective of conversion must not be only that of corporate survival, which would be the typical first

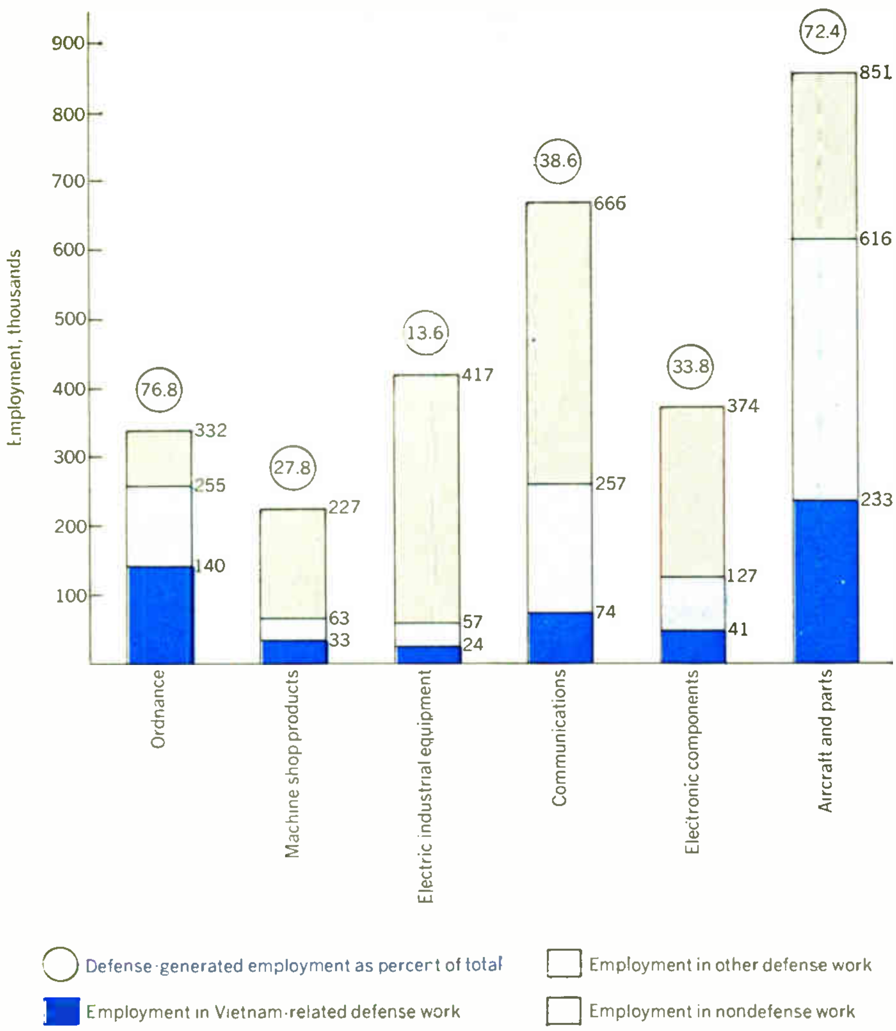


FIGURE 1. Employment share of total defense work and Vietnam-related defense work for selected industries, 1968. Each bar gives the number of jobs resulting from total defense work and the extra employment due to the Vietnam war. In addition to the industries shown, 165 700 more workers are employed on defense work in other parts of the electrical–electronic industry. (Source: R. P. Oliver, "Increase in defense-related employment during Viet Nam buildup," Monthly Labor Rev., vol. 93, pp. 3–10, Feb. 1970)

priority of management. Rather, conversion also must concern itself with the accommodation of the engineers and scientists who have long made their professional home in the military electronics industry. These have generally been considered as consisting of two groups: those in the thick of technology as such—in research, development, testing, and manufacturing—whose work content is still very much involved with theory, new products, and technical minutiae; and those for whom the technical content of the work has virtually vanished beneath a mountain of paper. It is the first group that may be quite difficult to transfer in sufficient numbers, and it is one of the objectives of this article to propose a scheme for solving the problem. Alternative employment for the latter group is more closely related to the transferability of management talent and has been discussed at length elsewhere.[6]

Foreign competition

The United States' strong role in international trade in machinery and other durables has long been a fact of history, and so it is not surprising that the continuation of that role should be assumed with quite a bit of complacency. American know-how, inventive genius, etc., have been around for a long time and have made us what we are today. This is not to say, of course, that we "invented" everything; all industrial societies have pseudo-historical publicity agents to make that claim on their behalf. Rather, the particular contribution of U.S. industry has been to take a device, wherever "invented," and by successfully organizing its production, transform it from a laboratory curiosity or highly specialized item into something that can be used in households, offices, or factories with high efficiency and at relatively low cost. Since the days of the treadle-operated sewing machine, there has not been an appliance that was not first developed into a commercially viable product in the United States—until the past decade, that is. Nor is this record confined to consumer durables. Diesel locomotives had long been manufactured in relatively small numbers both in the United States and elsewhere. It remained for U.S. firms to develop modular arrangements and standardization to the point where these machines could be virtually mass-produced.

All this is far from arguing that some "manifest destiny" ordains continued U.S. superiority in machinery and consumer durables. But the fact remains that the kinds of developments just described have been the principal basis for U.S. competitiveness in international markets and must be considered a point of departure when the present situation is examined.

The problem of competition from other countries does not relate only to the adequacy of the product development efforts of a specific industry; it is now unfolding against a major economic crisis as well. Its first element is that of domestic inflation, which, as is generally conceded, has resulted from superimposing an exceedingly costly war on a relatively prosperous economy, without the controls and restrictions that might have been effective, but no doubt would have further eroded political support for the war. The attendant expense has handicapped the United States gravely in world trade.

The second element is the balance-of-payments issue. It is true that we have not had to devalue the dollar, but this is due far more to mismanaged French economy and German forbearance (the purchase by the German government of large amounts of U.S. government securities, thereby increasing our national indebtedness) than to our good management. The key to the problem at this time is the deterioration in the trading account (U.S. exports minus imports). For most of the last 20 years, the surplus in that account was large enough to pay for a negative tourism balance, and, more important quantitatively, for our very large military expenditures abroad. The net cash outflow for new overseas investments has been subject to government-imposed controls that have generally held it to about the same amount as receipts from previous investments, so that this is not a major contributing factor. In the past few years, however, the trading account surplus dwindled and then became a deficit. As of the second quarter of 1969 it was running at an annual rate of $2 billion a year[7] and, although it turned again into a small surplus in the third quarter, this was only because of exports financed by the U.S. government. These overall statistics appear even worse when the United States' competitors are viewed country by country. The U.S. long had a large trade surplus with Japan, for instance, but this too turned into a deficit of almost $1 billion by 1968[8] and, if past trends continue, may reach as much as $2 billion within two years or so.

The impact of foreign trade on the electronics industry tends to be concentrated in certain specific areas—components, radio, tape recorders, electronic desk calculators. Of radios, 85 percent are now imported; in its newspaper campaign in connection with the 1969–1970 strike, General Electric noted that its Utica radio receiver plant was the last significant one left in the United States. Some 90 percent of tape recorders are imported; nearly all small television sets are imported. It is noteworthy that these items were first created as viable commercial products by Japan—a break in the U.S. record described earlier. In fact, imports of electronic items are so predominantly Japanese that to the U.S. electronics industry "foreign competition" really means Japanese competition.[9] This competition already promises to be a serious factor in what appears to be the only major new consumer product, the television tape recorder.

Experience with desk calculators is especially interesting, not only because it reflects some highly sophisticated and successful product developments, but also because it illustrates a marketing strategy likely to lead to greater penetration of a wider market.

The desk calculator–computer spectrum may be divided into (1) simple four-rule calculators, (2) more elaborate calculators, (3) programmable desk calculators, (4) minicomputers, (5) large computers, (6) sophisticated subsystems for large computers, and (7) novel software. The conventional U.S. view of non-U.S. competition is that the latter might be conceded the simpler end of the line, leaving U.S. industry in charge of the "high-technology" end. If there is to be competition, it will have to be on a price basis and U.S. firms will not enter that particular field of battle. But this in turn implies a rather basic change in the orientation of one of our most important manufacturing industries. It is quite true, of course, that purely on the basis of manufacturing costs, plus high administrative overhead and marketing costs, the U.S. electronics industry finds it difficult to compete for many markets. However, the great strength of technologically oriented industries in the United States has always been in the development of sophisticated production systems, often suiting design to the exigencies of turning a product out in large quantities. In a real sense, industrial excellence must rely on the ability to keep the production line going simultaneously with research on the product itself and on manufacturing methods. It is this ability that has made "American mass production" a byword and example in all industrial countries.

A concentration on high technology means a different product strategy entirely. As a long-term strategy, moreover, it suffers from several basic drawbacks. It ignores the fact (1) that pioneering is risky and does not enjoy a reputation for profitability; (2) that such arrangements are necessarily unstable; (3) that there is nothing disreputable about price competition, which most consumers tend to regard as the stuff bargains are made of; and, perhaps most important, (4) that such arrangements are impossible to impose unilaterally.

If one applies the "high-technology" preemption to the desk calculator and computer products, then clearly the four-rule machines ought to be solidly foreign by now. A visit to the 1969 Business Equipment Manufacturers Association trade show in New York demonstrated that this has indeed happened. There was nothing to compare to the offerings of Hayakawa-Sharp, Toshiba, Canon, and Sony, the Japanese firms that dominated that product completely. Minicomputers (and larger ones) included Japanese models, the Dutch Phillips, and the Italian Olivetti and Montecatini-Edison.

This does not mean that imports were as dominant in advanced fields as in the simpler ones, but the evidence was clear that non-U.S. producers had imposed no lines of demarcation on themselves. Moreover, capability in advanced computers is a principal objective of modern industrial nations. In addition to the widespread international operations of IBM, most countries seek national systems of their own. The efforts of France through Machines Bull and others are well known. In the case of Japan, a singularly comprehensive and effective arrangement has been made, which, by uniting industry, government, and the universities, has been able to circumvent the normally rather compartmentalized structure of Japanese business. The Japanese Electronic Industry Development Association,[10] together with the Agency for Industrial Science and Technology, is currently sponsoring a major computer development consortium. It seems highly probable that Japanese producers such as Fujitsu, one of the participants, will become significant in the field of large computers.

This sort of government-sponsored research effort on behalf of ultimately commercial products extends to components. Japanese firms that have been so successful in making transistors producible by the millions are now on the way to doing this with integrated circuits as well, including large-scale types (LSI) and metal oxide semiconductors (MOS). Some important theoretical work in the field is being carried out by the Kansai Electronic Industry Development Center, which is supported by the Ministry of International Trade and Industry (MITI) and includes among its participants Osaka University, Hayakawa-Sharp, Matsushita-Panasonic, and other large firms.[11] Indeed, Japanese IC industry has proceeded to the stage at which its entry into the export field is imminent.[12]

In all these developments Japanese research has complemented and implemented U.S. research, and many of the developments have taken place in cooperation with U.S. firms. Joint working agreements, patent cross-licensing, and, most important, offshore manufacturing arrangements, have been concluded between Friden and Hitachi, American Microsystems and Ricoh, Monroe and Canon, Honeywell and Nippon Electric, and Autonetics and Hayakawa, to name a few. Of course, such arrangements are not Japanese alone; "American" television sets and other appliances are being, or shortly will be, made in Hong Kong, South Korea, or in that most recent accession to the ranks of industrial boom areas, Singapore.[13] Japanese industry is now supplying components to other manufacturing countries such as Germany, but as labor costs and labor shortages increase, Japanese industry has begun to import components, parts, and subsystems from its Far Eastern neighbors.

Electronics virtually is emulating the migratory habits of a much older industry, textiles. This industry moved from England to Rumania and Russia in Europe; to India, China, and Japan within Asia; and, within the United States, from New England to the Piedmont, then to the Deep South, to Puerto Rico, and to reliance on Far Eastern imports. It does not require clairvoyance to envisage the typical reaction of management to conversion: Move to cheaper plants within the U.S., resort to joint manufacturing abroad, or move the commercial products out altogether. The Grumman Aircraft Engineering Corporation illustrates the point. Its only significant civilian aircraft is the Gulfstream executive turboprop, which was designed in Bethpage, N.Y., at the company's principal plant, but is manufactured in Savannah, Ga.; a recent report has the company looking into the possibility of assembly in Singapore.[13] Obviously, as cutbacks become a reality, such stratagems do little for the employment prospects of those now in the defense industry. Thus, it is necessary to examine the shortcomings of the present technical achievements of the electronics industry and to find ways in which defense-related talents might be converted to their solution.

The quality problem

There can be little doubt that in electronics, as well as in most other industries, defects in quality constitute a major problem. "They just don't make 'em like they used to" may once have been a bit of nostalgia best ignored, but this no longer is true. Sloppiness in automobile manufacture has been amply documented, to say nothing of safety defects. Repair facilities for automobiles are likely to reach a point of critical shortage and, of course, the television repair area has long been scandal ridden. Difficulties with warranties and guarantees, the inability of manufacturers to secure product liability insurance, the replacement of the legal doctrine of caveat emptor with that of merchantability and implied warranty, and extensive government intervention in product design itself have all drawn attention to the manufacturers' responsibility. This concern has been produced by exactly the kind of massive popular dissatisfaction with product quality to which we have just referred; the U.S. Federal Trade Commission finds itself bombarded by consumer complaints as never before, and, in fact, the political constituency subsumed under "consumerism" has become a strong force.

From a technical viewpoint, the chief culprits are probably cost cutting (i.e., corner cutting) and the pushing of technology beyond its reasonable limits for robust and reliable apparatus. Both of these, in turn, reflect inadequacies in manufacturing and quality control, but also, and more basically, a failure to understand fully the capabilities of modern products and processes. They also are linked to the basic cost structure of the industry and to its ability (and lack thereof) to meet foreign competition. In the present inflationary period, wages and salaries are rising much faster than productivity, which, in turn, is a function of product and process design. Something manifestly has gone wrong when there are so many complaints at the same time that U.S. industry is also being systematically bested by its competitors.

These developments strongly suggest that the scientific resources of the U.S. electronics industry are not being used to best advantage to solve these problems. This misapplication has three aspects. The first, perhaps most important, is the diversion of industry talent to weaponry. Some 80 percent of electronics engineers, according to EIA statistics, are supported by weapons or space contracts. They are channeled there by wage differentials of as much as a third or more, which effectively constitutes rationing by the purse.

The second aspect is related to the first. Responding to various fashionable "gaps," U.S. engineering schools have tended to phase out instruction in producibility and manufacturing in general. As a result of the "applied science" binge of many of the schools, such knowledge is rare indeed among young engineers and most of them have been successfully brainwashed out of the field. As a teacher of both operations research and manufacturing processes, the writer can only note somewhat ruefully that at a time when the mathematical models of production systems become ever more elegant, and operations research taxes the greatest capabilities of our computers, such relatively simple and classical subjects have been neglected in the curriculums. In part as a result of this development, other countries are either abreast of what the U.S. is doing, or are even ahead. There may be a momentary out-of-phase situation but it does not last and is not, in any event, necessarily in "our" favor. If it is true, for instance, that some electronic manufacturing methods still elude the Japanese at this point, it is also true that their steel industry utilized oxygen lances and continuous casting more rapidly than the U.S. did.

The third element in the crisis is simply that the "fallout" of military technology has failed to materialize to the extent claimed or hoped for. But what should one have expected? Surely no industry can seriously rely on such a circuitous route to the future. A nonelectronic

FIGURE 2. Proportion of selected occupations in defense-generated jobs, 1968. The chart shows the six occupations most extensively involved in defense work and related to electrical/electronic engineering. Engineers as a whole have 20 percent of their number in defense work and most technically related craftsmen are similarly affected. (Source: M. A. Rutzick, "Skills and locations of defense-related workers," Monthly Labor Rev., vol. 93, pp. 11–16, Feb. 1970)

example will illustrate the point: Heart valves are made of a raw material similar to that used for ladies' girdles, but nobody argues that the way to get better heart valves is to do more girdle research.

We have thus identified (1) the conversion problem as a need for U.S. electronics firms to consider new products and market strategies, (2) the foreign competition problem in terms of a loss of technical and managerial strength in a once highly successful industry, and (3) a quality problem to which the resources of the industry should be able to respond. Clearly, the effective implementation of conversion can free exactly those resources required to come to grips with the other two problems. The mechanics for doing that will be discussed in the following.

Organizing a solution

To coordinate and, in part, finance the kind of technological operation called for, it is proposed to establish a National Technology Foundation, which would act as a sponsoring agency for research on raw materials, design methodology, and applications, and would also assume that portion of "pure" research now sponsored by the Department of Defense that does not properly belong with the National Science Foundation. Its work agenda could tentatively cover the following, much of which is directly applicable to the electronics industry and, indeed, to all electrical manufacturing:

1. Material and product design studies with special reference to
   - (a) Economics and cost effectiveness of materials, including substitutability.
   - (b) Design methodology, e.g., stress analysis in products in which it is not now widely used, such as enclosures of equipment; circuit analysis and development, including the use of computers.
   - (c) Systematic application studies of new materials and methods to replace the present random and inefficient technological fallout, e.g., integrated and molecular circuits.
   - (d) Reliability and robustness studies based on the kinds of test programs proposed but seldom fully carried out on materials of all kinds[14]; applications to electronic circuitry.
2. Standardization and simplification
   - (a) Feasibility studies of modularization in producer and consumer durables.
   - (b) Cost studies, e.g., on economy of scale.
   - (c) Organizational studies for the systematic manufacture of modules, including the legal environment.
3. Processes
   - (a) Automation systems for manufacturing and service industries.
   - (b) Maintenance and safety problems, including new automation systems justifiable mainly for safety reasons.
   - (c) Productivity and producibility studies, both intra-industry and international.
   - (d) Capacity studies on machine tools and metal-working machinery and on electronic component/circuit manufacture, including new material–machine configurations.
   - (e) Rationalization and reduction of job shop manufacturing; development of versatile small-quantity production systems, "group tech-

nology"[15] (the use of common physical characteristics in order to group the products into manufacturing subsystems; e.g., all shafts are made on one group of machines).

4. Information gathering and processing
   (a) Information systems on design.
   (b) Cost-effectiveness registry of materials, subsystems, components, etc.
   (c) Operating models for businesses, including inventory, market studies, etc.

This list is not meant to be exhaustive, but it does provide a useful beginning. Several of the items specified are likely to become of greater interest in the near future. For example, safety on the job promises to be the focus of some major legal initiatives, including a new Federal safety law complete with standards. Obviously, one way of solving safety problems is by reducing the number of people doing the jobs concerned, i.e., by automating. Another solution lies in the improvement of reliability. All these should be issues within the purview of the NTF.

The actual work could be done by qualified defense contractors, but the NRF's purpose would lie in finding jobs for people rather than assuring the survival of corporations, which, as has been noted, have numerous ways of staying afloat. Certain universities, as well as research institutes (including some new ones), would qualify. Such organized research is in line with the Japanese examples mentioned earlier as well as with similar efforts in various branches of industry in Great Britain, the Netherlands, and West Germany. The communist countries have long used such forms of research organization and these efforts are in part responsible for the formidable Soviet machine tool industry, which makes extensive use of standardization, modularization, and cost-effectiveness studies. Further expansion of such efforts was proposed in a recent study,[16] the general findings of which are worthy of careful consideration in other industries.

The results of NTF's work would be freely accessible, with patents available to everyone for licensing at low rates, and all receipts going into a special NTF trust fund. This may sound somewhat self-defeating in the sense that the whole scheme is designed to solve problems first of all for U.S. industry, but, actually, the argument for exclusiveness does not have much merit in view of the extensive international cross-licensing and information exchange that is already extant. In any event, such arrangements tend to be two-way streets. Further, many of the technical–economic findings of the NTF would necessarily be based on U.S. costs and other conditions and thus would have limited usefulness elsewhere.

The employment that NTF contracts would provide to (hopefully) some of the better-qualified defense scientists and engineers might be only temporary. Certainly, those involved should be encouraged to seek permanent jobs among organizations with whom they deal while working on an NTF research project. This situation is not dissimilar to that of industrial consultants who regularly find executive positions with their clients. And at least the NTF assignments should provide on-the-job training in a context of technical progress, especially with respect to commercial producibility. If, as a result of such efforts, cost consciousness is raised from the image of sordid money grubbing that it now has among many "sophisticated" engineers, so much the better. In fact, it will be quite essential for this to happen because, in large measure, indifference to costs lies at the root of our product and trade problems.

As to the financing of the effort, patent license fees should help, but some direct governmental support, at the beginning at least, will be necessary. Contributions by defense firms or transfer payments by the Department of Defense should be considered as additional sources of funds. During hearings on Sen. George McGovern's proposed National Conversion Act (S.1285) on December 1 and 2, 1969, Walter Reuther proposed that one quarter of all profits of defense contractors should be earmarked for conversion planning. A levy on defense contractors is thus not a new idea.

The concept of the NTF as outlined here differs from various proposals for information systems and studies to deal with the general impact of technological progress. One such proposal, also by Walter Reuther, is for an "early warning system" to "gather information on a continuing basis [on] automation, atomic and solar energy, new materials, new products and other technological innovations." The areas of special interest would be the impact on employment, industrial location, industrial trends, educational needs, and international trade.[17] The NTF proposal goes much further, however. It suggests the actual implementation of technological needs as perceived, rather than as a response to existing problems, supplemented by such technological forecasting as may be feasible. The NTF also would be a contracting agency, doing the jobs for which firms or other institutions, acting singly, do not have the resources. Some of the clearinghouse functions of the agency that Reuther proposes (and which the NTF would also have to encompass) already exist in the Business and Defense Services Administration of the U.S. Department of Commerce.

Obviously, the NTF and similar concepts are not the only answer to the problem of conversion. Some firms will speedily change to commercial production, although whether in sufficient volume may be doubted. Individuals now in the defense industries may be able to find other jobs; this applies, for instance, to some skilled workers and to much of the clerical force. Others, in administrative tasks, may be able to participate in urban–surburban–rural planning and reconstruction. Finally, if a missile engineer should wish to become a medical doctor, for example, he should be encouraged to study for this, if otherwise qualified, with such scholarship support as may be required. The present article thus limits itself to only one of several possible modes of solution, which are by no means mutually exclusive.

Caveats and dangers

The problems and remedies discussed here have been identified before and the response to them has often been a mixture of cynicism, indifference, and defeatism that has made concerted action unusually difficult. This is quite apart from ideological "gut reactions," which, on many occasions, have inhibited even a discussion of the problems. Firms, unions, and communities have rarely put forward responsible plans, and, since they felt they had nothing to contribute, have refused even to attend conferences on these subjects, particularly on conversion. When there is nothing to show, such gatherings tend to become, in the sense of Hans Christian Andersen, emperors' fashion shows.

One aspect of the defeatism extends throughout the spectrum of problems discussed here: It is that U.S. costs and general business profligacy are so ingrained and out of line that nothing can be done. Of course, it is quite true that the defense industry has shown a conspicuous inability to meet limitations of time and budgets.[18] It is not unreasonable to view with suspicion the creators of the B-58, the F-111, the XB-70, the C-5A, and all the other weapons that failed to work properly, cost too much, and took too long; but that is the basic dimension of the conversion job, and the technical resource that we must now direct to other urgent tasks.

Finally, one can respond to an attitude of helplessness by considering the alternatives—and these are unpleasant indeed. Conversion could be handled simply by firing those affected and accepting the monumental dislocations this would entail. With the multiplier effects of such changes, many communities would be especially hard hit because conversion and the competition problem share the characteristic of attacking specific products and localities. It is easy to visualize political pressures against any defense cuts, and further growth in domestic extremism. It will be particularly necessary to guard against the kind of product change *within* defense work that would use up even more technical and research resources. It is in precisely that area that the shortage of talent is greatest and nothing can be done about international competition without investing exactly that kind of resource.

The more particular and immediate response to an import problem is, of course, protectionism. It has taken more than a generation to reduce the tariff and autarchy walls of the 1920s and 1930s to the more manageable scale of the General Agreements on Tariff and Trade (GATT). In any event, the U.S. electronics industry does not speak with a united voice on the subject because so much of it already makes use of plants outside the country; even component manufacturers, who are normally the most protectionist sector of the industry, have begun to "follow their markets abroad." Besides, protectionism would do nothing to help U.S. electronic exports; on the contrary, by sheltering domestic inefficiencies, it would soon reduce exports to the vanishing point, except for a few special products. And even these may be expected to decline as we become ever less able to cope with the quality problem. On that issue one can also envisage a future full of litigation (including soldiers suing the manufacturers of defective weapons!) and legal controversy, in the absence of technical efforts of the kind discussed here.

Clearly, there are major planning and action problems ahead, but, as shown here, it is possible to have the problem of conversion provide at least a partial answer to the problems of meeting foreign competition and of quality control. As Shakespeare says in *As You Like It:*

> "*Sweet are the uses of adversity:*
> *Which, like the toad, ugly and venomous,*
> *Wears yet a precious jewel in its head.*"

REFERENCES

1. For a report on this development, including its political implications, see Rovere, R. H., "Letter from Washington," *New Yorker*, pp. 169–177, Nov. 1, 1969.

2. Albright, J., "Laird plans a 40% troop cut," *Newsday*, p. 4, Dec. 22, 1969.

3. "Military markets start to sag," *Business Week*, p. 140, Oct. 25, 1969; see also *New York Times*, p. 1, Oct. 27, 1969, and p. 1, Oct. 22, 1969.

4. "Budget knives are nicking NASA," *Business Week*, pp. 21–22, Jan. 3, 1970.

5. See, for example, Melman, S., *The Depleted Society*. New York: Holt, 1967. See also the series of studies done by the U.S. Arms Control and Disarmament Agency: *Defense Industry Diversification*, 1966; *The Implications of Reduced Defense for the Electronics Industry*, 1965; *The Transferability and Retraining of Defense Engineers*, 1967. The studies are of uneven merit and, in the view of the writer, tend to minimize the problems involved.

6. Berkowitz, M., *Conversion of Military Oriented Research and Development to Civilian Uses;* Melman, S. (ed.), *The Defense Economy;* Ullmann, J. E. (ed.), *Potential Civilian Markets for the Military Electronics Industry*. New York: Praeger, 1970. These are three of five volumes in a series, *Conversion of Industry from a Military to a Civilian Economy*.

7. *Survey of Current Business*, Sept. 1969, p. 37.

8. U.S. Dept. of Commerce Foreign Trade Rept. FT 455 (1968 Annual) and FT 155 (1968 Annual).

9. *Electronic Trends International*, vol. 3, May 1969 (pt. II) and July 1969.

10. Japan Electronic Industry Development Assoc., *Annual Report 1968*. Tokyo: Kikai Shinko Kaikan, 1969.

11. Sugata, E., and Namekawa, T., "Integrated circuits for television receivers," *IEEE Spectrum*, vol. 6, pp. 64–74, May 1969.

12. "Now Japan discovers IC," *Business Week*, p. 39, Dec. 13, 1969.

13. "Singapore: the robust waif," *Business Week*, pp. 66–67, Dec. 6, 1969.

14. Freudenthal, A. M., and Gumbel, E. J., "Minimum life in fatigue," *J. Am. Statist. Assoc.*, vol. 49, pp. 575–597, 1954. Much of the extensive work on this subject relies on a few limited series of experiments.

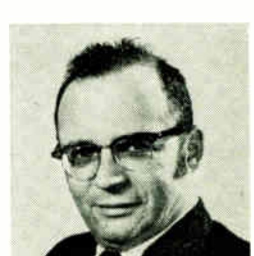
15. "A way to make diversity pay off," *Business Week*, p. 152, Oct. 18, 1969.

16. L'vov, D. S., *Principles of Engineering Design Economics*. Boston Spa, England: National Lending Library for Science and Technology, 1968.

17. Testimony of Walter Reuther before the U.S. Congress Joint Economic Committee, "The military budget and national priorities," 91st Congress, 1st Session, vol. 1, p. 438, June 1969.

18. For an early review, see Scherer, W., and Peck, M. J., *The Weapons Acquisition Process*. Cambridge, Mass.: Harvard University Graduate School of Business, 1960, chap. 3. For a recent critique see Stubbing, R. A., *Improving the Acquisition Process for High Risk Electronics Systems*. Distributed by U. S. Dept. of Commerce, Clearinghouse for Federal and Technical Information, Springfield, Va., No. AD 681 518, 1969.

John E. Ullmann is professor of management and chairman of the Department of Management, Marketing and Business Statistics at Hofstra University in Hempstead, N.Y. He received the B.Sc. degree in mechanical and civil engineering from the University of London in 1948 and the M.S. and Ph.D. in industrial engineering from Columbia University in 1951 and 1959 respectively. He is a licensed professional engineer in New York and a member of the National Labor Panel of the American Arbitration Association. Before his current appointment, he taught at Stevens Institute of Technology and has been visiting faculty member at New York University and Columbia. After 14 years as plant engineer and project engineer in industrial construction, Dr. Ullmann turned to various technical–economic planning problems. He has been a consultant to private firms and government agencies on design economics, industrial marketing, plant location, industrial development in an urban setting, transportation, and statistical problems. He has written many papers and is co-author of "Manufacturing Management: An Overview." In 1964, he directed a study of the conversion prospects of the defense electronics industry.



Piezoelectric effects in ferroelectric ceramics

Ferroelectric ceramics, the most strongly piezoelectric of all materials, exhibit a number of useful effects. Being nearly isotropic, these materials lend themselves to simple and revealing analytical treatment

Richard Holland **Jet Propulsion Laboratory**

Ferroelectric ceramics exhibit piezoelectric effects and optical birefringence that are electrically controllable. Consequently, these ceramics are applicable to a wide variety of devices. An article in the February issue has already described electrically actuated light valves and electrically controlled optical filters. The present article deals with the piezoelectric behavior of ferroelectric ceramics. Since, macroscopically, these materials are nearly isotropic, a discussion of their piezoelectric behavior can provide uncluttered—and surprising—insight into the theory of piezoelectric energy conversion.

Ferroelectric materials derive their name from a behavior analogous to that of ferromagnetic materials. In particular, for ordinary room temperature, the large-signal relation between electric field E and electric polarization P is hysteretic (Fig. 1).

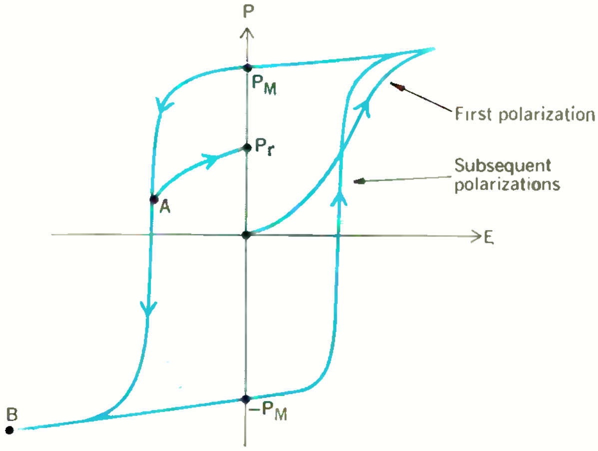


FIGURE 1. Electric field vs. polarization for a ferroelectric ceramic. This result pertains to temperatures below the Curie point, where hysteresis exists.

At high temperatures, polarization in a ferroelectric tends to have nearly linear dependence on the electric field. Only when the temperature is below some threshold value—called the Curie temperature or Curie point—does the dielectric behavior become hysteretic. The Curie point varies for different materials.

A ceramic that is ferroelectric is an aggregate of many tiny crystalline grains, randomly oriented with respect to one another. The origin of the hysteresis in a ferroelectric ceramic can be understood by considering the crystal structure of one of these grains. Figure 2 shows a unit cell of barium titanate, a typical ferroelectric, above its Curie point (120°C). This unit cell is cubic, with a titanium ion at the body-center position.[1] Below the Curie point, however, the titanium ion is offset from the body center toward one of the six face centers, and because the titanium ion has a net charge, the cell has a dipole moment. The cell below the Curie point is not truely cubic; it is about one percent longer in the dipole direction than along the other two axes.

Grain sizes are frequently in the range of 1 to 10 μm. Within a grain, adjacent cells tend to be oriented in the same direction. Groups of adjacent cells oriented alike are called "domains." Domains are likely to be smaller than grains, because free energy of the grain is higher if the same alignment of dipoles extends throughout the

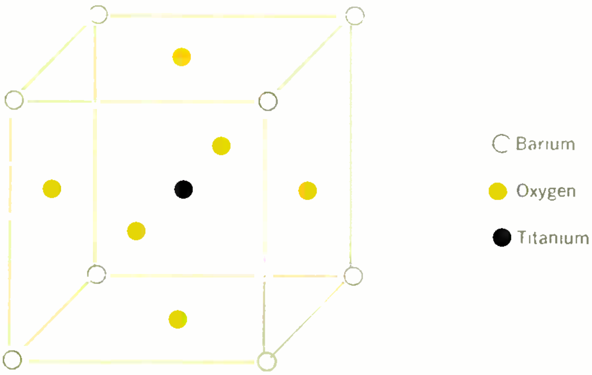


FIGURE 2. A single unit cell of barium titanate at temperatures above 120°C (the Curie point). In this temperature region, barium titanate forms a cubic crystal.

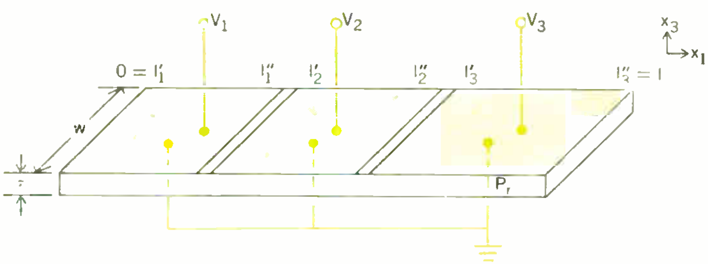


FIGURE 3. A ferroelectric ceramic bar with three electrode pairs along its length. Remanent polarization is parallel to x_3.

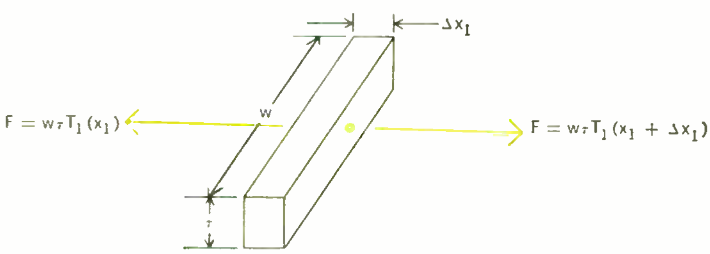
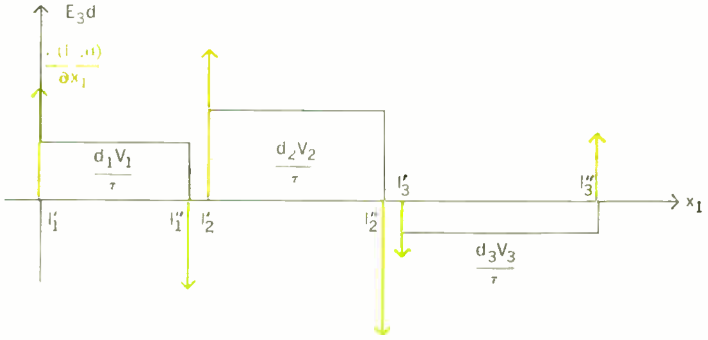


FIGURE 4. A thin slice taken out of the ceramic bar of Fig. 3.

FIGURE 5. The functions E_3d and $\partial(E_3d)/\partial x_1$ vs. x_1 for the bar in Fig. 3. The quantity E_3d is a set of three plateaus, and $\partial(E_3d)/\partial x_1$ is a set of six impulse functions.



grain.[2] Thus the grains of a ferroelectric ceramic that has never been subjected to a large electric field are usually broken up into a multiplicity of domains, oriented nearly at random; on a macroscopic basis, the ceramic in this condition must be isotropic.

When a very large electric field is applied, domains with favorably oriented dipoles tend to grow in size at the expense of all the other domains. This is an irreversible change, in that the titanium ions do not all return to their original locations when the large electric field is removed. Consequently, the material retains a net remanent polarization (P_M in Fig. 1) and a macroscopic elongation, even when the large polarizing field is no longer applied. Also, certain dielectric and elastic properties become anisotropic in a polarized ceramic. However, on a macroscopic basis, the ceramic properties must retain a plane of isotropy normal to the macroscopic axis of remanent polarization.

It is possible, of course, to obtain a remanent polarization P_r that is intermediate between the maximal values, $+P_M$ and $-P_M$, in Fig. 1. This state is achieved by supplying the polarizing charge through a very-large current-limiting resistance (say several thousand megohms) and then turning off the trickle of current at the instant the polarization reaches point A on the hysteresis curve in Fig. 1. In this way, all the domains will not complete their switching to point B. The ratio P_r/P_M is called the normalized remanent polarization. This ratio, which can vary between $+1$ and -1, is a fundamental variable characterizing the condition of the ceramic.

Definition of piezoelectricity

Consider the mechanically free ferroelectric ceramic bar shown in Fig. 3 with three pairs of electrodes on its major faces. (We can introduce the basic concepts of piezoelectricity by describing the length-extensional response of this bar under small-signal conditions.) Let us assume the material previously has been polarized in the thickness or x_3 direction by application of large voltages across the three electrode pairs. Also assume the width w and the thickness τ of the bar are much less than the bar length l and the acoustic wavelength.

Particle displacement in the x_1 direction will be represented as u_1. The first component of strain is then defined as the derivative of u_1 with respect to x_1

$$S_1 = \frac{\partial u_1}{\partial x_1} \qquad (1)$$

This is the only strain component that will be of concern to us here. Stress T_1 in the x_1 direction is defined as the force per unit area acting in the $+x_1$ direction on a cross section of the bar taken normal to x_1. In an ordinary, or nonpiezoelectric, linear elastic material, stress and strain are related by Young's modulus

$$S_1 = T_1/Y \qquad \text{(Hooke's law)} \qquad (2)$$

Similarly, in a linear nonpiezoelectric material, D_3 and E_3 (the electric displacement and field in the thickness direction) are linearly related by the dielectric constant*

* If multidimensional effects, which are insignificant in this discussion, were to be included, we should designate our Y as Y_{11} and our ϵ as ϵ_{33}. In the presence of anisotropic effects, Y_{33} and ϵ_{11}, the other Young's modulus and dielectric constant, will differ from Y_{11} and ϵ_{33}.

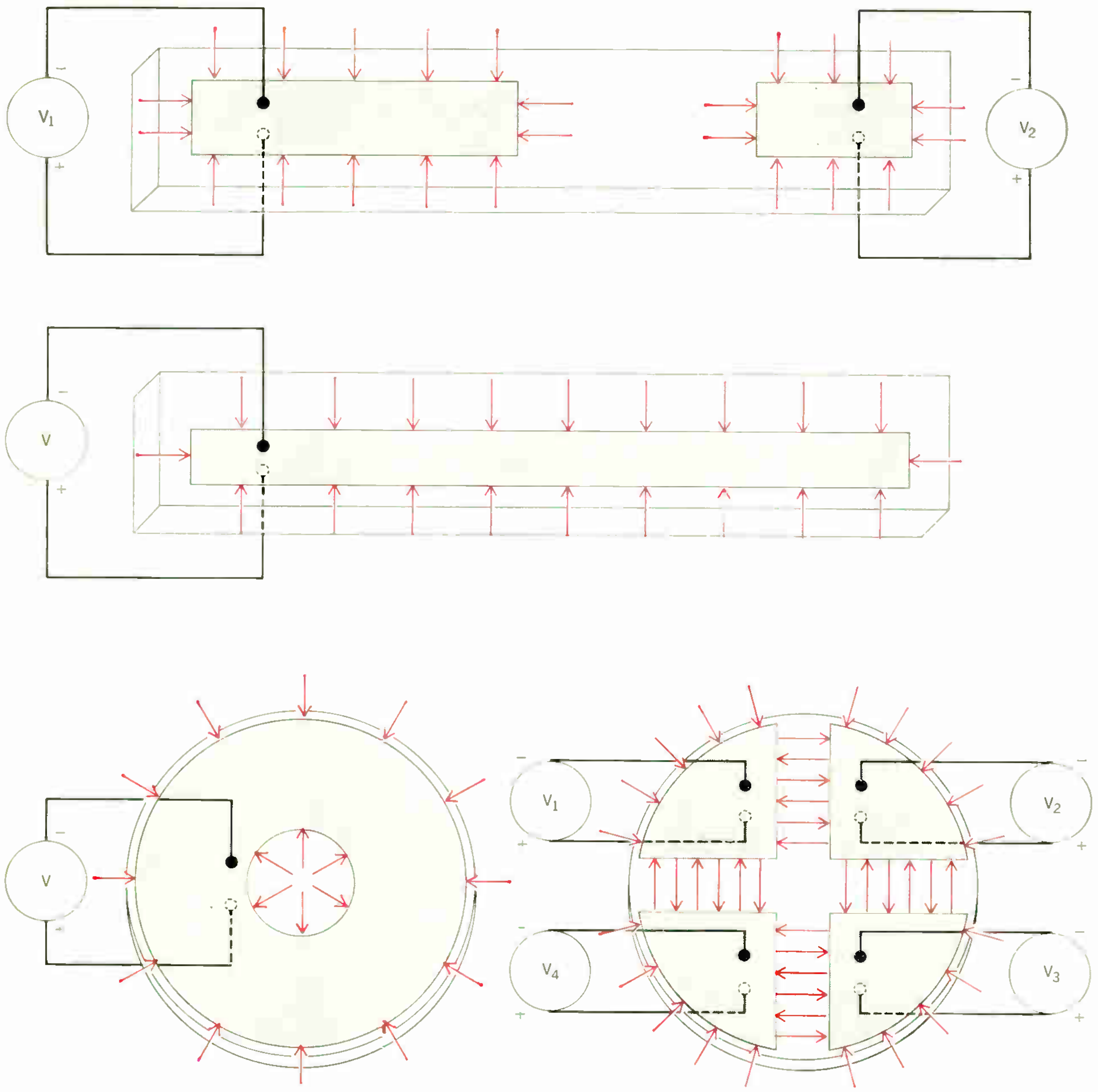


FIGURE 6. The piezoelectric force distribution for arbitrarily electroded and shaped ferroelectric ceramic wafers.

$$D_3 = E_3\epsilon \tag{3}$$

In linear piezoelectric materials, the constituent equations (2) and (3) become coupled by the piezoelectric coefficient d,

$$S_1 = T_1/Y + E_3 d \tag{4a}$$

$$D_3 = T_1 d + E_3\epsilon \tag{4b}$$

The equality of the two d coefficients in (4a) and (4b) is guaranteed by one of the thermodynamic Maxwell relations.[3]

Unpolarized ferroelectric ceramics do not exhibit piezoelectricity under small-signal excitation. However, after being polarized, ferroelectric ceramics generally possess small-signal piezoelectric properties, i.e., d in Eqs. (4) is nonzero. In addition to piezoelectricity, these ceramics have a number of other peculiar and useful attributes. For example, the small-signal dielectric coefficient ϵ is usually at least several hundred times greater than ϵ_0, the permittivity of free space. Also, the three small-signal coefficients Y, d, and ϵ in Eqs. (4) are functions of remanent polarization (P_r in Fig. 1). Generally speaking, Y and ϵ are nearly even functions of P_r, and vary about 15 percent between $P_r = P_M$ and $P_r = 0$. On the other hand, d is nearly an odd function of P_r, and changes sign near $P_r = 0$.

Sources of sound in piezoelectric materials

Let us assume the bar shown in Fig. 3 is excited sinusoidally at a frequency for which width and thickness, but not length, are much less than an acoustic wavelength. If we neglect electrical fringing, the problem is one-dimensional, with x_1 the only coordinate in use.

Consider a thin cross-sectional slice of this bar having length Δx_1, as shown in Fig. 4. The net force on this slice in the $+x_1$ direction is

$$F = w\tau \frac{\partial T_1}{\partial x_1} \Delta x_1 \tag{5}$$

This net force will induce an acceleration $\ddot{u}_1$ in the x_1 direction that is given by

$$\ddot{u}_1 = \frac{F}{\Delta x_1 \rho w \tau} = \frac{\partial T_1}{\partial x_1} \cdot \frac{1}{\rho} \tag{6}$$

where ρ is the density of the bar. If Eqs. (1) and (4a) are combined with (6), one obtains the wave equation for particle displacement along x_1

$$\frac{\partial^2 u_1}{\partial x_1^2} - \frac{\rho}{Y} \ddot{u}_1 = \frac{\partial(E_3 d)}{\partial x_1} \tag{7}$$

The component on the right of this equation represents the forcing or driving terms. If this term is zero, i.e., if d or E is zero over the entire bar, no mechanical motions are piezoelectrically induced.

This piezoelectric driving term, $\partial(E_3 d)/\partial x_1$, may be represented by an equivalent mechanical force. The result of this representation is most curious. In particular, if we have a nonpiezoelectric bar subjected to a force density $f(x_1)$ acting on the bar in the $+x_1$ direction, particle displacement obeys the equation

$$\frac{\partial^2 u_1}{\partial x_1^2} - \frac{\rho}{Y} \ddot{u}_1 = -\frac{f(x_1)}{Y} \tag{8}$$

Comparison of (7) and (8) thus indicates that the piezoelectric drive $\partial(E_3 d)/\partial x_1$ may be represented equivalently by a force density distribution of

$$f(x_1) = -Y \frac{\partial(E_3 d)}{\partial x_1} \tag{9}$$

Stop for a moment and consider the nature of this force distribution. The three electroded regions on the bar in Fig. 3 may differ in applied voltages V_i and remanent polarizations P_{ri} (P_{ri} is P_r of Fig. 1 at electrode i). However, within each of the three regions, V_i and P_{ri} are normally constant. Moreover, d_i (d at electrode i) depends only on P_{ri}. Thus, d_i also will normally be constant within each of the three regions. This means that $E_3 d$ as a function of x_1 will be a set of three plateaus, as illustrated in Fig. 5.

But Eq. (9) indicates that the derivative of $E_3 d$, not $E_3 d$ itself, governs the equivalent force. As Fig. 5 illustrates, this derivative is a set of impulse-singularity or δ-functions, one acting at each electrode edge.*

In other words, the equivalent sources of sound in a piezoelectric bar at frequencies in the realm of length-extensional response are singularity functions located only at the electrode edges. Intuitively, one would expect this equivalent source to be distributed uniformly as a body force density under each of the activated electrodes.

* The idealization of no electrical fringing at the electrode edges is implicit in this description of the derivative $\partial(E_3 d)/\partial x_1$. However, as long as the actual fringing dimensions are less than $1/16$ of an acoustic wavelength, this idealization is an excellent representation of reality.

FIGURE 7. Algorithm for deriving the electrode admittance matrix of the bar in Fig. 3.

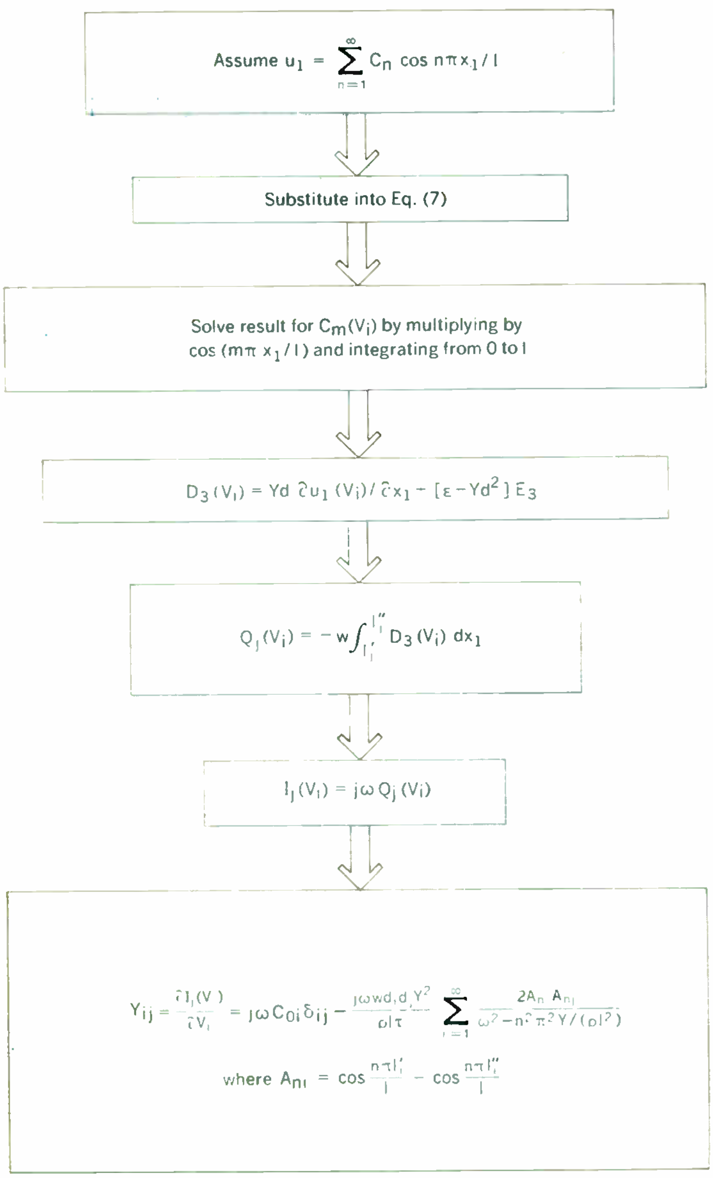
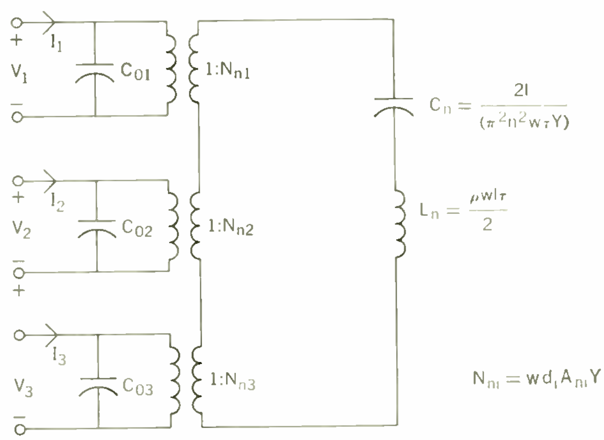


FIGURE 8. Equivalent circuit near acoustic resonance n for the ferroelectric ceramic bar in Fig. 3.



However, this problem is a striking instance of how intuition can lead one astray.

The concept that piezoelectric drive can be represented as an equivalent force acting only at the electrode edges may be extended from one dimension to two. In other words, in a thickness-polarized ferroelectric ceramic wafer of arbitrary shape and electrode arrangement, piezoelectric drive is still representable as a singularity force distribution acting only around the edges of each electrode. Figure 6 illustrates this statement—as we can see, the equivalent forces in the two-dimensional case are directed within the wafer plane and perpendicular to electrode-edge contours.[4,5]

This device of representing piezoelectric drive in thin plates by force singularities at the electrode edge applies to plates of other piezoelectric materials as well as to ferroelectric ceramics. However, in the case of plates made from less symmetric piezoelectric materials, a plane of material isotropy may not coincide with the plane of the plate. When this plane of isotropy is missing, the equivalent force singularities will not be oriented perpendicular to the electrode edges,[5,6] though their orientation will remain in the wafer plane.

Equivalent circuits and admittance matrixes

The complete linear response of the ceramic bar in Fig. 3 may be represented by the admittance matrix Y_{ij}, which interrelates the three currents and voltages. This admittance matrix may be derived in closed form by the algorithm diagramed in Fig. 7, which we shall now describe.

One begins by assuming that the particle displacement u_1 may be represented by a cosine series. That assumption is then substituted into Eq. (7). Next, the Fourier coefficients C_m are evaluated as functions of the applied voltages V_i from the resulting equation. This is done by multiplying the resulting equation by cos $(m\pi x_1/l)$ and integrating over the length of the bar. We now have the particle displacement as a function of the applied voltages, and of the bar dimensions and coefficients. Next, T_1 and S_1 are eliminated from Eqs. (1) and (4), so that D_3 may be expressed in terms of u_1 and E_3, which we already know as functions of the V_i.

We can then compute the free charge $Q_j(V_i)$ on an arbitrary electrode (say electrode j) as a function of the V_i by integrating $D_3(V_i)$ over that electrode. The current at that electrode $I_j(V_i)$ is merely the time derivative of the charge there. Finally, Y_{ij} is obtained by differentiating $I_j(V_i)$ with respect to V_i.

The actual equation for Y_{ij} given at the bottom of Fig. 7 has an interesting physical interpretation. The first term is just the ordinary driving-point capacitance of the various electrodes. It represents the admittance one would see for the bar in Fig. 3 if no piezoelectric effects were present. The second term in the equation represents the piezoelectric or motional component of the admittance. It consists of an infinite series of resonant terms: at each frequency for which the bar is an integral number of half acoustic wavelengths long

$$\omega_n = \frac{n\pi}{l}\sqrt{\frac{Y}{\rho}} \tag{10}$$

one term goes to infinity if losses are neglected.

The strength of resonance n as described by the formula for Y_{ij} depends on the product $(d_iA_{ni})\cdot(d_jA_{nj})$. Note that A_{ni} as defined in Fig. 7 is the difference in value of the displacement eigenfunction of mode n, u_1 = cos $(n\pi x_1/l)$, when evaluated at the two ends of electrode i. One can thus regard the motional component of Y_{ij} as a "communication" between electrode i and electrode j. The eigenmodes of the bar act as a "medium" for that communication. How strong the communication is via mode n depends on d_iA_{ni} and d_jA_{nj}. These quantities are measures of how strongly electrodes i and j respectively are piezoelectrically coupled to mode n.

A distributed-medium problem such as that of the bar we have been discussing cannot be represented over all frequencies by a finite number of lumped *R*s, *L*s, and *C*s. However, in most cases of practical interest, one operates a system such as this in the immediate vicinity of a resonance (say ω_n). In that case, all the terms in the infinite series for Y_{ij} are negligible in comparison with the nth term. Under these conditions, Y_{ij} is representable by the equivalent circuit shown in Fig. 8.

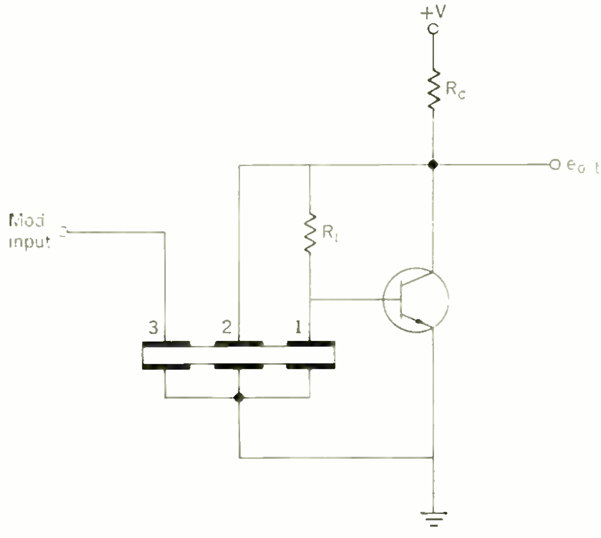
The techniques described here for analyzing multiple-electrode bars may be generalized to treat quite complicated two-dimensional plate structures such as those in Fig. 6. Details of this work are published elsewhere.[4-6]

Three-dimensional piezoelectric structures, however, are beyond the realm of generalizability of these techniques. The difficulty with three-dimensional structures is that one cannot decouple the electrostatic problem from the elastic problem, as we have done with the bar. Rather, the electrostatic and elastic problems must be solved simultaneously. Techniques for doing this will be described briefly later on.

Device applications

A number of device applications exist for the type of structure described in Fig. 3. One area of application pertains to the synthesis of high-Q high-frequency miniaturized tuned circuits. The straightforward approach to this synthesis problem would be to use miniaturized inductors and capacitors. Unfortunately, if an inductor is

FIGURE 9. An electrically tuned transistor oscillator with a three-electrode-pair ferroelectric ceramic bar in the feedback loop.



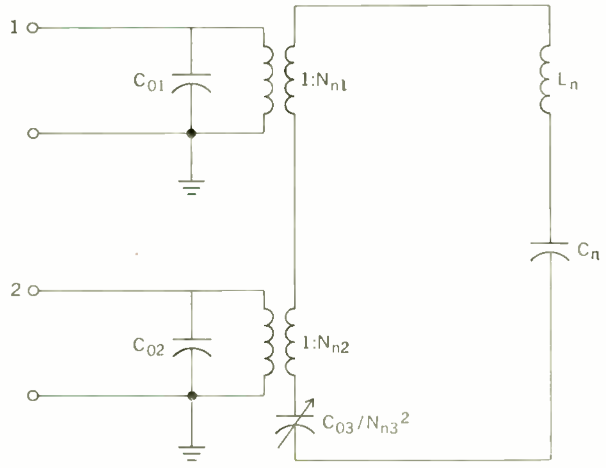
scaled down by a factor α, its loss tangent will increase as α^2, so that this approach quickly becomes unsatisfactory.[7,8] Using active circuits is not particularly desirable either, since this route leads invariably to grave stability problems.[7,8] Simple piezoelectric resonators, however, seem to be an ideal solution to the difficulty. Being passive, they are necessarily stable. Also, the loss tangent does not increase when the resonator is scaled down for operation at higher frequency. The sample in Fig. 3 would be classified as a simple piezoelectric resonator if just one electrode pair were present, rather than three.

Based on the equivalent circuit shown in Fig. 8, a second rather obvious application for this type of configuration is as a tuned transformer. Another use for multiple-electrode ferroelectric ceramic resonators is in electrically tuned oscillators.[9] An electrically tuned oscillator circuit of this type is shown in Fig. 9. The frequency of oscillation in this case will be near the resonance of the three-electrode-pair bar in the feedback circuit. As we shall now demonstrate, this resonance may be electrically shifted.

Assume electrode region 3 of the bar is connected to a high-impedance charge source (labeled Mod. input in Fig. 9), which can change the remanent polarization of the ceramic at that electrode pair. Let us now reflect the dielectric capacitance C_{03} of that electrode pair across its electromechanical transformer N_{n3}, so that the equivalent circuit of Fig. 8 becomes specialized to the present case (see Fig. 10). In Fig. 10, note that this reflected capacitance C_{03}/N_{n3}^2 now appears in series with the frequency-determining series resonant circuit, and raises the overall frequency. Moreover, varying the remanent polarization of the ceramic at electrode region 3 will alter the piezoelectric coefficient there (d_3), and hence, according to the formula for N_{n3} in Fig. 8, will alter the value of this reflected capacitance. For many ferroelectric ceramics, the possible range of N_{n3} in relation to C_{03} and C_n is such that 5 percent or more shifts in resonant frequency can be obtained in this way. Analogous variable-frequency three-electrode resonators in geometries other than a bar can utilize the piezoelectric effects of a ferroelectric ceramic more efficiently to obtain possible frequency shifts of 15 percent. An example of a resonator geometry falling into this more efficient category consists of a thin disk with three concentric electroded regions.*

Other examples of multiple-electrode ferroelectric devices utilizing piezoelectric effects include FM discriminators and nondestructive readout memory elements[5,10,11] Ferroelectric ceramic miniaturized components generally possess a great advantage over their semiconductor counterparts whenever radiation tolerance is required in addition to compactness. Nuclear weapons and space applications circuitry are two situations that require components with these attributes.

FIGURE 10. Equivalent circuit for the three-electrode-pair ferroelectric ceramic bar with remanent polarization of electrode-region-3 variable.



The Lagrangian formulation of piezoelectricity

As was stated earlier, three-dimensional piezoelectric problems are considerably more difficult mathematically than two-dimensional ones. This is true because the electrostatic and elastic components of three-dimensional problems are mixed in a way that does not permit their decoupling and separate treatment. Circular cylinders of commensurate diameter and thickness, and rectilinear parallelepipeds are two examples of this three-dimensional problem class.

If these complex three-dimensional configurations had no practical facets, we would not need to be concerned with their intractability. This is not the case, however. Thin ceramic plates or bars, where a third dimension is lacking, are necessarily fragile and easily shattered. The addition of a reasonable thickness dimension greatly improves the structural integrity. This is especially important when the device may be subjected to extreme mechanical shocks of acceleration—as is true in a nuclear weapon or in a space system.

A method has recently been developed for the theoretical study of these difficult three-dimensional piezoelectric geometries.[12-16]† This method is basically a variational or Lagrangian formulation.

Four space variables are necessary to describe uniquely the electromechanical state of a piezoelectric solid: three components of particle displacement, u_1, u_2, u_3; and the electrostatic potential ϕ. A piezoelectric Lagrangian density $\mathcal{L}$ is some function of those four variables that has the following property:

Define the integrated Lagrangian L as the integral of $\mathcal{L}$ over the piezoelectric solid in question

$$L = \iiint_V \mathcal{L}(u_1,u_2,u_3,\phi)\, dV \tag{11}$$

Further add arbitrary infinitesimal, but twice differentiable, perturbations to the four independent variables

$$\begin{aligned} u_1 &\rightarrow u_1 + \delta u_1 \\ u_2 &\rightarrow u_2 + \delta u_2 \\ u_3 &\rightarrow u_3 + \delta u_3 \\ \phi &\rightarrow \phi + \delta\phi \end{aligned} \tag{12}$$

Finally, compute the change in L caused by these perturbations,

$$\delta L = \iiint_V [\mathcal{L}(u_i + \delta u_i,\phi + \delta\phi) - \mathcal{L}(u_i,\phi)]\, dV \tag{13}$$

* See Ref. 5, sect. 2.2.2, p. 60.
† See Ref. 5, chap. 4, pp. 125–178.

If a necessary and sufficient condition that δL vanish for any and all perturbations δu_i and $\delta\phi$ is that u_i and ϕ obey the three-dimensional elastic-wave and electrostatic equations and boundary conditions, then $\mathfrak{L}$ is a Lagrangian density.

It turns out that the density function that is represented by

$$\begin{aligned}\mathfrak{L} = &\text{ Kinetic energy density } (\tfrac{1}{2}\rho\dot{\mathbf{u}}\cdot\dot{\mathbf{u}})\\ &+ \text{ dielectric potential energy density } (\tfrac{1}{2}\mathbf{D}\cdot\mathbf{E})\\ &- \text{ elastic potential energy density } (\tfrac{1}{2}\mathbf{S}:\mathbf{T}) \qquad (14)\end{aligned}$$

meets the above condition for being a piezoelectric Lagrangian density.

This sophisticated Lagrangian formalism is at present the ultimate theoretical tool of linear piezoelectricity. With its use, formulas have been derived for the admittance or impedance matrix of an arbitrarily shaped and electroded piezoelectric body of arbitrary composition and inhomogeneity.[13,*] Its use also permits computation of the natural acoustoelectric modes and resonance spectrum of such an arbitrary piezoelectric body.

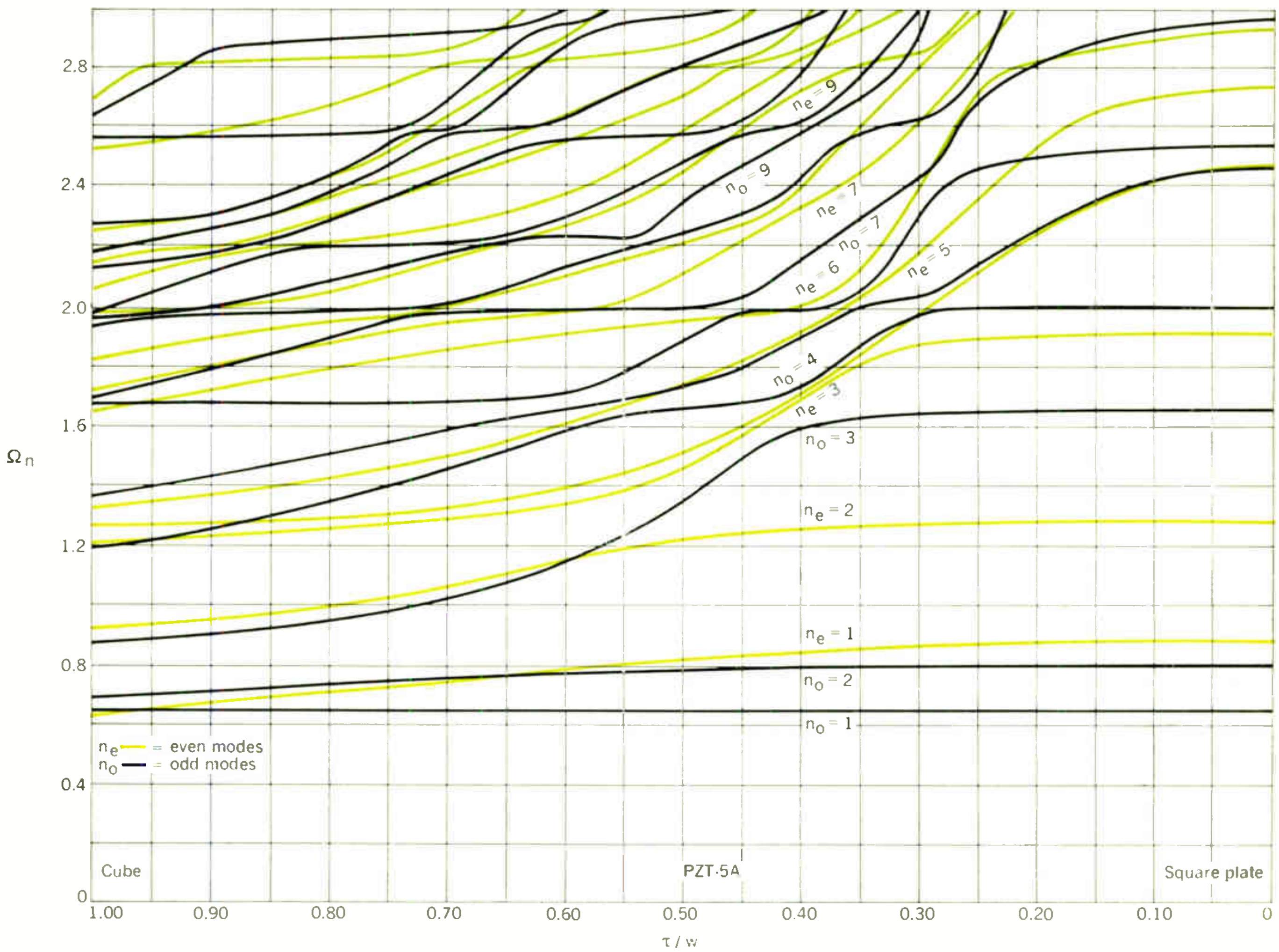
* See Ref. 5, chap. 4, pp. 125–178.

A ferroelectric rectangular parallelepiped

One of the most frustrating, yet conceptually simple, problems of theoretical mechanics is to determine the natural modes of a free elastic rectilinear parallelepiped, even in the absence of piezoelectricity. As an illustration of the effectiveness of the piezoelectric Lagrangian formulation, this formulation has been used to determine the natural modes and spectrum of a ferroelectric ceramic parallelepiped with piezoelectric effects included.[17]

Ferroelectric ceramic parallelepipeds can support acoustic modes that, depending on whether u_i is an even or odd function of (x_i,x_j,x_k), are separable into eight "symmetry families." The family in which u_i is odd in x_i but even in x_j and x_k corresponds to *dilation* (as opposed to shear, flexure, shear-flexure, dilation-flexure, etc.). It is this dilation family we have selected for study here. In order to restrict the problem to a system with only one variable parameter, let us further concentrate just on parallelelpipeds of arbitrary height with square cross section. Height will be assumed to be along x_3, so that the square cross section is in the x_1x_2 plane. The height-to-width ratio is then the desired single variable. Figure 11 shows the spectrum of dilation resonances

FIGURE 11. Normalized spectrum of the dilation family for a lead-zirconate-titanate ceramic (PZT-5A) square parallelepiped. The thickness-to-width ratio is varied from unity (cube case) to zero (thin square-plate case). Even modes are indexed by n_e, and odd modes by n_o.[17]



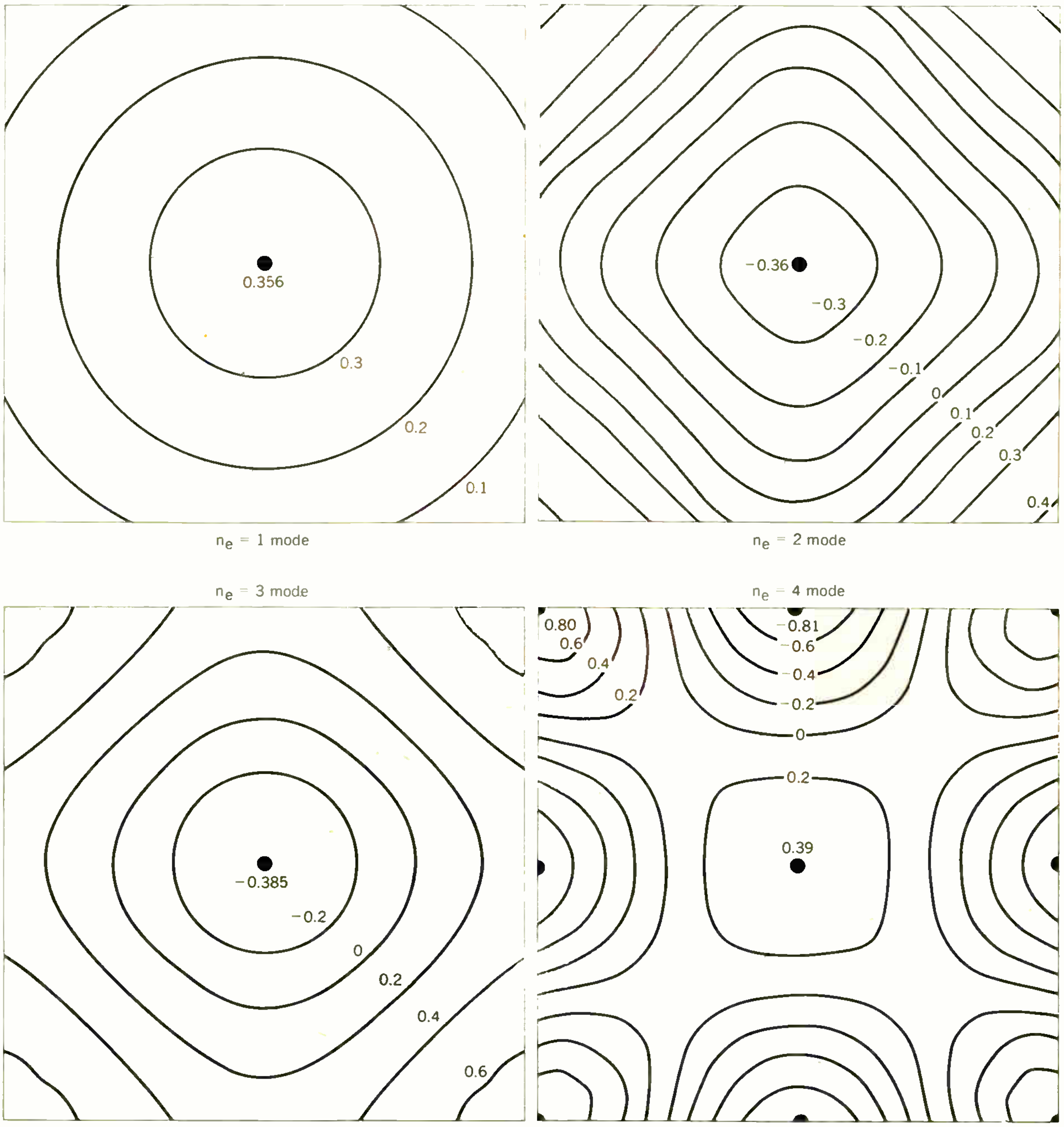


FIGURE 12. The u_3 displacements at $x_3 = 0$ of the first four even modes of a ferroelectric ceramic (PZT-5A) square parallelepiped with height/width (τ/w) ratio of 1/2[17] Note that rotation by 90 degrees leaves these mappings unchanged.

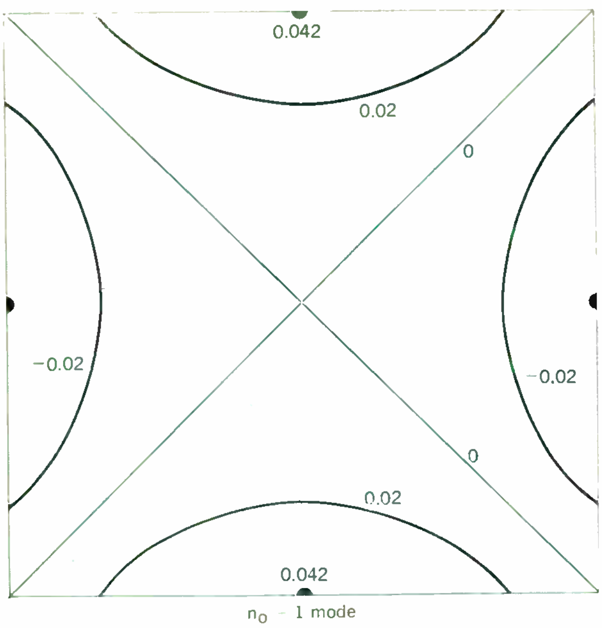
of such a parallelepiped as the height τ to width w ratio is varied from unity (cube case) to zero (thin square-plate case). The parallelepiped is assumed to be composed of the lead-zirconate-titanate ferroelectric ceramic PZT-5A[18, *] and to be poled in the height or τ direction. Normalized dimensionless frequencies Ω_n in this illustration are related to the actual resonant frequencies ω_n by

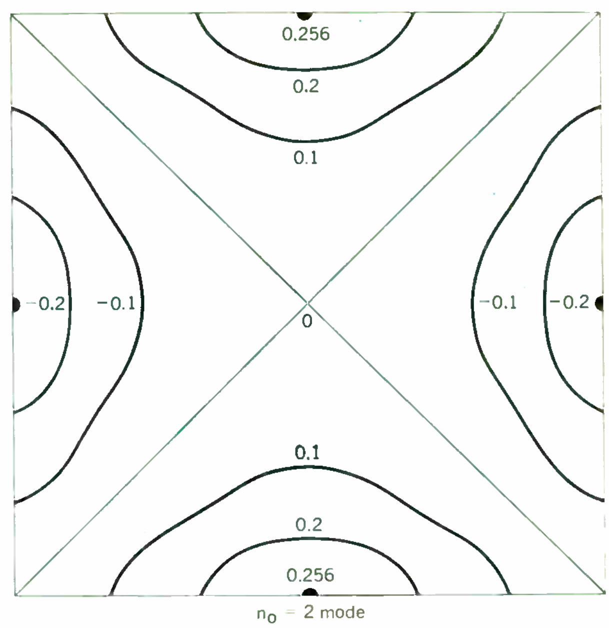
$$\Omega_n = \omega_n \frac{w}{\pi} \sqrt{\frac{\rho}{c_{33}^E}} \qquad (15)$$

Here c_{33}^E is an elastic stiffness or stress/strain ratio. It is defined as stress/strain for a plate compressed along the polarization direction with the resulting piezoelectrically induced E field short-circuited out, and with all motion in the plane perpendicular to the compressing stress prevented.[3, 19]

The dilation family of resonances in a square ferroelectric parallelepiped is split up into two subfamilies, the even and the odd modes, as Fig. 11 indicates. This distinction is illustrated by reference to Figs. 12 and 13: Figure 12 shows the u_3 displacements of the first four even modes at $x_3 = 0$; Fig. 13 presents the same information for the first three odd modes. These seven illustrations may be thought of as "fingerprints" of the seven

* Clevite Corporation trademark.





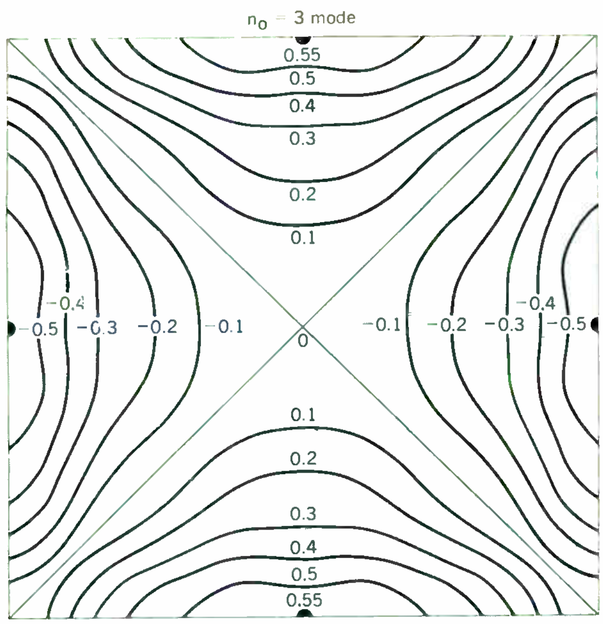
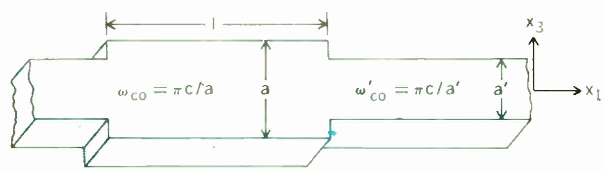


FIGURE 13. The u_3 displacements at $x_3 = 0$ of the first three odd modes of a ferroelectric ceramic (PZT-5A) square parallelepiped with height/width (τ/w) ratio of 1/2.[17] Note that rotation by 90 degrees effectively multiplies the data by −1.

FIGURE 14. An enlarged electromagnetic waveguide section bounded by two narrower sections. For ω between ω_{co} and ω_{co}', trapped waves can reflect back and forth in the enlarged section without penetrating into the narrower sections.



modes in question. Note that in the odd modes the main diagonals are nodal lines.

Figures 12 and 13 are also based on the lead-zirconate-titanate ferroelectric ceramic PZT-5A; the τ/w ratio selected for presentation in these diagrams is ½.

Energy trapping

High-frequency piezoelectric resonators usually are thin plates in which the thickness dimension τ is an odd integral number of acoustic half-wavelengths. The trouble with simple "thickness-mode" resonators of this type is that some very-high-order overtones of the length and width resonance usually lie quite close in frequency to the desired thickness response.[20] These adjacent resonances manifest themselves as spurious responses, which give rise to a cluttered admittance vs. frequency characteristic, rather than a single clean maximum at the thickness resonance.

In recent years, a principle called energy trapping has been applied to overcome this trouble, both in quartz[21-24] and ferroelectric ceramic[25,*] thickness-mode resonators.

This energy-trapping principle is most easily described by means of an electromagnetic waveguide analogy. Refer to the waveguide arrangement depicted in Fig. 14, where an enlarged waveguide section of width a is located between two narrower sections of width a'. In the enlarged central section, no electromagnetic waves will propagate at frequencies below the cutoff of the first TE wave

$$\omega_{co} = \pi c/a \qquad (16)$$

(Here, c is the speed of light.) However, at frequencies above ω_{co}, the first TE wave will propagate in the enlarged section. Similarly, in the narrower sections, no wave will propagate below the cutoff frequencies of the first TE wave in those sections

* See Ref. 5, chap. 5, pp. 179–224.

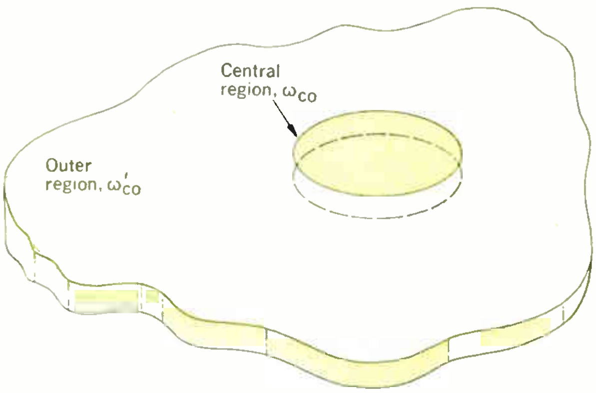


FIGURE 15. Basic acoustic energy-trapping configuration in a thin plate. The plate is so constructed that ω_{co}, the fundamental acoustic thickness-wave cutoff frequency in the central region, is lower than ω_{co}', the corresponding outer-region cutoff frequency.

$$\omega_{co}' = \pi c / a' \tag{17}$$

Note that ω_{co}' is a higher frequency than ω_{co}. Consequently, if one somehow manages to inject electromagnetic energy into the central waveguide section at a frequency between ω_{co} and ω_{co}', it will reflect back and forth in that section. However, it will be unable to escape into the narrower waveguide sections to the right or left. In other words, one can trap energy in the central section if all possible escape routes have a cutoff frequency above that characterizing the excitation.

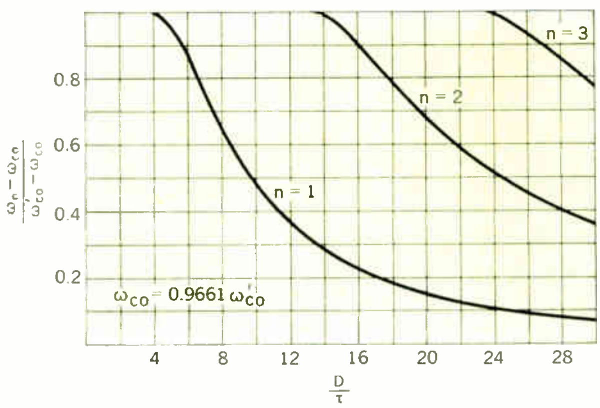
In the field of acoustics, this energy-trapping concept is utilized by constructing an elastic plate consisting of one central region completely surrounded by a second outer region (Fig. 15). The two regions individually are specified to be of uniform thickness. Elastic plates of finite thickness, like electromagnetic waveguides, can support thickness-mode waves. These acoustic waves either attenuate or propagate in the plane of the plate, depending on whether the exciting frequency is above or below some cutoff frequency. Consequently, if the central region of the plate is fabricated so that its lowest thickness-wave acoustic cutoff frequency ω_{co} is below the corresponding frequency ω_{co}' of the outer region, it is possible to trap acoustic energy in the central region. To do this, the exciting frequency, of course, must be between ω_{co} and ω_{co}'.

The question arises as to how one may achieve this lowering of the fundamental thickness-wave acoustic cutoff frequency in the central region. Two methods are in common use, frequently in combination.

The first of these methods is directly analogous to the electromagnetic case. One need merely thicken the central region relative to the outer region, and ω_{co} will automatically be reduced below ω_{co}'. If the plate is piezoelectric, this thickening may be achieved advantageously by depositing electrodes on the desired region. These electrodes will then serve the dual purpose of lowering ω_{co} relative to ω_{co}', and of providing electrical access for injecting or withdrawing energy at the trapping region.

The second method for lowering ω_{co} is restricted to piezoelectric plates. This method operates on a much less intuitive basis than the first method. In particular, if infinitesimally thin electrodes are applied to the central region and not to the outer region, one will find the desired lowering of ω_{co} is automatically achieved, even though the central region is not finitely thickened by the electrodes. The magnitude of this frequency lowering depends on the piezoactivity of the plate: For a weakly active material such as quartz, the lowering may be only 0.3 percent or less.[26] On the other hand, for a strongly active material such as some ferroelectric ceramics, the lowering may exceed 20 percent.*

FIGURE 16. Trapped-energy resonance spectrum of thickness waves for a PZT-5A plate (see Ref. 5, chap. 5, pp. 179–244). These results are based on a circular trapping region of diameter D and plate thickness τ. Note that the resonant frequencies ω_n are plotted as the fractional spacing of ω_n between ω_{co} and ω_{co}'.



Suppression of spurious resonances by energy trapping

Let us now describe the manner in which the energy-trapping principle may be used to prevent spurious responses in thickness-mode resonators. Consider again the electromagnetic waveguide arrangement in Fig. 14. At some frequency slightly above ω_{co}, one standing wave can exist in a resonant condition along the dimension l in the x_1 direction. At a somewhat higher frequency, two resonant standing waves are permitted; at a still higher frequency, three; etc. Eventually, of course, frequency will exceed ω_{co}', and then there can be no further standing-wave resonances, as the energy escapes as fast as it is injected. The total number of permitted standing-wave resonances of this type between ω_{co} and ω_{co}' is proportional to the length l of the enlarged region and to $(\omega_{co}' - \omega_{co})$, the cutoff frequency separation in the two regions. Thus, if $l(\omega_{co}' - \omega_{co})$ is small enough, only the first resonance of this series may exist.

An identical situation exists in the case of piezoelectric trapped-energy resonators. However, in this case, the resonances beyond the first show up as the undesired spurious responses mentioned in the previous section. Consequently, if one makes the electroded energy-trapping region small enough, all the spurious responses may be completely eliminated. This small-electrode condition is a generalization of the previously mentioned

* See Ref. 5, chap. 5, pp. 179–244.

concept that a short waveguide section of length l (Fig. 14), and hence a small $l(\omega_{co}' - \omega_{co})$, will eliminate all the trapped waveguide resonances except the first.

It is perhaps useful to illustrate these principles with a specific numerical example. Suppose the plate shown in Fig. 15 is composed of the ferroelectric ceramic PZT-5A. Further, suppose the central region is circular, of diameter D, and covered by infinitesimally thin electrodes. Let this region be maximally polarized in the thickness direction ($P_r = P_M$ in Fig. 1). Additionally, let the outer region be electrically depolarized ($P_r = 0$, Fig. 1). If the thickness of both regions is equal and of value τ, we will find $\omega_{co} = 0.9661\omega_{co}'$ for this arrangement.* In other words, the central region has a lower fundamental thickness-wave cutoff frequency than the outer region. Thus the condition for energy trapping is fulfilled by the electroded central region.

Figure 16 shows the trapped-resonance spectrum for this configuration as a function of the diameter-to-thickness ratio D/τ.* Note that below $D = 3.2\tau$ there are no trapped resonances, and between $D = 3.2\tau$ and $D = 13.0\tau$ there is just one trapped resonance. Consequently, in this configuration an electrode diameter D between 3.2τ and 13.0τ will yield a thickness resonator free of spurious secondary responses.

Multiple-electrode energy-trapping devices

In conclusion, let us discuss a class of advanced applications of the energy-trapping concept. Early energy-trapping devices were all simple thickness-mode piezoelectric resonators with single electrodes on the top and bottom of a plate. However, recently, more sophisticated devices have been developed with two or more trapping regions located in close proximity to each other. Such arrangements permit energy to "tunnel" from one trapping region to another. These multiregion trapped-energy configurations have enough variable parameters that it is possible to use them for synthesizing virtually any desired filter response. For example, an entire lattice filter section with a single resonant circuit in each branch may be replaced by a single monolithic device consisting of two adjacent energy-trapping regions on a piezoelectric plate.[27, 28]

Still more complicated lattice filters with multiple resonant circuits in each branch may be synthesized by trapped-energy devices with more than two trapping regions.[29-31]

* See Ref. 5, chap. 5, pp 179–244.

The research upon which this article is based was supported by the United States Atomic Energy Commission.

REFERENCES

1. Mason, W. P., "Electrostrictive effect in barium titanate ceramics," *Phys. Rev.*, vol. 74, pp. 1134–1147, Nov. 1948.

2. Känzig, W., "Ferroelectrics and ferroelectronics," in *Solid State Physics*, vol. 4. New York: Academic Press, 1947, pp. 97–124.

3. Berlincourt, D. A., Curran, D. R., and Jaffe, H., "Piezoelectric and piezomagnetic materials and their function in transducers," in *Physical Acoustics*, vol. 1, pt. A, W. P. Mason, ed. New York: Academic Press, 1964, pp. 169–270; see especially pp. 182–193.

4. Holland, R., "The equivalent circuit of a symmetric N-electrode piezoelectric disk," *IEEE Trans. Sonics and Ultrasonics*, vol. SU-14, pp. 21–33, Jan. 1967.

5. Holland, R., and EerNisse, E. P., *Design of Resonant Piezoelectric Devices*. Cambridge, Mass: M.I.T. Press, 1969, pp. 42–89.

6. Holland, R., "Analysis of multiterminal piezoelectric plates," *J. Acoust. Soc. Am.*, vol. 41, pp. 940–952, Apr. 1967.

7. Newell, W. E., "Tuned integrated circuits—a state-of-the-art survey," *Proc. IEEE*, vol. 52, pp. 1603–1608, Dec. 1964.

8. Newell, W. E., "Ultrasonics in integrated electronics," *Proc. IEEE*, vol. 53, pp. 1305–1309, Oct. 1965.

9. Land, C. E., "Transistor oscillators employing piezoelectric ceramic feed-back networks," *1965 IEEE Internat'l Conv. Record*, vol. 13, pt. 7, pp. 51–68.

10. Schueler, D. G., "Ferroelectric ceramic logic and NDRO memory devices," *1966 WESCON Conv. Record*, vol. 10, pt. 3, paper 3/4.

11. Land, C. E., "Small-signal applications of monolithic multiport piezoelectric devices," *1966 WESCON Conv. Record*, vol. 10, pt. 3, paper 3/5.

12. EerNisse, E. P., "Variational method for electroelastic vibration analysis," *IEEE Trans. Sonics and Ultrasonics*, vol. SU-14, pp. 153–160, Oct. 1967.

13. Holland, R., and EerNisse, E. P., "Variational evaluation of admittances of three-dimensional piezoelectric structures," *IEEE Trans. Sonics and Ultrasonics*, vol. SU-15, pp. 119–132, Apr. 1968.

14. Lloyd, P., "Equations governing the electrical behavior of an arbitrary piezoelectric resonator having N electrodes," *Bell System Tech. J.*, vol. 46, pp. 1881–1900, Oct. 1967.

15. Tiersten, H. F., "Natural boundary and initial conditions from a modification of Hamilton's principle," *J. Math. Phys.*, vol. 9, pp. 1445–1450, Sept. 1968.

16. Allik, M., and Hughes, T. J. R., "Finite element method for piezoelectric vibration," *Internat. J. Numerical Methods in Eng.* (to be published).

17. Holland, R., "Resonant properties of piezoelectric ceramic rectangular parallelepipeds," *J. Acoust. Soc. Am.*, vol. 43, pp. 988–997, May 1968.

18. Berlincourt, D., "Variation of electroelastic constants of polycrystalline lead titanate zirconate with thoroughness of poling," *J. Acoust. Soc. Am.*, vol. 36, pp. 515–520, Mar. 1964.

19. "Standards on piezoelectric crystals, 1949," *Proc. IRE*, vol. 37, pp. 1378–1395, Dec. 1949.

20. Shockley, W., Curran, D. R., and Koneval, D. J., "Energy trapping and related studies of multiple electrode filter crystals," *1963 Proc. 17th Ann. Symp. Frequency Control*, pp. 88–126.

21. Horton, W. H., and Smythe, R. C., "The work of Mortley and the energy-trapping theory for thickness-shear piezoelectric vibrators," *Proc. IEEE (Letters)*, vol. 55, p. 222, Feb. 1967.

22. Mindlin, R. D., and Lee, P. C. Y., "Thickness-shear and flexural vibrations of partially plated, crystals plates," *Internat'l J. Solids and Structures*, vol. 2, pp. 125–139, Jan. 1966.

23. Koneval, D. J., Gerber, W. J., and Curran, D. R., "Improved UHF filter crystals using insulating film techniques," *1966 Proc. 20th Ann. Symp. Frequency Control*, pp. 103–130.

24. Curran, D. R., and Koneval, D. J., "Energy trapping and the design of single- and multi-electrode filter crystals," *1964 Proc. 18th Ann. Symp. Frequency Control*, pp. 93–119.

25. Kaname, Y., Nagata, T., and Nakajima, Y., "Piezoelectric ceramic trapped energy resonators," *Proc. Barium Titanate Study Meeting of Japan*, pp. 31–33, May 19, 1966; English translation, *SC-T-68-1544*, Sandia Corp., Albuquerque, N.Mex.

26. Bleustein, J. L., and Tiersten, H. F., "Forced thickness-shear vibrations of discontinuously plated piezoelectric plates," *J. Acoust. Soc. Am.*, vol. 43, pp. 1311–1318, June 1968.

27. Mailer, H., and Beuerle, D. R., "Incorporation of multiresonator crystals into filters for quantity production," *1966 Proc. 20th Ann. Symp. Frequency Control*, pp. 309–342.

28. Onoe, M., and Jumonji, H., "Analysis of piezoelectric resonators vibrating in trapped-energy modes," *Electronics and Communications in Japan*, vol. 48, pp. 84–93, Sept. 1965.

29. Sykes, R. A., Smith, W. L., and Spencer, W. J., "Monolithic crystal filters," *1967 IEEE Internat'l Conv. Record*, vol. 15, pt. 11, pp. 78–93.

30. Sykes, R. A., and Beaver, W. D., "High frequency monolithic crystal filters with possible application to single frequency and single side band use," *1966 Proc. 20th Ann. Symp. Frequency Control*, pp. 288–308.

31. Onoe, M., Jumonji, H., and Kobori, N., "High frequency crystal filters employing multiple mode resonators vibrating in trapped energy modes," *1966 Proc. 20th Ann. Symp. Frequency Control*, pp. 266–287.

Richard Holland's biography appeared on page 78 of the February issue.

The early history of electronics

V. Commercial beginnings

The electronics industry got its start with attempts to bridge bodies of water radiotelegraphically, culminating in the spanning of the Atlantic by Marconi in 1901. The fifth installment of our continuing series shows how the young industry was bedeviled by priority arguments and litigation

Charles Süsskind *University of California*

The first four articles in this series traced the scientific developments underlying the earliest tentative attempts to exploit the new results of electromagnetic theory for practical purposes.[1] Almost to a man, the main contributors to these developments were connected with one academic establishment or another. Only one was not: an amateur who became the first entrepreneur of the new technology, young Marconi—newly arrived in London from his native Italy in February 1896, not quite 22 years old, and determined to make his fortune in the city that was not only the capital of the most powerful and richest maritime nation in the world but also the headquarters of a far-flung empire.

First "wireless" company

On June 2, 1896, Marconi applied for the first radiotelegraphy patent,[2] which was granted on July 2, 1897. On July 20, 1897, a company was formed, the Wireless Telegraph and Signal Co. Ltd., "to acquire from Signor Guglielmo Marconi certain letters patent... and rights in connection with telegraphy or otherwise, and to carry on the business of engineers and manufacturers of all kinds of telegraphic apparatus." The initial capitalization was £100 000; the first chairman was Henry Jameson Davis (1854–1936), a cousin of Marconi, whose mother had been born Anna Jameson and was related both to that family of Dublin whisky distillers and to the Haigs of Scotland. (It had been Davis who had arranged Marconi's first introduction to Preece, the engineer-in-chief of the British government's telegraph service.)

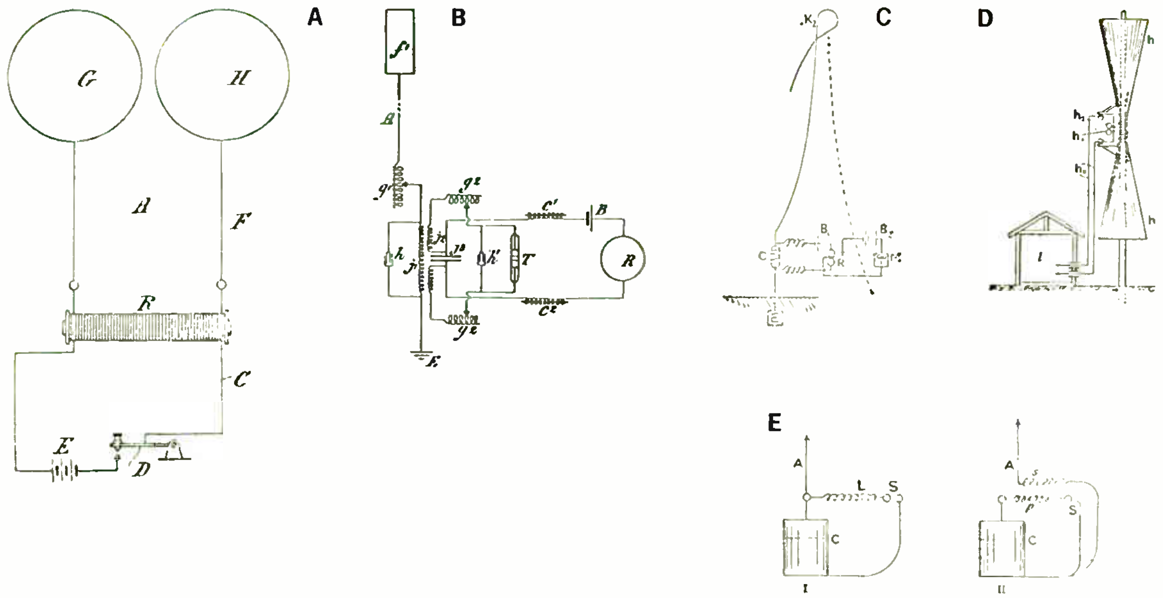
Expectations in the City ran high. The money was quickly subscribed; there was even, as one contemporary account put it, "a great flutter in the dove-cotes of telegraphy, and holders of many millions of telegraph securities, and those interested in the allied industries, began to be alarmed for the safety of their property."[3] Cable shares trembled. Editors and chairmen of one telegraph company after another were forced to explain to stockholders at annual meetings that wireless telegraphy did not, as yet, provide much competition for cable systems.[4]

Wireless Telegraph and Signal Co. was the first commercial enterprise formed to exploit radiotelegraphy; three years later it became Marconi's Wireless Telegraphy Company Ltd. The objectives were to acquire control of Marconi's patents and to exploit them in all countries. However, Marconi had patriotically reserved the right to exploit his system in his native land. In July 1897, Marconi returned to Italy and put on for government observers at Spezzia a demonstration of signaling in which he achieved a distance of 20 km between warships.

The first permanent installation was established at the Isle of Wight, connecting that point with Bournemouth, some 22 km away. Again a four-ball exciter was employed, with one outer ball connected to ground and the other by a wire to an insulated metal net covering a rectangle some 40 meters long that could be raised by a pulley and suspended vertically from a slightly higher mast. Telegraphic messages could be reliably sent at about 12 words per minute. This plant was visited by many notables, including Lord Kelvin, who sent telegrams to Preece and to Stokes and insisted on paying a shilling for each message to emphasize that the system was ready for commercial use. The date of Kelvin's visit (June 3, 1898) is thus the date of the first commercial radiotelegram.

That same spring, Marconi's company arranged a demonstration for Lloyd's of London between Ballycastle in the north of Ireland and Rathlin Island, 12 km away, where a vertical wire was simply suspended from a lighthouse. The experiment was successful, despite the deleterious effect of the lighthouse on the antenna, and served to demonstrate that the equipment did not require skilled personnel for its operation. Soon after that, Marconi's system was employed to report the yacht races at the Kingstown Regatta for a Dublin paper.

Following this success, Marconi received a mark of royal favor: he was asked to install stations that would enable the royal yacht *Osborne* to communicate with the Isle of Wight. With this installation, Marconi was able to study the effect of intervening hills for the first time, since the two stations were not within sight of each other even when the yacht was moored in Cowes bay, about 3 km from the shore station. Queen Victoria was delighted with the new invention and exchanged some 150 messages with Edward Prince of Wales, who was on the yacht. Communications with a shore station were maintained when the yacht moved about; intervening hills appeared to present no obstacle. These successes augured well for the young inventor and his company, which now decided to double its capitalization (at its second meeting on October 7, 1898), at the same time restricting Marconi to



Evolution of radiotelegraphic apparatus before the turn of the century. A—In Marconi's original patent (British 12 039, 1896), spark-gap spheres are shown connected across induction coil and comprise antenna. B—Marconi's more advanced circuit in his famous "four-sevens" patent (British 7777, 1900) shows inductively coupled resonant circuits (*h* and *h'* are adjustable condensers "preferably in the form of two metallic tubes separated by a dielectric and sliding telescopically on each other as in this way their capacity can readily be varied with accuracy to tune the circuits"). C—In receiver used in Marconi's 1901 transatlantic experiments, however, no tuning was used except for accidental capacitance variations as kite rose and fell with wind. D—Meanwhile, Lodge had patented tuned transmitter he called "syntonic radiator" (British 11575, 1897) with inductance connected across tips of two flaring "capacity areas," which strangely prefigure today's VHF antennas. E—Braun's earliest patent (German 111 578, 1898) described his inductive coupling (right) as preferable to directly coupled antennas.

a voting power of 40 percent of the issued capital.

The company next received a request from the headquarters of the British coast guard service, Trinity House, for a demonstration of communications between a lighthouse at Dover and East Goodwin Sands Lightship, 20 km away. Again, the sailors had no trouble learning how to use the instruments. During a gale in January 1899, a heavy sea struck the lightship and destroyed part of her bulwarks; the mishap was promptly reported to headquarters, which dispatched immediate assistance. The first real marine disaster in which radiotelegraphy played a part occurred in the spring of 1899, when the steamer *R. F. Matthews* collided with the lightship in a fog and lifeboats were summoned to the rescue by radio.

Marconi continued to improve his equipment. On March 2, 1899, he read a paper before the IEE, an occasion that was awaited with great interest.[5] More than 300 persons were turned away at the door and the paper had to be read a second time some days later. The basic design was much the same as before, but Marconi described a number of engineering improvements and disclosed that he was also trying directional reflector antennas.

In the same month, Marconi tackled the task of linking England with France by radio. On the English coast, the Dover installation was used for this purpose, and a new station was erected at Wimereux near Boulogne. The link was a success from the start. A message to Branly, acknowledging him as the discoverer of the coherer effect, was sent a day after communications were established. The link caused quite a stir. John Ambrose Fleming (1849–1945), the professor of electrical engineering at University College in London who was to play an important part in the development of electronics himself, wrote an enthusiastic letter to *The Times* describing the equipment and explaining its simplicity and low cost.[6]

The scientific community was scarcely less fascinated. The British Association meeting for 1899 had been scheduled for Dover; it was the centenary of the invention of the Voltaic pile. An antenna was erected atop the town hall tower in Dover and during a public evening lecture by Fleming on "the centenary of the electric current," messages were exchanged with the French Association, then meeting at Boulogne.[7]

America beckoned next and Marconi traveled to the United States to report the races for the cup between British and American contenders, sponsored by the *New*

York Herald, during which thousands of words were transmitted between ship and shore.

The distances achieved by the Marconi system had been continually increasing. During the July–August 1899 Royal Navy maneuvers, several ships had been equipped with the apparatus, and messages had been exchanged over sea paths as long as 135 km. After the station at Wimereux had been established, it was found that its broadcasts were being picked up by East Goodwin Sands Lightship and, on one occasion, as far away as the receiving station established at Chelmsford in Essex, some 135 km away over land and sea. It was becoming obvious that as stations multiplied, something would have to be done to prevent interference. Moreover, everyone was becoming increasingly aware of the fact that radiotelegraphy was not private: anyone with a receiver could listen in. Directional radio beams, for all of their advantages in conserving transmitter power or extending range, were only a partial answer to the problem of privacy.

The means by which this problem was solved, resonant tuning, was conceived by Oliver Lodge as an extension of the syntonic-jar experiments described in our third installment. He recognized that an open circuit radiated too freely to sustain the continuous-wave oscillations necessary for tuning, whereas a closed circuit (although it could be tuned more easily) was, as he put it, "a feeble radiator and a feeble absorber"; he concluded that "the two conditions, conspicuous energy of radiation and persistent vibration electrically produced, are at present incompatible."[8] But by 1897 he had worked out a compromise, a circuit whose essential feature was that an inductance (in the form of a coil of wire) was added to the capacitance between the antenna and the ground, to form a resonant circuit. Tuning could be effected by varying the number of turns in the coil or by adjusting the height and length of the antenna. The rapidly damped wavetrains characteristic of the primitive spark-gap transmitters, in which virtually the entire energy of the spark discharge was consumed in the first half-cycle of the oscillation, were replaced by waves damped so slowly as to approach the continuous-wave condition.

This important improvement was embodied in a series of patents by Lodge and Muirhead; the first specification was filed on May 10, 1897,[9] followed by a number of specifications of further improvements.[10] Long before the first patent was issued, Lodge displayed his apparatus publicly at a meeting of the Physical Society[11] and at a *conversazione* of the Royal Society.[12]

The originality of Lodge's invention may be fully appreciated only if it is realized that at the time he submitted his first specification, Marconi's basic patent had not yet been made public. On the other hand, Marconi continued to make other improvements in his apparatus and covered several by British patents.[13] The most important of them was No. 7777 of 1900 ("Improvements in apparatus for wireless telegraphy")—the so-called "four-sevens" patent, in which Marconi described a different method of tuning. This was a more advanced circuit than that of Lodge; its essence was that the antenna was connected to ground through two coils, of which one was an adjustable inductance (for tuning) and the other was coupled inductively to a third winding, across which was connected the remainder of the transmitting circuit. The oscillation transformer formed by the two coupled coils neatly dealt with the problem of providing a closed oscillating circuit without ruining the radiating properties of the antenna. As we shall see, that circuit may not have been original with Marconi.

The properties of oscillating coupled circuits also had been extensively treated by several scientists during the preceding years, largely as the result of the renewal of interest in oscillatory discharges that followed upon the discoveries of Hertz and the spectacular demonstrations of Tesla. It is quite improbable that Marconi was aware of the extensive analytical background (all in German) represented by the investigations of scientists such as Vilhelm Friman Koren Bjerknes (1862–1951),[14] Anton Overbeck (1846–1900),[15] Josef (Ritter von Armingen) Geitler (1870–1923),[16] and Max Wien (1866–1938).[17]

The "four-sevens" patent proved to be tremendously important to the Marconi Co. until expiration in 1914. Even though they had held the earlier patent, Lodge and Muirhead were never able to force the Marconi Co. to take out a license under their patent. However, its existence had one important consequence. Relations between the Marconi Co. and the British Post Office had rapidly worsened after the company was founded in 1897. Preece felt that Marconi should have dealt only with the government and had actually recommended that a reasonable purchase price should be set. Marconi was perfectly willing to negotiate, but when no concrete offer was forthcoming, he lost patience. No one could blame Marconi for being "tired of kicking his heels on the inhospitable doorstep of the British Treasury," as one contemporary commentator put it.[18] After that, not even the personal friendship and mutual respect between Marconi and Preece could avert the break. The company made another effort in 1899, when it offered the government outright use of all Marconi patents for government purposes throughout the British Empire for a license fee of £50 000 per year. This offer also did not prove to be acceptable to the Post Office—one of the reasons being the existence of Lodge's patent.

But Lodge did not get any real satisfaction until 1911, when he managed to have his patent prolonged for another seven years, a dispensation obtainable under British patent law if the inventor is held to have had insufficient opportunity to benefit from the patent during its normal life. In the resulting welter of mutual infringement suits, Preece now openly supported Lodge. The suit was settled out of court. Under the terms of the agreement, the Marconi Co. acquired the patents owned by the Lodge–Muirhead syndicate, which was thereupon dissolved.

Radiotelegraphy in Germany

Marconi's only important competition during the early years came from Germany. There, the electronics industry had its beginnings in the efforts of two separate groups of workers. The first was headed by Adolf Slaby, whose researches at Charlottenburg had received a fresh impetus from his presence at one of Marconi's early demonstrations across the Bristol Channel in the summer of 1897. Marconi was not at all happy about his guest. "I was practically compelled to allow my experiments... to be witnessed in all their details by Prof. A. Slaby... Previous to these tests I did not know Prof. Slaby, nor had I expressed the slightest desire to allow him to witness my experiments, but did so out of courtesy to the British Post Office officials, who introduced him and who made

Karl Ferdinand Braun (1850–1918) has not received the recognition due him outside his native Germany, and even there he is known mainly for inventing the cathode-ray oscilloscope with which his name is associated. Until recently, there was no book-length biography of this co-winner (with Marconi) of the 1909 Nobel Prize in physics, awarded to him mainly on the basis of improvements in radio transmitters and receivers that underlay such lasting contributions as inductive antenna coupling and resonant-circuit tuning and that gave Germany an early start in telecommunications technology second only to Britain's.

Braun studied at Marburg and Berlin and, after holding teaching positions in several German cities, in 1885 became professor of experimental physics in Tübingen, where he worked on electrolysis and electrical measurements. In 1895 he moved to Strasbourg (then a German university), where he developed the CRT oscillograph and became interested in radiotelegraphy. Beside his "coupled transmitter," which saw relatively quick industrial application, he worked on directional antennas and applied the results of some of his earlier research on crystals to the use of such materials for detection (replacing the coherer); and he contributed to the use of radio transmission as beacons for navigation.

He visited the United States in 1909, and later (after World War I broke out in Europe) to testify in a patent suit. The suit was postponed, and when the United States entered the war, Braun had to remain, living in his son's home in Brooklyn, N.Y., where he died on April 20, 1918.

A detailed biography by Friedrich Kurylo, *Ferdinand Braun: Leben und Wirken des Erfinders der Braunschen Roehre, Nobelpreistraeger 1909* (Munich: Heinz Moos Verlag, 1965), was commissioned by the three firms that were the principal beneficiaries of Braun's work: Telefunken, Siemens & Halske, and Hartmann & Braun. Its translation into English awaits a similar philanthropic impulse.

the request . . .," Marconi wrote later.[19] His misgivings proved to be well justified. Slaby resumed his experiments with the help of his assistant, Graf von Arco, and soon acquired the support of a major electrical manufacturer, the Allgemeine Elektricitäts-Gesellschaft (AEG), the successor of the German Edison Company. Kaiser Wilhelm II took a personal interest in the new invention and presently arranged to place men and materials of the armed forces at Slaby's disposal. Slaby and von Arco made a number of improvements, the most important of which related to the understanding of the current distribution along the antenna wire and the resulting optimization of the location of the point at which connection to the rest of the circuit was made.[20] In 1900, the two men were able to put on an elaborate demonstration of a tuned, multiplexed system for the Kaiser's benefit.[21]

The other, and doubtless more important, group was that surrounding Karl Ferdinand Braun (1850–1918) at the University of Strasbourg. Braun is best remembered as the inventor of the cathode-ray oscilloscope (*Braunsche Röhre* in German), but it was the development of a transmitter and receiver rivaling Marconi's that earned him the Nobel Prize for physics in 1909 (which he shared with Marconi). Braun had become interested in radiotelegraphy as a result of being called in to give an expert opinion on a method of induction telegraphy through water; during the trials he noticed that ac (which had just been introduced in Strasbourg) worked better than interrupted dc, presumably because of skin effect, and he decided to find out if higher frequencies would not work even better. They did not: trials were made in rivers and other winding waterways, and the high-frequency waves would not travel around bends. But, in trying to improve the source, Braun hit upon an effective resonant circuit and coupling method that, when transferred to a radio transmitter, proved to be superior to Marconi's.

In Marconi's original circuit, the spark gap was inserted between antenna and ground. Braun connected an *LC* circuit across the spark-gap generator and coupled inductively to a secondary winding inserted between antenna and ground. Not only was the antenna no longer at high voltage (which was dangerous and subject to disabling short circuits), but the circuit could be tuned and the coupling and the transformer ratio could be adjusted for optimum effect. There is some evidence that Marconi's first transmitter improvements, which also consisted of inductive coupling, were freely borrowed from Braun's patents,[22] although Braun's other improvement, transmitter tuning (to which he presently added a tuned receiver), was thought to be a flaw at first—Marconi's system was considered simpler, since it did not "require" tuning.

Even though Braun's patents antedated Marconi's "four-sevens" patent, the German backers of the company formed to exploit the Braun system did not contest the infringement; they were too busy worrying about their archrival, Slaby. For a time the Braun group, which before long came to include the German industrial giant Siemens & Halske, even considered joining forces with Marconi against Slaby and AEG.

The existence of two rival groups in Germany was highly undesirable from the viewpoint of the German authorities. Not only was their position vis-à-vis the increasingly powerful Marconi Co. weakened, but they were also forced to make technical decisions that a single manufacturer would have had to make as part of his own responsibility. The situation was exacerbated by the fact that the German army favored Braun, whereas the navy preferred Slaby. Considerable pressure was brought to bear, from the Kaiser down to the government departments concerned, to make the two groups cooperate. Outright collaboration between two rival manufacturers was not feasible and the matter hung fire until 1903, when a suitable compromise was worked out. In that year, the two companies formed a new company to which they transferred all their work in radiotelegraphy, including their patents and their lists of customers; von Arco became the technical director. The name of the new firm was Gesellschaft für drahtlose Telegraphie m.b.H., but it quickly became known by its telegraphic address, *Telefunken*, which also became its trademark.[23]

The creation of a single concern with strong government support considerably strengthened the German position. The relative merits of the Braun–Siemens and Marconi systems and questions of priority were much debated.[24] The Marconi Co. on its part made no attempt to sue Telefunken for patent infringement either—whether because of diffidence over the validity of Braun's claims or because of the official support he was receiving, it is difficult to say. It should be noted, however, that after the Marconi Co. had acquired the Lodge–Muirhead patents in 1911, Telefunken came to an "arrangement" with Marconi Co., with both sides tacitly admitting patent infringements.[25] Until then, relations between the British and German interests were characterized by frequent recriminations and mutual accusations. For instance, the following anecdote is told by Marconi's associate, R. N. Vyvyan: "On the success of Marconi's experiments across the Atlantic he was invited to Rome, as the guest of the Italian Government. The German Emperor had arrived the previous day, also staying at the Quirinal as a guest of the King. At dinner the night of Marconi's arrival, the Emperor, after discussing wireless turned to Marconi and said 'Signor Marconi, you must not think I have any animosity against yourself but the policy of your Company I object to.' Marconi replied, 'Your Imperial Majesty, I should be overwhelmed if I thought you had any personal animosity against myself but the policy of my Company is dictated by myself.'"[26]

It is a curious fact that neither in Europe nor later in America did the established manufacturing companies play a direct part in the commercial development of radiotelegraphy. As W. R. Maclaurin has pointed out, neither the telegraph and cable industries nor the telephone and electrical manufacturing industries had the vision to participate directly in the initial phases of the new development; the advance came largely from new concerns and new capital.[27]

Transatlantic radio

In June 1900 Marconi returned from a trip to the United States determined that he would attempt to bridge the Atlantic by radiotelegraphy. Such an undertaking was largely an act of faith, since there seemed to be no scientific reason that radio waves, which had been shown to be exactly like light waves in every respect except wavelength, should follow the curvature of the earth. Though scientists had speculated about the existence of an ionized layer in the atmosphere, no one had tied it to communications; the distances achieved were thought to depend on the large height of antennas and upon refraction. Marconi's determination to go ahead with the experiment despite the small probability of success has been ascribed variously to ignorance of scientific principles and unprecedented courage. The truth is probably that, like many another practical man unencumbered by much formal technical training, Marconi was not averse to proving the scientific wiseacres wrong—a kind of motivation that has been known to lead to useful inventions before and since, although its many failures are mostly forgotten.

The preparations for the test took over a year. The first task was to replace the laboratory-size equipment typical of earlier tests by a well-engineered power plant capable of providing around 10 kW, together with the switches and other circuit components that would have to be designed for use with this unusually high power level. This portion of the project was entrusted to J. A. Fleming, who had become a consultant to the Marconi Co. A site was selected at Poldhu Bay near Mullion in Cornwall, a few miles from the westernmost tip of England. This location, on a windswept Cornish cliff far from any town, had the additional advantages that it would not interfere with other electrical installations and would keep out the curious: Marconi had no intention of risking his newly won reputation by publicizing his fantastic experiment before its successful completion.

Construction began in the autumn of 1900, but it was not until the following spring that telegraphic tests could be carried out between Poldhu and stations on the Isle of Wight and in the southern part of Ireland. The transmitting antenna consisted of an array of wires forming an inverted cone with its apex near the ground, suspended from a circular array of twenty wooden masts each about 60 meters high and arranged in a circle 60 meters in diameter. No one knew with certainty what the wavelength of the generated oscillations would be. Nearly a year was occupied in erecting the array and a similar, receiving one in the United States, at South Wellfleet, on Cape Cod, Mass. Then disaster struck. In September 1901 a tremendous storm swept the English coast and toppled over most of the masts. The U.S. station suffered a similar fate soon thereafter. A simpler transmitter array was erected next in which only two masts were employed. Some 50 wires were suspended from a stay between the mast tops; the wires were about a meter apart at the top and converged fanlike to a point below.

At the end of November, Marconi sailed with a couple of assistants to Newfoundland, where he improvised a receiving station on a plateau called Signal Hill, near St. John's. There was no question of erecting masts here: balloons or kites would have to be used for suspending the receiving antenna. The wind proved to be too strong for balloons, the first of which soon ripped loose and was lost. The first kite was likewise lost, but the second proved to be more successful and rose to a height of some 130 meters, carrying the antenna wire with it.

To stretch the capabilities of the receiver to its fullest, Marconi abandoned the Branly-coherer circuit with a paper-tape recorder or "inker" and instead used a novel sort of coherer that later became known as the Italian Navy coherer, since an early form of it had been evolved through operational use in that service. It was similar to the previously mentioned coherer first demonstrated by Tommasina and had been brought to Marconi's attention by his friend, and later collaborator, Marchese Luigi Solari (1875–1957). A similar coherer had also been independently suggested by a pair of British submarine-cable engineers, A. C. Brown and G. R. Neilson. The device was not only very sensitive but, to a large measure, self-restoring. It consisted of two plugs of iron or carbon, separated by a gap of a few millimeters into which a drop of mercury was introduced. A small battery was placed in series with this coherer; the relay normally used to operate the telegraphic inker was replaced by an ordinary telephone receiver.

Marconi now cabled instructions to the crew in England to begin the prearranged schedule of transmissions, which consisted of the single code letter *s* (···) transmitted at regular intervals from 3 to 6 P.M. each day. On the following day (December 12, 1901) he heard the faint signals for the first time, but not throughout the entire period of transmission. The irregularity of reception was caused by the fact that the height of the kite supporting the receiver antenna varied up and down with the wind, thus varying the capacitance (and hence the resonant frequency) of the receiver circuit. The circumstance of the kite's bobbing up and down was very likely a piece of good luck, since it may have brought the receiver into resonance with the transmitter during at least part of the period of transmission. But if Marconi was aware of his good fortune, he said nothing about it. Nor did he realize at that time that the probability of success would have been increased considerably by scheduling the transmission after dark.

The fact that the telegraphic recorder had to be replaced by a telephone receiver for increased sensitivity led to some skepticism on the part of a few observers, who maintained that Marconi and his assistants must have mistaken the random clicks produced by atmospheric electricity for the prearranged signals. Some of the skeptics remained unconvinced for a couple of months, until adequately witnessed experiments carried out by Marconi aboard the homeward-bound U.S. liner *Philadelphia* showed beyond doubt that signals from the Poldhu transmitter were being received over distances comparable to the nearly 3500 km to Newfoundland. (It was on this voyage that Marconi first noticed the substantial difference between the distances attainable at night and during the day.) But one group of observers—the Anglo-American Telegraph Co., whose submarine cable terminated in Newfoundland—had no doubts whatever. The company had secured a monopoly on receiving transatlantic messages there that still had over three years to run, and promptly requested that Marconi terminate his experiments. At the hour of his greatest triumph, the young inventor was unceremoniously shooed out of Newfoundland.

The achievement of transatlantic radio communications was the crowning success of the early years of radiotelegraphy. Marconi was acclaimed throughout the world. He had become a legendary figure at the age of 27. "The initial success of Marconi appeals powerfully to the imagination," editorialized *The New York Times*. "It will be the fervent hope of all intelligent men that wireless telegraphy will very soon prove to be not only a 'scientific toy' but a system for daily and common use. The men of science point out the obstacles. They have commonly been deemed insuperable. The first triumph is an augury of future conquests."[31]

The eye of Marconi himself remained firmly fixed to commercial exploitation as he pointed out the advantages of a pair of stations compared with the cost of laying and maintaining thousands of milometers of submarine cable: "I believe the cost of what we now call cabling to England might be reduced at least twenty fold. The present rate is twenty-five cents a word. I do not see why eventually, with the wireless system this cost should not be reduced to one cent a word or less."[31] But many years were to elaspe before transatlantic radiotelegraphy would become competitive with the cables, and it never replaced them altogether. Rather, the first important commercial success of radiotelegraphy came in a related field: marine communications.

This study was sponsored in part by Grants GS-1428 and GS-2411 from the National Science Foundation.

REFERENCES

1. Süsskind, C., *IEEE Spectrum*, vol. 5, pp. 90–98, Aug. 1968; vol. 5, pp. 57–60, Dec. 1968; vol. 6, pp. 69–74, Apr. 1969; vol. 6, pp. 66–70, Aug. 1969.

2. Marconi, G., British Patent 12 039, June 2, 1896.

3. Fahie, J. J., *A History of Wireless Telegraphy*. Edinburgh and London: Blackwood, 1899.

4. See, for instance, *Electrician*, vol. 41, pp. 82–83, 1898.

5. Marconi, G., *J. IEE (London)*, vol. 28, pp. 273–297, 300–316, 1899.

6. *The Times* (London), April 3, 1899.

7. *Electrician*, vol. 43, pp. 760–768, 1899.

8. Lodge, O., *Electrician*, vol. 33, pp. 153–155, 186–190, 204–205ff, 1894.

9. Lodge, O. J., and Muirhead, A., British Patent 11 575, May 10, 1897. (Also U.S. Patent 609 154.)

10. Lodge, O. J., British Patents 16 405, July 10, 1897; 18 644, Aug. 11, 1897; 29 069, Dec. 8, 1897, 11348, June 3. 1901; 10 181, May 2, 1902; 13 521, June 14, 1902.

11. *Proc. Phys. Soc.*, vol. 16, pp. 58–61, 1898.

12. *Electrician*, vol. 41, pp. 71–72, 1898.

13. Marconi, G., British Patents 12 326, June 1, 1898; 6982, Apr. 1, 1899; 25 186, Dec. 19, 1899; 7777, Apr. 26, 1900.

14. Bjerknes, V., *Ann. Phys.*, ser. 3, vol. 44, pp. 74–101, 513–526, 1891; vol. 54, pp. 58–63, 1895; and vol. 55, pp. 121–169, 1895.

15. Overbeck, A., *Ann. Phys.*, ser. 3, vol. 55, pp. 623–632, 1895.

16. Geitler, J., *Ann. Phys.*, ser. 3, vol. 55, 513–524, 1895.

17. Wien, M., *Ann. Phys.*, ser. 3, vol. 61, pp. 151–189, 1897.

18. *Eastern Daily Press* (Norwich, England), June 16, 1897; cited by S. G. Sturmey in *The Economic Development of Radio*. London: Duckworth, 1958, p. 19.

19. Marconi, G., *Elec. Rev.*, vol. 84, p. 179, 1919.

20. German Patents 113 285, 116 071, 116 113, 124 154, 126 273, 127 730, 129 017, 129 892, 130 120, 130 723, 131 584–6, 133 718, all applied for during the years 1899–1901.

21. *ETZ*, vol. 22, pp. 38–42, 1901; *Electrician*, vol. 46, p. 475, 1901.

22. Braun, K. F., German Patents 111 578, Oct. 14, 1898; 109 378, Jan. 26, 1899; British Patents 1862 and 5104, 1899.

23. Siemens, G. *History of the House of Siemens*. Freiburg and Munich: Karl Alber, 1957 vol. 1 chap 11.

24. See, for instance, *Electrician*, vol. 46, pp. 778–779, 1901; and vol. 52, pp. 1033–1034, 1904.

25. Sturmey, S. G., *op. cit.* (Ref. 18), p. 20.

26. Vyvyan, R. N., *Wireless over Thirty Years*. London: Routledge, 1933, p. 51.

27. Maclaurin, W. R., *Invention and Innovation in the Radio Industry*. New York: Macmillan, 1949, chap. 2.

28. *Electrician*, vol. 51, pp. 360 and 502, 1903.

29. Brown, A. C., and Neilson, G. R., British Patent 28 955, Dec. 17, 1896.

30. *New York Times*, Dec. 17, 1901.

31. Dunlap, O. E., *Marconi: The Man and His Wireless*. New York: Macmillan, 1907, p. 108.

A biography of Charles Süsskind appeared on page 76 of the March 1970 issue.

New product applications

New concept in logic-card systems allows the user to select cards and do simple logic himself

A new and unique concept in logic-card systems permits a systems engineer to select the complex function cards he needs and then do the simple logic himself. Or, alternatively, the manufacturer will completely wire-wrap a system from a block diagram or schematic using computer-aided design and automatic wire-wrapping techniques.

The new concept is made possible by what is called the "Wrap-X" Printed Circuit Board. This module contains 16 wire-wrap sockets for either 14-pin ICs and/or 16-pin MSI chips. What's different about the system is that it is mechanically compatible with over 100 DTL, HTL, and TTL standard, off-the-shelf "function" logic cards such as decimal converters and shift registers.

The new system competes with the large-scale back-plane approach but is said to have significant economic advantage in small systems and prototype quantities of large systems. One reason is that complex function cards cannot be economically duplicated with back-plane.

The back-plane approach requires very complex wire-wrap interconnections to replace the function cards. Therefore, an extensive run list and many hours of engineering design time are necessary. In the Wrap-X system, however, most of the interconnect is plated. Also, in the back-plane, discrete components and semiconductors must be housed on plug-in platforms each of which requires a physical design before the run list can be generated. This approach also necessitates many hours of design and drafting time over and above the Wrap-X and standard-card approach.

Figure 1 shows the Wrap-X approach as used for a simple analog-to-digital comparator system. It is representative of most small control or measurement applications.

An illustrated brochure is available from Datascan, Inc., 1111 Paulison Ave., Clifton, N.J. 07013.

Circle No. 85 on Reader Service Card

FIGURE 1. Wrap-X system for a simple analog-to-digital comparator system uses two function cards and one Wrap-X card.