VOLUME 6 ☐ NUMBER 1 JANUARY 1969

IEEE spectrum

features

29 Spectral lines: Technological forecasting

Although technological forecasting is at present much less than a science, it is much more than a fantasy. It is carried on by many bodies, from governments to industries; for example, one company has plans, based on forecasts, for the era A.D. 2000

30 Technological forecasting in planning for company growth

Harper Q. North, Donald L. Pyke

Of all the scientific knowledge that will be available in the year 2000, only 20 percent has been discovered thus far. This suggests that today's technological explosion is only a burning fuse, with the real explosion yet to come

37 Public safety, the radio spectrum, and the President's Task Force on Communications Policy

Robert P. Eckert, Peter M. Kelly

Representing an investigative study and evaluation of the problems of mobile-radio communications congestion, this article, prepared for an important Presidential task force, also contains some important considerations for civil disturbances

47 Synthesizing active filters

Sanjit K. Mitra

An examination of past and present contributions reveals a welcome and growing trend toward the practical implementation of active filters

64 Stimulating change in engineering education

J. O. Kopplin

Part of the projected changes in education will be brought about by advances in communication technology. It appears likely that in the future teachers and students will meet in person only in the laboratory, around the conference table, and over coffee

67 The compatibility of materials

Walter H. Kohl

Materials for critical projects, such as the space program, must be chosen with care so that they will be able to perform their intended functions economically and safely in rugged environments

79 Ceramic capacitors for hybrid integrated circuits

Donald W. Hamer

In spite of its higher cost, the multilayer chip is the most popular type of ceramic capacitor because of the extremely wide capacitance range available (now seven full decades) and because of its extremely attractive mounting geometry

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

85 Telecommunications and electronics in Japan

Koji Kobayashi

After World War II, Japanese technology had to be developed anew from the damage wrought by the conflict. The most significant step in this direction was the importation of modern management and control techniques, mainly from the United States

89 Electrical features of the Grand Coulee third power plant

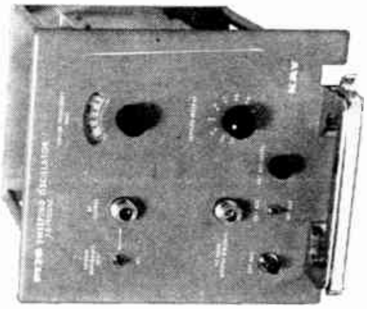
John V. Baptist, Roy Y. Nitta

A third power plant is expected to increase the total generating capacity of the Grand Coulee system to five times the present output, with individual generator ratings increased almost six times

the cover

Signal-flow graphs, one of which is freely represented on this month's cover, are widely employed in the field of network theory. For a discussion of active-filter network developments turn to the article beginning on page 47

departments



MICROWAVE SWEEP PLUG-INS TO 4GHz

PLUG-INS	FREQUENCY RANGE	POWER OUTPUT*		PRICE
P-120	.75 — 1.5 GHz	Leveled, 10mw	Unleveled, 13mw	$1100.
P-125	1 — 2 GHz	" 10mw	" 13mw	1150.
P-126	1.2 — 2.4 GHz	" 50mw	" 70mw	1275.
P-128	1.5 — 3 GHz	" 15mw	" 18mw	1495.
P-127	2 — 4 GHz	" 10mw	" 13mw	1490.

KAY ELECTRIC COMPANY
Maple Avenue • Pine Brook, New Jersey 07058

* Minimum into 50 ohms

Circle No. 3 on Reader Card

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

BOARD OF DIRECTORS, 1968*

S. W. Herwald, *President*
L. C. Holmes, *Vice President*
J. H. Mulligan, Jr., *Vice President, Technical Activities*
F. K. Willenbrock, *Vice President, Publication Activities*
W. G. Shepherd, *Secretary*
Harold Chestnut, *Treasurer*
W. K. MacAdam, *Junior Past President*

F. E. Borgnis
J. G. Brainerd
E. E. David, Jr.
R. G. Elliott
G. A. Fowler
R. W. Gillette
E. L. Harder
F. A. Hawley

Hubert Heffner
D. M. Hodgin
W. H. Huggins
S. F. Kaisel
H. R. Mimno
C. F. Savage

Shigeo Shima
J. N. Shive
W. T. Sumerlin
R. H. Tanner
J. G. Truxal
R. C. G. Williams

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

* *Term expires January 7, 1969*

HEADQUARTERS STAFF

Donald G. Fink, *General Manager*

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

Committees and Staff Secretaries

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

IEEE SPECTRUM is published monthly by The Institute of Electrical and Electronics Engineers, Inc. Headquarters address: 345 East 47 Street, New York, N.Y. 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 $1.00. Prices for nonmember subscribers available on request. Editorial correspondence should be addressed to IEEE SPECTRUM at IEEE Headquarters. Advertising correspondence should be addressed to IEEE Advertising Department, at IEEE Headquarters. Telephone: 212—752-6800.

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

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

BPA

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

VOLUME 6 □ NUMBER 1 IEEE spectrum JANUARY 1969

Spectral lines

Technological forecasting. The Organization for Economic Cooperation and Development (OECD) has 21 members, including Japan, Canada, the United States, and various European powers. These governments, recognizing technology as an influence in the transformation of society, hope to foresee its effects and to take them into account when formulating national goals and plans.

The Organization commissioned Dr. Erich Jantsch to report on technological forecasting as it exists at present; the result[1] is a murky but intriguing document of some 400 pages, including a bibliography with more than 400 entries. The volume shows that although technological forecasting is at present much less than a science, it is much more than a fantasy.

The art has two aspects. One is "exploratory" forecasting, which starts from knowledge available now and makes a probabilistic assessment of where that will lead to. "Normative" technological forecasting, on the other hand, decides on what the future ought to be like, and works backward to determine what should be done now in order to ensure that the future shall be like that. Exploratory forecasting starts here and looks into the future; normative forecasting decides on goals and devises routes for reaching them. Ultimately, the two procedures will presumably be coupled in a feedback loop.

The exploratory forecasting bases itself on opportunities—the normative, on needs. Normative forecasting for technology is beset by two distinctive difficulties. One is the establishment of a timetable for invention; the other is the problem of whether what we now foresee as a need will be actually appropriate at some time in the future. It is, for example, arguable that some of Great Britain's difficulties in 1967–68—devaluation, brain drain, adverse balance of trade, to name a few—have been generated by the success of programs of reform that were laid out by the Fabians about 70 years ago, when the United States and Russia were of negligible consequence as industrial powers.

A technological example of normative forecasting was the late President Kennedy's establishment of a program for putting a man on the moon in this decade. It is perhaps not coincidental that the project was announced at a time of some slackness in the economy, and of some prospect of underemployment of U.S. technological manpower. One result of the program has been to commit an appreciable fraction of the country's technical resources to the space program at the present time, when other problems—including mere terrestrial transportation—are clamoring for attention. Another result of the commitment to a lunar excursion has been the continued concentration on that goal at the expense of others.

Technological forecasting is at present an activity carried on by many governments. It is also carried on by lesser bodies, including capitalist industries. The American Telephone and Telegraph Company has plans, based on forecasts, for the era A.D. 2000. As president of Radio Corporation of America, the elder Sarnoff conspicuously set goals (e.g., color television with no moving parts) toward which the company's technology should strive. In this issue, H. Q. North and D. L. Pyke describe the beginning of an elaborate and systematic attempt by TRW, Inc., to guide its research and development efforts into channels where they will have maximum benefit to the company's competitive position (see pages 30–36).

However, it is normative forecasting by governments that has the greatest potential for impact on society. Capitalism of the sort now existing in the United States differs from socialism most fundamentally in the fact that a capitalist society has very many planners, whereas a socialist one has fewer. If normative forecasting at the national level takes firm root, presumably it can see that things develop in the desired way only by restricting the freedom of decision now granted to private planners. It seems likely, therefore, that normative forecasting by governments will diminish the difference between the economic patterns of the advanced nations; one result could be the lessening of tensions among these nations.

A less attractive implication of normative forecasting by governments is its relation to the liberty of the citizen. For example, any comprehensive vision that is to shape a nation's course for perhaps 20 years must include an assumption about the population that will exist at the end of the period. What happens if after a few years it becomes apparent that the population trend is out of step with the plan? Do you change the plan on which so much thought and economic effort have been expended, or do you take steps to alter the population trend? In principle, the government can have, say, a five-year plan that is fairly firm, but can allow some flexibility in its program for later years, which can be reassessed and modified as unscheduled trends develop. As the example of the space program shows, however, a project on the grand scale has a momentum of its own.

With suitable modifications that will be learned as the art develops, it seems likely that normative technological forecasting will work. If it works, the temper of our times guarantees that governments will use it. There thus promises to open up a whole new area of action for the professional engineering societies. By starting now on some normative forecasting about their own development, they can increase the influence and usefulness that they will have when the governments call for their opinions.

J. J. G. McCue

---

[1] Jantsch, E., *Technological Forecasting in Perspective*, Organisation for Economic Co-operation and Development, Paris, 1967.

Technological forecasting in planning for company growth

The ability to predict consumer demands and technological advances and successfully apply results to corporate objectives may determine whether a company survives in an exploding technology

Harper Q. North, Donald L. Pyke *TRW Inc.*

No one company has resources that are sufficient to permit it to counter all potential threats or to explore all available opportunities resulting from today's technology explosion. Accordingly, each company must become increasingly selective in its allocation of those resources committed toward insuring its future. This article describes the particular experience of TRW Inc. with technological forecasting—one of several techniques that a company may use in improving its selectivity.

The technological explosion, which affects some industries more than others, has been particularly violent in the electronics field, leading most (and it would probably be safe to say all) companies in this area to seek means of improving planning, not only for growth, but also for survival.

The Electronic Industries Association (EIA) reports[1] that sales in electronics have grown from $1 million in 1914 to over $20 billion in 1966. One wonders what the figure might have been had not the depression of the thirties put a ten-year dip in the growth curve. A normalized curve, representing the years from 1940 on, reveals that sales have literally *doubled* every five years—a fantastic rate of growth!

This explosion is partially explained by the growing number of technical people taking part in the growth. Derek J. De Solla Price[2] observes that 80 to 90 percent of all scientists who ever lived are alive today. This figure might approach 100 percent if applied to scientists and engineers in the field of electronics, since Lee De Forest's invention of the vacuum triode occurred as recently as 1906, and the remarkably short period of only 20 years has passed since Bardeen and Brattain published their now-famous letter[3] heralding the principles of the point-contact transistor.

A large fraction of the scientists and engineers from this growing pool find their way to industry, but the number available can scarcely keep pace with the demand. The limitations imposed by the short supply of outstanding people, coupled with the ever-present "profit squeeze," make it imperative that companies improve their R & D planning in order to stay competitive, and become expert at the job if they are to grow and remain profitable.

Complex tasks, such as the planning and management of research and development, are accomplished as a result of the operation of some program or system, whether it has been defined or not. The chart shown in Fig. 1 describes one such system.[4] Although very few R & D organizations follow precisely this idealized system, a review of its logic is worthwhile.

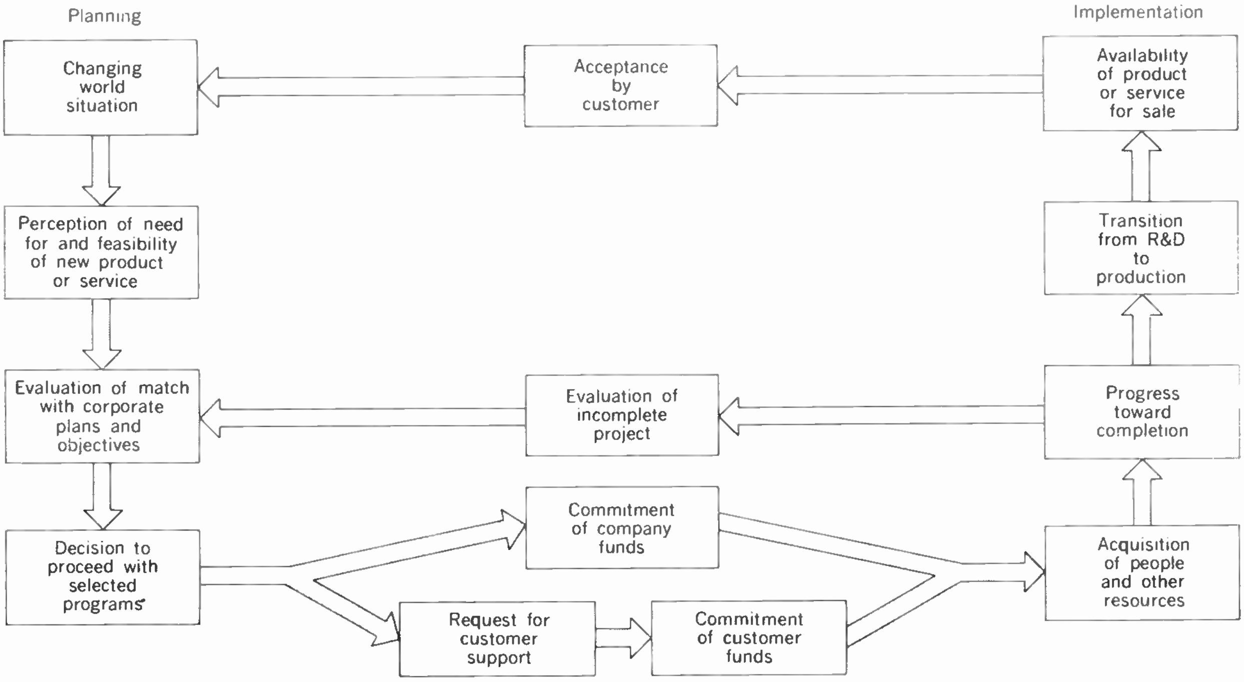


FIGURE 1. A research and development project life cycle.

Most developments are triggered by the identification of an unsatisfied need for a new product or service. If the proposed development provides a good match with corporate objectives and available resources, and if competitors are not too well entrenched, a decision to proceed is likely.

Support for the proposal can come from a commitment of either company or customer funds, depending upon the nature of the business. With all necessary resources assembled, the project is launched and, as work progresses, is subject to evaluation, which may lead to a cancellation of the decision to proceed. However, where progress is satisfactory and there is a transition from R & D to production, the end result is a salable product or service. Customer acceptance and usage contribute to a change in the world situation; thus the cycle is brought to completion.

The system is quite effective when the cycle keeps up with the rate of change in the relevant technology; it even provides some "slack" for the researcher in an industry whose technology is relatively stable. But the electronics industry is based upon some of the fastest moving technologies in today's world. Products are frequently obsolete before the design is even completed.

A common response to this situation is to increase substantially the investment in research and development on the theory that a fraction of the completed programs will lead to products with a short but profitable lifetime before obsolescence. This approach is wasteful from the standpoint of both manpower and money. Accordingly, the more practical response by industry is an attempt to improve its selection of R & D programs.

Technological forecasting is only one of a number of tools available to aid in the selection process; but as the rate of expansion in technology increases, so does the importance of some form of technological forecasting. In 1965, TRW Inc. began an experiment in this area that previewed a 20-year period.[5] This study (Probe I) used a modification of the Delphi technique, developed by Drs. N. Dalkey and O. Helmer of the Rand Corporation,[6] and later employed by Dr. Helmer and T. J. Gordon in a long-range forecasting study[7]—a condensation of which appeared in *News Front Magazine*.[8]

The technique seeks to take full advantage of the committee approach, while at the same time eliminating some of the disadvantages. Those who have dealt with committees of experts may have run into problems involving one or more of the following types of people:

1. The expert with a reputation who feels called upon to defend his publicly stated opinion.

2. The senior executive with whom subordinates are reluctant to differ.

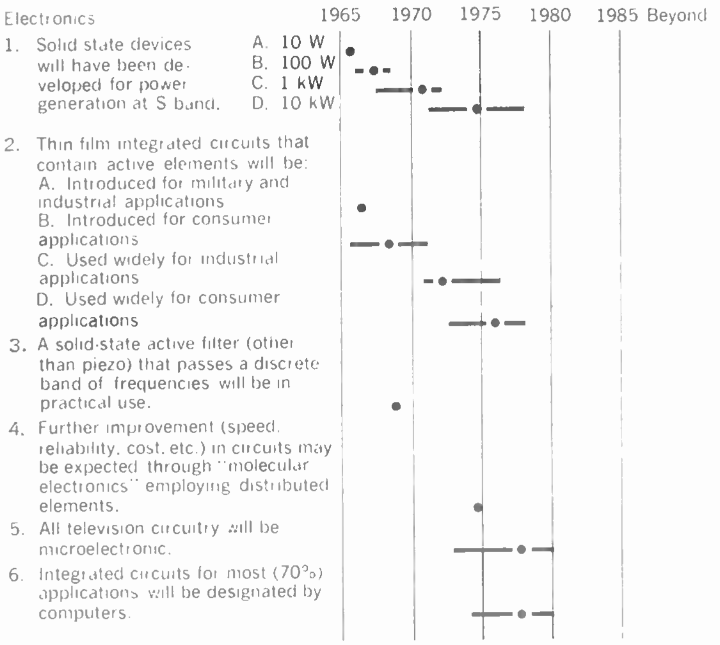


FIGURE 2. Sample of Probe I results for electronics.

3. The silver-tongued salesman who can "sell refrigerators to Eskimos."

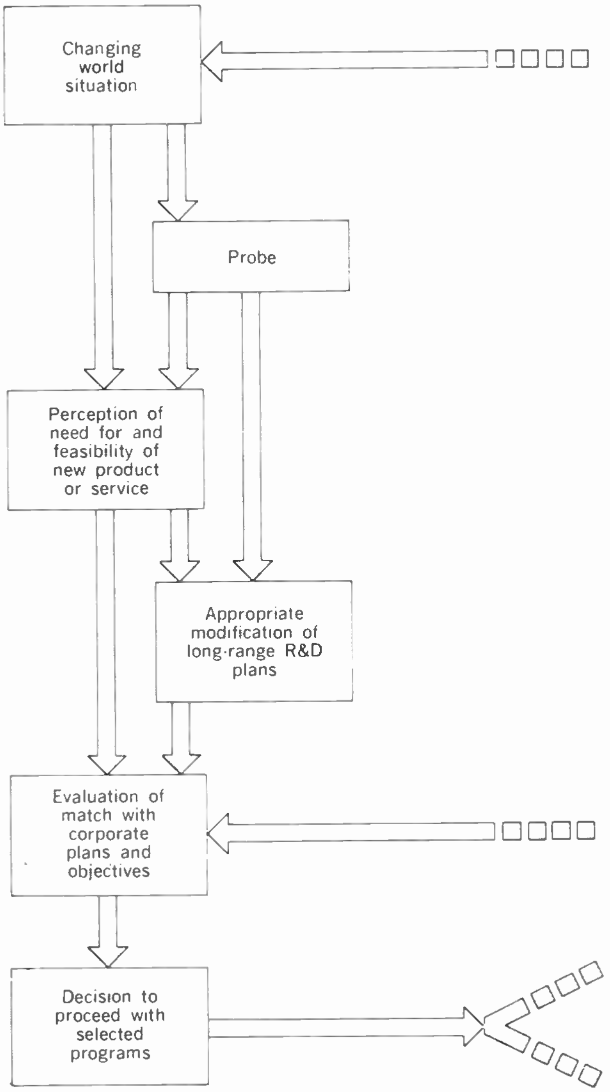
Undoubtedly, others could be added from the reader's own experience. The Delphi technique attempts to avoid these problems by dealing individually with every member of the committee, thus protecting the anonymity of each person.

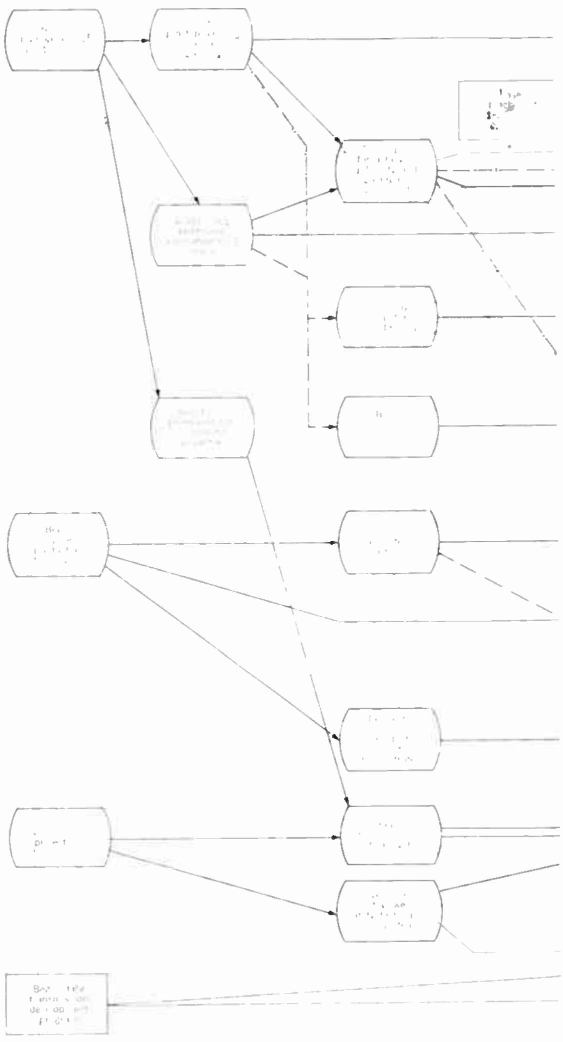
TRW's experimental study was completed in June 1966, and resulted in the publication, for internal use only, of a 50-page document[5] containing a forecast of 401 technical events that a panel of 27 experts felt would occur during the next 20 years and have a significant impact on TRW—its products, services, or processes.

A sample of results that may be of interest to electronics people is contained in Fig. 2. The median date is indicated by a dot, and the range of estimated dates is covered by the solid line. Where there is no spread, artificial concurrence can probably be assumed. These estimates were made over two years ago; they will be "updated" in Probe II, which will be described later.

Figure 3 shows the manner proposed to introduce

FIGURE 3. Planning subsystem of an R&D project life cycle.



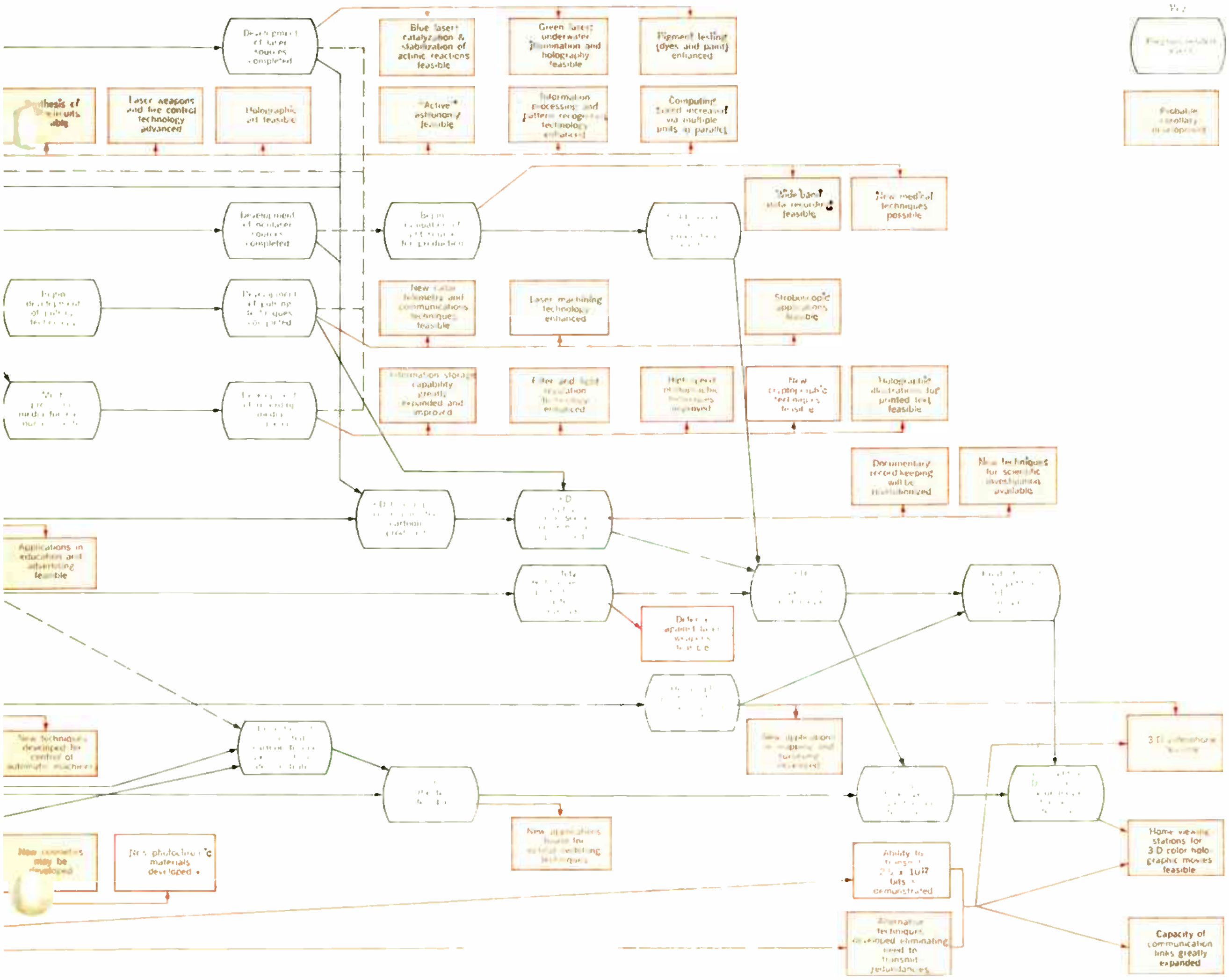


Probe I into the planning subsystem, i.e., the first four elements in the project life cycle described in Fig. 1. Some time ago, TRW recognized that corporate executives—even those who are highly trained and broadly experienced in technology—find it almost impossible to remain informed in all technologies of interest to their companies. Thus, Probe I was developed to aid corporate executives in their perception of the need for, and feasibility of, new products or services and the appropriate modification of long-range corporate plans. In addition, along with the revised corporate plan, Probe I was designed to aid R & D managers in their modification of long-range R & D plans, which in turn are matched with corporate plans and objectives by usual executive evaluation. These results would lead to a decision to proceed with specific programs.

We stated earlier that complex tasks are completed by some system, whether or not the system has been consciously defined. There is a corollary to that axiom, which goes something like this: the conscious description of a system is no guarantee that complex tasks will be completed in accordance with the process described. A review of research and development activities at TRW during the latter half of 1966 revealed that, whereas Probe I had provided an interesting checklist and a wonderful source of publicity for the company, there was no admission by anyone that it had been applied directly to R & D planning.

The following analogy helps to describe our conclusions concerning why the system did not function as planned. If one contemplates an automobile trip to a distant city, particularly one never visited before, he usually consults a map in planning the trip. So it was with technological events anticipated by our Probe I panel of experts. It was interesting, for example, to contemplate that "3-D color movies utilizing holographic techniques will be technically feasible by 1972"; but unless one considers the technological developments that must precede the event, he has no "road map" to use as a basis for planning an R & D program in this or any related technological area. Moreover, the "terminal city," the three-dimensional holographic color movie, is too big for any

FIGURE 4. SOON chart incorporating basic program for 3-D holographic color movies.



		Automotive								Electronics						Equipment							Systems					
	Probe II categories	Michigan	MRC	Ross Gear	Valve	Replacement	TP. Ltd.	Staff	Ramsey	Electronic components	Semiconductors	Capacitors	UTC	Electronic insulators	Globe Industries	Equipment labs	Accessories	J&O	Metals	Magna	Mark 46	Staff	ESD	PSD	SVD	SEID	SL	Instruments
Technologies	Electronics and electrooptics							X		X	X	X	X			X	X			X			X		X		X	X
	Materials (including coatings, fuels, and lubricants)		X		X				X	X	X	X	X	X	X	X	X	X	X	X	X			X			X	
	Mechanics and hydraulics	X	X	X		X	X	X	X	X					X	X	X							X	X		X	X
	Power sources, conversion, and conditioning		X		X			X	X	X	X		X		X	X	X				X		X	X	X			X
	Information processing							X			X											X	X		X		X	
	Instrumentation and control		X					X			X					X	X	X		X	X			X	X		X	X
	Manufacturing processes	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Systems and subsystems	Plant automation—production and business	X	X	X	X	X	X	X	X	X	X	X	X		X							X	X				X	
	Transportation	X	X	X	X	X	X	X	X	X	X					X	X					X	X	X		X	X	
	Defense and weapons (excluding missiles)		X	X				X		X	X				X	X	X	X				X	X	X	X	X	X	
	Aerospace (including missiles)		X					X		X	X	X	X	X	X	X	X	X	X			X	X	X	X	X	X	
	Oceans									X	X					X				X	X	X	X			X	X	X
	Personal and medical									X	X				X							X	X			X	X	X
	Urban and international										X		X									X	X			X	X	
	Environmental control	X	X					X			X				X	X							X	X	X		X	X

FIGURE 5. Areas of interest to TRW divisions as defined by Probe II.

one company to contemplate. It was for this reason we began to experiment with the preparation of "technological" road maps.

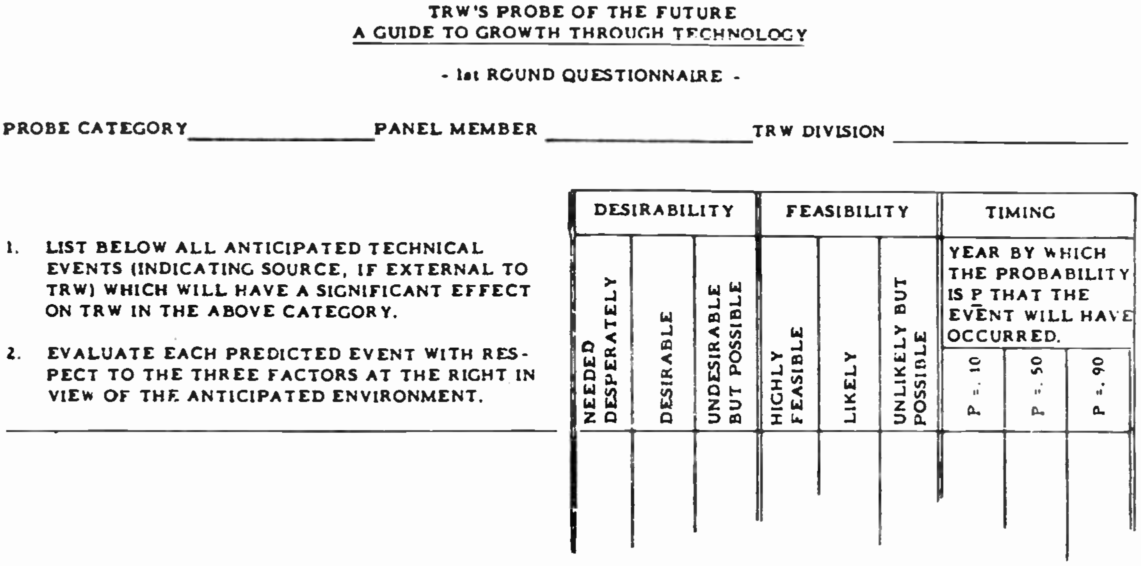
Beginning with a logic network (rounded boxes of Fig. 4), which contains a "map" of developments that must precede the technical feasibility of 3-D holographic color movies, and using this network as a skeleton, we went on to add (colored boxes in Fig. 4) those technological developments that, though not a prerequisite to the end event, were likely to occur as fallout from efforts to reach that event. The resulting chart (Fig. 4) has been labeled a SOON chart, an acronym for Sequence of Opportunities and Negatives. To be more precise, it is a Phase I SOON chart designed for product planners, indicating prerequisite developments, potential products, and by-products associated with a specific technical event.

A Phase II SOON chart, none of which has yet been prepared, is visualized that will contain excerpts dealing with the developments in a specific technology as drawn from those Phase I charts upon which that technology can be brought to bear. Such a chart would provide a weighted, time-phased indication of the developments likely to occur in the technology under study. Thus, whereas the Phase I chart provides a helpful tool for long-range product planners, the Phase II chart should provide a similar aid for those charged with the responsibility of planning technology development.

The complexity involved in formulating SOON charts is such that one hesitates to invest the effort required to construct them unless basic data provide a sound foundation for doing so. In view of Probe I's experimental nature and the possible improvements suggested by it, we decided that the desired foundation was not provided and that predictions were lacking in some of the areas in which they were most needed. As a result, Probe II was launched and is presently under way.

For this latter study, experts were chosen in each of the 15 categories shown in Fig. 5. These were selected as the minimum number into which we could fit the predictions of the panelists involved in Probe I. One or two categories were added and a few combined. The rest of the chart illustrates the matrix that was used to identify areas of interest with each of TRW's divisions. Where such interest was indicated by "group captains," one or more panelists were chosen from that division. By this method, each of TRW's divisions has become involved in both predicting and evaluating those events that may affect its future.

In implementing the asking of more penetrating questions, Probe II is being conducted in two rounds plus a follow-up as required. In round one, panel members are asked to list their forecast events on the form shown in Fig. 6, weighing each event in accordance with the factors shown. *Desirability* is to be considered from the viewpoint of the customer. Accordingly, this rating should reflect an estimate of the potential demand that would indicate the importance of the event from a marketing standpoint. *Feasibility*, on the other hand, is to be considered from the viewpoint of the producer, reflecting an estimate of the technical difficulty likely to be encountered in prerequisite developments. *Timing* reflects both an estimate of the date by which the probability is 0.5 that the event will have occurred and the degree of uncertainty associated with that estimate; i.e., that date by which there is a reasonable chance that the event may have occurred ($P = 0.1$), and that date by which the event is almost certain to have occurred ($P = 0.9$).



TRW'S PROBE OF THE FUTURE
A GUIDE TO GROWTH THROUGH TECHNOLOGY

- 1st ROUND QUESTIONNAIRE -

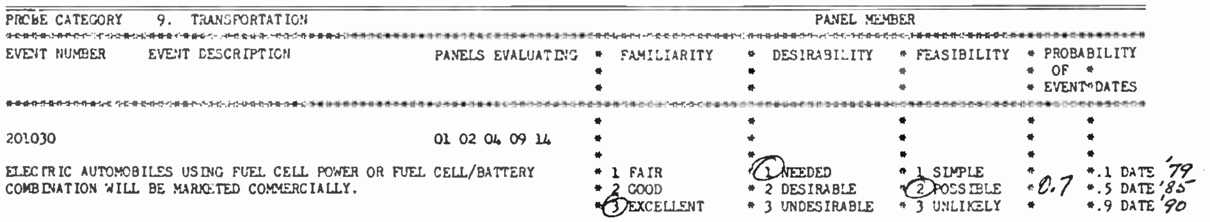
PROBE CATEGORY________________ **PANEL MEMBER** ________________ **TRW DIVISION** ________________

1. LIST BELOW ALL ANTICIPATED TECHNICAL EVENTS (INDICATING SOURCE, IF EXTERNAL TO TRW) WHICH WILL HAVE A SIGNIFICANT EFFECT ON TRW IN THE ABOVE CATEGORY.
2. EVALUATE EACH PREDICTED EVENT WITH RESPECT TO THE THREE FACTORS AT THE RIGHT IN VIEW OF THE ANTICIPATED ENVIRONMENT.

DESIRABILITY			FEASIBILITY			TIMING		
						YEAR BY WHICH THE PROBABILITY IS P THAT THE EVENT WILL HAVE OCCURRED.		
NEEDED DESPERATELY	DESIRABLE	UNDESIRABLE BUT POSSIBLE	HIGHLY FEASIBLE	LIKELY	UNLIKELY BUT POSSIBLE	P = .10	P = .50	P = .90

FIGURE 6. Round-one questionnaire in Probe II study.

FIGURE 7. Sample response of round-two pilot study.



```
PROBE CATEGORY    9.  TRANSPORTATION                                      PANEL MEMBER
*****************************************************************************************************
EVENT NUMBER      EVENT DESCRIPTION               PANELS EVALUATING  *  FAMILIARITY   *  DESIRABILITY   *  FEASIBILITY   *  PROBABILITY
                                                                     *                *                 *                *   OF  *
                                                                     *                *                 *                *  EVENT* DATES
*****************************************************************************************************
                                                                     *                *                 *                *      *
201030                                            01 02 04 09 14     *                *                 *                *      *
                                                                     *                *                 *                *      *
ELECTRIC AUTOMOBILES USING FUEL CELL POWER OR FUEL CELL/BATTERY      * 1 FAIR         * [✓ 1] NEEDED    * 1 SIMPLE       *      *.1 DATE '79
COMBINATION WILL BE MARKETED COMMERCIALLY.                           * 2 GOOD         * 2 DESIRABLE     * [✓ 2] POSSIBLE *0.7   *.5 DATE '85
                                                                     * [✓ 3] EXCELLENT* 3 UNDESIRABLE   * 3 UNLIKELY     *      *.9 DATE '90
```

Round one has yielded over 2100 events. Of these, about 1750 survived initial editing, which eliminated obvious duplications, vague statements, and descriptions of trends rather than events. Final editing by panel chairmen is expected to eliminate another 500.

The printing of round-two questionnaires and the processing of subsequent data are being carried out by computer. Each panel member will be provided with a composite list containing descriptions of all events predicted by his own panel, plus those of other panels that relate to his category.

A pilot study, utilizing two panels, has been completed. In this study, panelists evaluated events with respect to the same three factors used in round one. In addition, recognizing that the expertness of a specialist necessarily varies according to the issue, we requested the panelists to indicate their familiarity with the technologies relevant to each event on a scale ranging from *layman's knowledge* at one end, to *specialty knowledge* at the other. Subsequent experience with pilot panels led us to suggest further that panelists ignore those events about which they "know nothing." Another refinement emerging from the pilot study permits panelists to record their estimate of the probability that the event will ever occur.

Figure 7 shows a sample event and a typical response. The number 201030 references the event statement with a particular source document. The numbers 1, 2, 4, 9, 14 indicate that this event will be exposed to the panels concerned with electronics, materials, power conditioning, transportation, and urban problems. The response shows that the evaluator considers that: (1) he is a specialist in the relevant technologies, (2) event occurrence is essential to or will be demanded by a significant segment of the public, (3) a substantial technical effort will be required, (4) the probability is 0.7 that the event will occur, and (5) assuming the event *will* occur, he would anticipate it during the period from 1979 (0.1) to 1990 (0.9), with the most probable (0.5) date being 1985.

When round two has been completed and the data analyzed, those panelists whose responses deviate too far from the mean will be contacted by their chairmen to determine whether the response was due to a misunderstanding or the fact that the responder was aware of information not considered by other panelists.

It is expected that, on a statistical average, each event will have been evaluated by approximately 40 experts representing three or more different viewpoints. Present plans call for the production of the following items, all of which will be labeled "TRW Private":

1. Fifteen monographs, one for each category, will be distributed both to panelists involved and company employees needing information concerning a specific category. Each monograph will contain a list of all events developed during our probe that may have an impact upon TRW's operation in that particular category.

2. Probe II will involve a document similar to its predecessor (but more complete), where each predicted event is listed only once under that category in which the impact upon TRW is likely to be most significant. As such, this probe will be used principally by group executives and their planners. For their convenience, predicted events will be indicated for group attention in columns marked for that purpose.

TRW planners whose organizations are likely to be affected by significant events will be urged to prepare SOON charts, or their equivalent, concerning those events as a guide to both product planning and the orientation of long-range R&D programs in the relevant technologies.

We are presently concerned about the practicality of formally charting a large number of events to the degree of detail that is shown in Fig. 4. The task appears to be too formidable for the time available to planners and might lead to their abandoning the whole idea. Accordingly, we are considering the introduction of a computer-based approach to achieve at least a portion of the results that might be obtained from SOON charting—an approach by which scientists and engineers familiar with fields associated with a particular end event would predict only the major technological developments required to reach essential milestones with dates of achievement consistent with the predicted date of the end event. These could be interrelated in a composite computerized network that would provide us with a capability to: (1) make an initial cross check on the accuracy of forecast dates for different end events that are dependent upon the same technological development, and (2) maintain a continuing surveillance over the accuracy of the network. Thus we should be able to monitor developments in those technologies of importance to TRW and, for that matter, to its competitors. The extension of this technique to other variables is, of course, possible.

It is probably unnecessary to state that technological forecasting is not a panacea that will assure corporate success for its user. Thus far, in TRW, technological forecasts have been used only as checklists whose value we plan to expand. We feel that any advance notice of technological forecasting via the Delphi technique is a more useful guide than has been encountered elsewhere.

Everyone has his way of dramatizing the accelerated pace of technology. A TRW translation of another observation, from a source already cited,[2] is that of all the scientific knowledge that will be available in the year 2000, only 20 percent has been discovered thus far. This is a very sobering thought for any who are still under the impression that, although technological refinements may greatly increase in the future, most of the basic science is known today. It suggests, in fact, that the "technology explosion," which we have reviewed in retrospect, is merely a "burning fuse" and that the real explosion is yet to come.

Revised text of a paper presented at the 1968 Western Electronic Show and Convention (WESCON), Los Angeles, Calif., Aug. 20–23, 1968.

REFERENCES

1. "Year Book 1967," Electronic Industries Association, p. 2.
2. De Solla Price, D. J., *Little Science, Big Science*. New York: Columbia University Press, 1963.
3. Bardeen, J., and Brattain, W. H., "The transistor, a semiconductor triode," *Phys. Rev.*, vol. 74, July 15, 1948.
4. Roberts, E. B., *The Dynamics of Research and Development*. New York: Harper & Row, 1964, p. 304.
5. North, H. Q., and Pyke, D. L., "A probe of TRW's future—the next twenty years," TRW private document, July 1966.
6. Dalkey, N., and Helmer, O., "An experimental application of Delphi method to the use of experts," *Management Sci.*, vol. 9, pp. 458–467, 1963.
7. Gordon, T. J., and Helmer, O., "Report on a long-range forecasting study," P-2982, The Rand Corp., Sept. 1964.
8. Helmer, O., and Gordon, T. J., "Probing the future," *News Front Mag.*, vol. 9, Apr. 1965.

Harper Q. North (F) received the B.S. degree in physics from the California Institute of Technology in 1938, the M.A. degree in 1940, and the Ph.D. degree in physics from the University of California in 1947. After an early research career with the General Electric Company, Dr. North joined the Hughes Aircraft Company in 1949, and became director of the Hughes Semiconductor Division in 1953. In 1954, he founded Pacific Semiconductors Inc. (now TRW Semiconductors) and was company president until 1961. He has been TRW Inc. corporate vice president, research and development, since 1962. A Fellow of the American Physical Society, technical board member of the SAE, and member of the USASI board of directors, Dr. North is a past president of the Electronic Industries Association and the recipient of its 1966 Medal of Honor. He holds many fundamental patents in the semiconductor field and has been a consultant to the U.S. Department of Defense since 1950. Dr. North presently serves as chairman of its advisory group on electron devices.

Donald L. Pyke obtained the B.S. and M.S. degrees in engineering from Purdue University in 1946 and 1953, respectively. He has spent nine years in the academic field as an educator and administrator, including such varied posts as assistant professor, assistant dean, and acting dean of engineering at Dartmouth College's Thayer School of Engineering. After serving in the capacity of college representative at John Wiley & Sons, Inc., and technical assistant to the vice president of plans and operations at Systems Development Corporation, Mr. Pyke joined TRW Inc., where he has held the positions of director of technical staff development, director of administration—Intellectronics Labs, and assistant to the vice president of research and development during the last 12 years. He is currently active in such organizations as the ASEE and the Southern California Industry Education Council.

Public safety, the radio spectrum, and the President's Task Force on Communications Policy

Although the major metropolitan centers of the United States are in urgent need of a wider mobile-radio spectrum, much can be done to relieve the congestion by increased coordination of municipal communications and greater use of channel sharing

Robert P. Eckert, Peter M. Kelly *Kelly Scientific Corporation*

The President's Task Force on Communications Policy, with the financial support and technical cooperation of the Office of Law Enforcement Assistance, U.S. Department of Justice, retained the services of the Kelly Scientific Corporation to investigate and evaluate the problems of mobile-radio communications congestion among users of the Public Safety Radio Service. The contractual effort covered a period of approximately 20 weeks, ending on June 28, 1968. The results of the investigation are reported in this article; they include important considerations for application to such timely problems as riot control and other civil disturbances.

In evaluating the mobile-radio communications problem, the following tasks concerning Public Safety Radio Service (PSRS) radio spectrum uses in the three major metropolitan areas of Chicago, Los Angeles, and New York were considered to be of primary importance:

1. Development of facts on present usage.
2. Estimation of costs of present usage.
3. Estimation of user requirements for communications over a three- to five-year period.
4. Establishment in broad outline of alternate approaches to PSRS radio spectrum management and usages.
5. Development of approximate cost and effectiveness comparisons of the approaches developed.
6. Development of quantitative radio spectrum requirements based on those approaches that appeared practicable.

It was decided early, in a review of the time and manpower resources available, and the breadth and complexity of the task, that a very pragmatic approach to the problem would be required and a limited accuracy of results accepted. The goal was the development and interpretation of basic factual data that could provide a sound basis for decision makers in the government.

It was assumed that, in the future, local governments will function in their relationships with one another very much as they have in the past and that police and fire dispatchers would continue to operate very much as they do now. Further, although generous allowances were made for unanticipated requirements, no effort was made to investigate completely the implications of new technology. What was sought was a justifiable estimate of the probable PSRS mobile-radio channel requirements for the three-to-five-year time frame, and estimates of the amount of radio spectrum space required to meet those needs.

An effort was made to obtain a comprehensive and complete estimate of communications requirements from users in the three urban areas mentioned. The user community involved well-informed agencies and others not nearly so well informed on their anticipated needs. The stated requirements were used, therefore, to check the validity of requirements calculated on the basis of population and political organization. In reporting the results to the President's Task Force, it was observed that the requirements as determined could vary as much as a factor of two from those that might be developed out of a more exhaustive study. Such results can be considered to be within reasonable limits since land-mobile-radio networks, effectively dominated by a single dispatcher, possess considerable flexibility and can absorb a wide range of communications traffic before overloading.

Further, it was pointed out that availability to the Public Safety community of additional radio spectra would not by itself automatically insure that a sequence of actions would follow that would result in more effective mobile-radio communications for handling riots and other major emergencies. The riot problem will be discussed separately, and a number of causes identified for the communications problem in such situations.

Results of the study

Tables I, II, and III give the projected radio spectrum requirements of the immediate future for three areas

(Chicago metropolitan area, Los Angeles County, and New York City and environs) with two different basic assumptions on administration and usage of this portion of the radio-frequency spectrum. These principal results, in some cases, do not total exactly because of rounding-off effects. The basic report develops guidance for determining needed spectrum space if a part of the spectrum other than UHF is desired.

A third model used in projecting requirements is that of a common-user system applied to the entire PSRS mobile-radio communications requirements of a complete major metropolitan area. This model, as developed, is more speculative in nature than those presented in the tabulations. The timing needs for planning, design, and development are such that the concept, if it should prove to be acceptable, cannot reasonably be expected to impinge in any practical way upon the Public Safety user community within five years. Within the framework of these statements it is speculated—but by no means proved—that common-user systems of the general type described here could meet the mobile-radio communications requirements of the Public Safety community without the need for additional radio spectrum resources.

I. Summary of Chicago area requirements

Agency or Area	Approximate Additional PSRS Mobile-Radio UHF Spectrum Required, MHz: Present Usage	Approximate Additional PSRS Mobile-Radio UHF Spectrum Required, MHz: With Sharing
City of Chicago	1.5–4.2	1.5
Cook County (excluding Chicago)	7.4	0
DuPage County	1.2	0
Lake County	2.9	0
McHenry County	1.5	0
Will County	1.0	0
Total	15–18	1.5

II. Summary of Los Angeles area requirements

Agency or Area	Approximate Additional PSRS Mobile-Radio UHF Spectrum Required, MHz: Present Usage	Approximate Additional PSRS Mobile-Radio UHF Spectrum Required, MHz: With Sharing
Los Angeles County	1.7–2.3	1.7
Los Angeles City	0.4–2.3	0.4
Independent municipalities	4.2	4.2
Total	6.3–8.7	6.3

III. Summary of New York City area requirements

Agency or Area	Approximate Additional PSRS Mobile-Radio UHF Spectrum Required, MHz: Present Usage	Approximate Additional PSRS Mobile-Radio UHF Spectrum Required, MHz: With Sharing
New York City	4.0–9.8	4.0
Nassau County	1.1	1.1
Suffolk County	0.5–0.7	0
Westchester County	1.0	0
State of New Jersey	21.2	3.2
Total	27.3–33.4	8.3

For proper interpretation of all of these results, the basic report[1] should be read in detail, with particular attention paid to the validity and shortcomings of the basic data and the basis for the many approximations made.

From Tables I–III and the previous comments, it is evident that:

1. The greatest immediate needs for additional PSRS spectrum requirements are in the area of metropolitan New York.

2. Substantial savings in spectrum requirements are possible through greater user coordination, that is, through sharing.

3. The data presented on common-user systems are not oriented toward providing immediate relief of the communications congestion problem, but rather provide insight into possible long-term solutions.

4. Particular attention should be directed to the counties around Chicago, for which extreme police mobile-radio congestion is reported, but for which the results given in Table I show sufficient spectrum if sharing is practiced.

The thrust of the report is to present information that the decision makers in government can use in developing their plans of action. A variety of possible actions were presented and no case was made for choosing one over another. It was pointed out that the present spectrum shortage and the need for local government communications personnel to justify their usefulness to their superiors is leading rapidly to what a number of such personnel characterize as a "frequency war." This frequency war is being carried on at the highest levels of State and Federal government and shows promise of destroying local area cooperation and putting impossible pressures on the FCC. It was exacerbated recently when additional ultra-high frequencies were made available and a few of the more alert communicators in PSRS were able to obtain a major share of them. The range of suggested actions is shown in Table IV.

Nomenclature

Three terms are used throughout this article that should be defined:

1. *Channel*—A channel as used here is an exclusive capability for two-way voice communication between mobile units of a PSRS agency and its dispatcher. Each channel is free of interference from other agencies except under sharing agreements. It may involve one frequency assignment in a simplex system or two in a duplex system (see definitions 2 and 3).

2. *Frequency*—A frequency refers to an assignable bandwidth capable of supporting voice-frequency transmissions within the regulations of the FCC.

3. *Simplex/duplex*—The report is oriented toward radio spectrum requirements and hence is only concerned with whether a channel requires one (simplex) or two (duplex) frequencies. Thus, duplex as used in the report covers any channel that uses two frequencies. Generally, in deriving information from FCC records, it has been

IV. Range of possible actions to be taken

Hypothesis	Recommendation	Qualifying Remarks
Immediate action on spectrum reallocation is more urgent than immediate control of spectrum usage.	Reallocate spectrum immediately in accordance with the requirements developed under present usage in Tables I, II, and III.	Clearly no long-term solution; there is no assurance that interagency coordination for riot control will result.
The establishment of control is more urgent than immediate action on spectrum reallocation.	Make no reallocation to PSRS until FCC administrative controls have been strengthened.	Long delays in establishing such controls can work hardships, particularly in New York City and in the counties around Chicago.
Partial relief is desired while improved administration of spectrum is developed.	Reallocate spectrum to Los Angeles and New York areas on basis of sharing as shown in Tables I, II, and III. Determine time delays involved in Chicago area in establishing coordinated networks based on sharing. If time delays are too lengthy, provide some relief for Chicago area in excess of that listed in Table I.	A shortage of spectrum, if not carried to the extreme, does have the effect of encouraging more efficient use of the spectrum available. To date, it is the only factor working toward this end.
Long-term solutions are required for that time when no further radio spectrum is available.	Develop administrative machinery and technological basis for common-user systems.	This recommendation, to be effectively implemented, must be acted upon by local and Federal government working in concert.

assumed that HF and VHF systems are simplex and UHF systems are duplex.

Present usage of the PSRS spectrum

The three metropolitan areas studied differ in some major parameters. Table V compares the three cities in terms of their important features; for example, both Chicago and New York are seen to be densely populated whereas Los Angeles is less densely populated and is spread over a larger area.

A second feature of significance in comparing the three cities is the nature of their boundaries. Chicago and New York are homogeneous with no (New York), or few (Chicago), independent municipalities encompassed by the major metropolis. Los Angeles, by contrast, is a complicated patchwork containing within its boundaries portions of Los Angeles County and a number of independent municipalities.

New York City, until relatively recently, was not so advanced in its use of PSRS, at least in police mobile-radio communications, as Chicago, which modernized its system in 1960, or Los Angeles, which, because of its large area, has always relied more heavily upon mobile patrol, and hence on radio, than upon the foot patrolman, who is still an important factor in the eastern cities. Under Mayor Lindsay, a major communications improvement program has been instituted and, as a result, the present projected demands of the City of New York are more clearly defined than those of any other large user.

In the New York area, three counties, Nassau, Suffolk, and Westchester, and one state, New Jersey, were examined in addition to the City of New York. No clear-cut pattern of usage prevailed. One finds counties with coordinated police services, others with sharing of communications facilities undoubtedly brought about by the lack of available spectra, and still others characterized by a large number of small independent networks. In the New York area, the greatest evidence of the "frequency war" developing between users was found. Elsewhere, people are competing for spectrum space, but talking to each other; in New York, the problem is more intense.

V. Comparison of the three cities

City	Population, millions	Area, square miles	Population Density per Square Mile
Chicago	3.55	222	16 000
Los Angeles	2.82	463	6 000
New York	7.78	320	26 000

In the outlying counties of the Chicago area, a large number of small users are sharing communications facilities, particularly in the police community. The fact that communications congestion exists is self-evident, since it is not uncommon for police officers to be instructed to telephone the dispatcher after completing an assignment in order to relieve the radio network congestion. Rather straightforward calculations show that relatively mild sharing would alleviate the situation. Since sharing already exists, this result is difficult to understand. It is conjectured that, in networks with perhaps 20 agencies sharing and 20 dispatchers at work, the network is no longer dispatcher-dominated, since the term implies the use of a single dispatcher. With a large number of dispatchers, the basic assumption used in the later calculations may no longer be valid. If such is the case, then sharing may require a single dispatching center.

Los Angeles County, with emphasis upon the Los Angeles Basin, was the third area studied. This area is dominated in Public Safety communications by the City of Los Angeles and the County of Los Angeles. The situation in Los Angeles County is considerably complicated by the many small municipalities existing in proximity (see Fig. 1). The results of the user survey of projected requirements conducted in the Los Angeles area are summarized in part in Table XIV.

In all three major metropolitan areas examined, channel loading was sampled. The pattern is one of uneven loading of channels with the police mobile-patrol channels being most heavily loaded and the other Public Safety channels, including the special-purpose police channels, being much more lightly loaded.

Basis for projection of approximate user requirements

The advantage of a detailed survey of user requirements is that it does, in fact, provide details on the needs of the user. The disadvantage of such a survey is that, in each case, it presupposes the user has planned ahead, is knowledgeable regarding communications and his own community's needs, and can realistically estimate his projected requirements. This is evidently not true in all cases.

The advantage of an approximate-rule-of-thumb approach is that it does make provision for the user who may not be aware that he will have future needs. By its nature, however, a broad-brush approximate approach does not provide for local conditions. In this investigation, a mix of the two approaches was taken to arrive at user requirements. In cases of doubt, the higher level of requirement is consistently taken in order that an upper level of PSRS radio spectrum requirements can ultimately be derived.

The methods of approximation used are as follows:

1. With 13 Public Safety channels per 600 000 population, apportion the total channels among police, fire, and other PSRS users in a 6:3:4 ratio respectively.
2. Double the police and fire projected requirements to obtain an estimate of the requirements for emergency situations.

These figures are then multiplied by what are referred to here as local-condition multiplicative factors to account for the fact that a large population base may be split into a number of small jurisdictions. The method of arriving at these rules of thumb and of determining the modifying factors are described in the following paragraphs.

Present practice for police departments is to employ one mobile patrol unit for about 2000 residents, although there is quite a variance in the exact ratio of mobiles per capita. The facts relating to present practice are developed in the basic report. It has been noted that 50 to 100 cars are employed on each channel in the more densely populated precincts, 60 cars per channel being a rule of thumb accepted by metropolitan police departments. This ratio of 60 cars per channel is supported by a mathematical study[2] of the Boston Police Department, which shows that approximately 60 cars per dispatcher is the optimum division of manpower between the cars and the Boston dispatcher center. The discussion here is concerned with patrol vehicles on the road during peak periods of a normal week and does not include reserves or emergency conditions.

VI. Alternate approaches to estimating police channel requirements

Source	Magnitude of Estimation
Reasoning presented in the text	1 channel per 120 000 population
Crime Commission work	1 channel per 150 000 population
Crime Commission work	1 channel per 120 000 population

VII. Rule-of-thumb estimates for police channel requirements

	Estimate	Comments
Lower limit	1 channel per 100 000 population	Probably valid for all normal requirements including peak periods
Upper limit, for major emergencies	1 channel per 50 000 population	Not exceeded by present requests from police; hence, probably a safe upper limit

Thus, an exclusive channel should be assigned to every 60 cars and the 60 cars will serve about 60 × 2000 people depending upon the local crime rate. This figure of one police channel per 120 000 population is in approximate agreement with similar figures arrived at in the report of the President's Crime Commission. By this rule, the Los Angeles City Police Department should have 25 channels, which compares well with the fact that this police department has 13 car-to-base frequencies and says that eight more frequencies are needed.

Alternate methods of estimating police channel requirements that are in good agreement with these results have been published in the literature and are quoted here[3]:

"Approximate rules of thumb for radio channel requirements might be arrived at on the basis of population. Sample statistics taken in the course of the work indicate that on the average, over a 24-hour day, a police patrol vehicle receives about one call per hour.[4] Further, on the basis of annual calls for police service from the public, it appears that it takes a population of about 25 000 to generate the average one call per hour for a mobile patrol vehicle.* Thus, tentatively and subject to a more detailed evaluation, it may be said that it takes at least one patrol vehicle on the street at all times for each 25 000 of population. Further, studies by the Rand Corporation indicate that a police conversation lasts on the average from 25 to 60 seconds.† This checks well with the viewpoint accepted by the FCC that about 50 to 60 mobile vehicles make a maximum load for a single radio channel.

---

*Supplied by Dr. J. Kidd, National Science Foundation, on the basis of data derived from the City of Chicago.

†In an informal conversation, Al Hiebert, Rand Corporation, supplied this figure with the caution that it is based on inadequate evidence. All other measurements known to the Science and Technology Task Force have been of the number of transmissions rather than the number of conversations and so are much less useful.

Finally, these figures are based on an average loading, but a communications system must be designed for the peak load. Assume a 5:1 ratio between peak and average loading. Thus, the figures would indicate the assignment of one police radio channel for mobile patrol purposes for every 150 000 of population, on the basis of 30 cars per channel. The needs for supervisory vehicles, detective channels, and other special-purpose channels would, of course, increase this number.

"An alternative method of arriving at a population rule of thumb, suggested by Dr. Blumstein, is as follows:

"Assume an average of 2.5 police personnel per 1000 population. Assume further that half the force is on mobile patrol beats and that with one-man patrol cars it takes about five police patrol personnel for 24-hour coverage of a beat. This leads to a figure of one patrol beat per 4000 population. If we use the preceding approach, which considers the peak-to-average loading ratio, a figure of one patrol car per 5000 population is obtained. Again, assuming 30 cars to a radio channel leads to the assignment of one radio channel per 120 000 population."

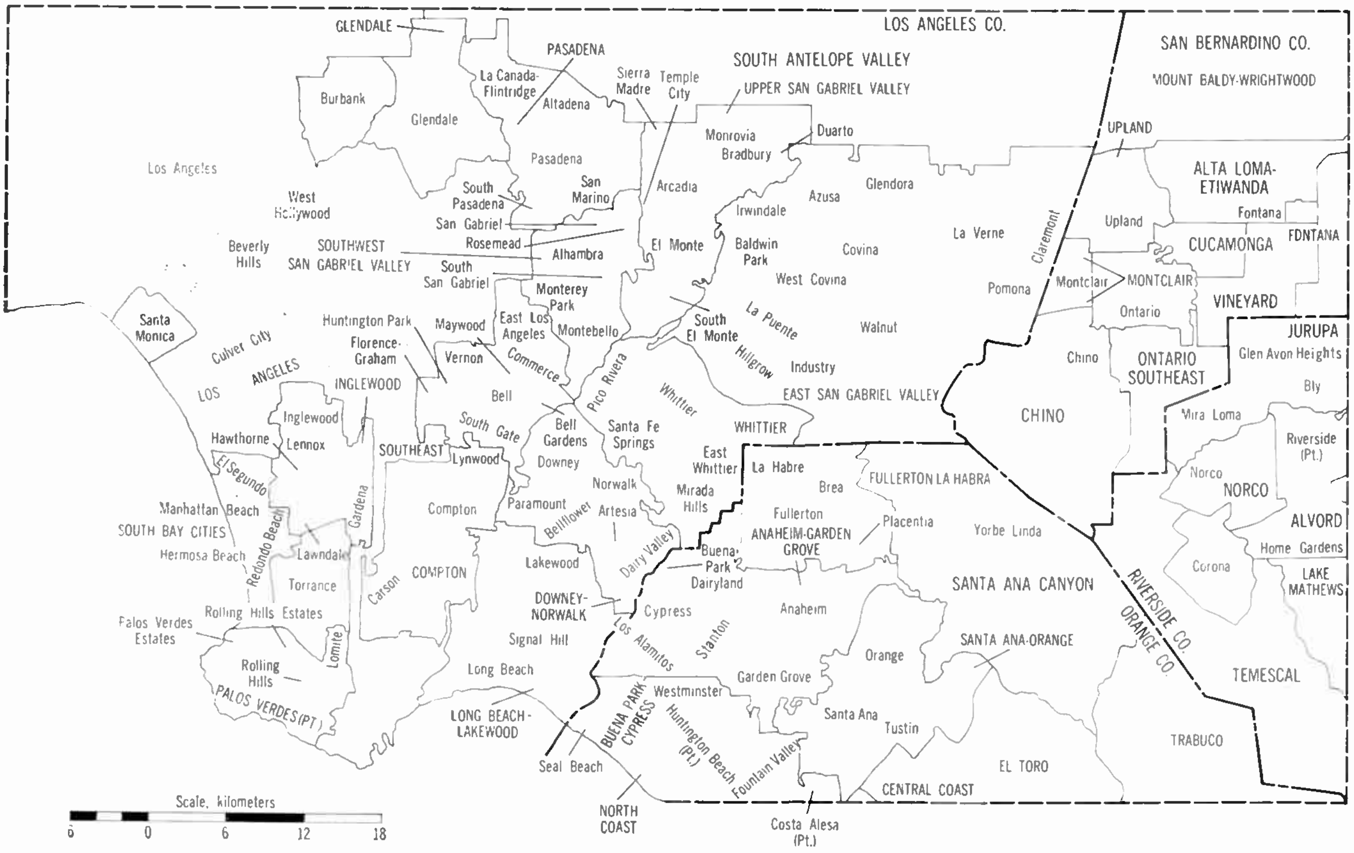
The foregoing discussion is summarized in Table VI.

In accordance with comments previously made regarding the absence of consideration of the impact of new technology, it is desirable to have some assurance that the projections do, in fact, allow some margin for additional usage. Accordingly, the basic rule for estimating channel requirements that will be used here is taken as one channel for every 100 000 in population. It can be seen from Table VI that this should serve to provide a safe margin on the basis of present usage.

Finally, consider the requirements for riot control or any other major emergency. The Office of Law Enforcement Assistance of the Department of Justice has made a study, backed by a survey of 12 major cities, of the need for personnel radios in the control of major civil disturbances.[5] It is found in such situations that if all police personnel work 12-hour shifts, all days off are canceled, and normal manpower assignments outside the area of unrest are reduced by 50 percent, then, for every 1000 total police personnel, a maximum of 342 uniformed police officers would be available for the emergency area. Further, the tactical radio network would require one personnel transceiver for each five police officers in the emergency area; counting the additional requirements for command transceivers, this amounts to a total of 78 for the original base of 1000 total police personnel.

If one assumes a channel capacity of only 30 units per channel in an emergency situation—as opposed to 60 for normal operations—then between two and three additional channels are required per 1000 police personnel for emergency tactical networks. On the basis of two police personnel per 1000 population, this is the requirement for a population of 500 000. Since the normal channel requirements were set at one per 100 000, the emergency requirement adds another ⅖ to ⅗ channel for that population base. For convenience and to insure an adequate margin, this article uses the ratio of one additional channel for emergency needs per 100 000 population. Hence,

FIGURE 1. Map of small municipalities in the Los Angeles area.



the overall projection basis is as shown in Table VII.

Note that, in each case (for both normal and emergency requirements), the rule for projection of requirements has been selected to provide a margin above the figure that the analysis provided. A reviewer of the material on fire department channels has pointed out the importance of using, in normal conditions, those channels assigned for emergency purposes in order to insure that they will be operational in the emergency. That comment applies with equal force to the police area. In a later discussion of sharing, it will be assumed that police, in an emergency, may be assigned channels normally used by nonpolice agencies and in this way obtain some of the additional channels required. In such situations, the emergency mode of operation should be tested periodically in a "fire-drill" procedure to insure operational readiness.

The approach used in estimating fire channel requirements* is to develop detailed estimates for various populations. These estimates are then compared with the 2:1 ratio of police to fire needs, respectively, which are derived both from Los Angeles area data and FCC data. The actual usage requirement is sufficiently close to the 2:1 ratio to permit the latter—one channel per 200 000 population—to be used for estimation purposes.

Channel needs for fire departments can be related to population, although there is less background available on the subject than was the case for police. For the U.S. as a whole, there were 10 800 fires per million population in 1966. For the same year, the number of alarms per million was 20 100 or approximately twice the number of fires. There is, however, considerable variance in the rates as reported by individual cities, but the average experience in the areas of New York, Los Angeles, and Chicago is not far from the national average. Assuming that fires and the need for fire radio channels subside after one hour on the average, quantitative conclusions can be reached concerning the likelihood of simultaneous fires and the corresponding need for multiple channels. Mathematics suitable for the analysis just indicated is covered in a number of textbooks on queueing theory; for an example, see Cox and Smith.[6] The analysis is carried out by assuming that one fully loaded channel is needed for each fire, and that the average incidence of fires is 1.2 per million people per hour. This rate of occurrence is in accordance with experience in the areas of New York, Los Angeles, and Chicago. Furthermore, the probability of more than one, and of more than two, simultaneous fires, and so on, can now be calculated.

Table VIII shows the number of fire channels necessary in order to keep the likelihood of more simultaneous fires than channels at one percent or less. This number of channels appears in the "Fire Fighting" column.

Table VIII also shows the additional number of channels required for dispatch operations. The assumption relative to dispatch channels is that every alarm, including false alarms or fires of minor nature, necessarily involves transmission of a number of emergency messages over an average span of ten minutes. The alarm must be broadcast and acknowledged by appropriate responding units. The first arriving company must give an on-scene situation report, as must the assigned chief officers. If required, additional help must be ordered. Companies are given their specific assignments by radio while en route in many instances, and companies relieved of duty at the scene may be placed on standby for future use by the dispatcher. The exact number of dispatch channels needed is calculated by associating a ten-minute period of radio activity with every alarm. The likelihood of more simultaneous alarms than fires is held to one percent or less; twice as many alarms as fires were assumed.

Finally, Table VIII shows the total number of channels required in the fire radio service as the sum of fire-fighting channels and dispatching channels. For comparison purposes, the rule-of-thumb estimates are also listed.

The model represented by Table VIII predicts fire channel requirements generally higher than the present practice. For instance, the seven million people living in the County of Los Angeles are served by county and municipal fire departments, which have a total of 24 frequencies assigned to them. However, within this total of seven million, the channels are not evenly divided. The unincorporated areas and contract cities served by county fire departments have a total population of 1.8 million and only three frequencies. According to Table VIII, about six additional frequencies are needed, which is more than the four additional frequencies requested by Chief Engineer Keith E. Klinger of the Los Angeles County Fire Department. Similarly, the City of Los Angeles, with a population of 2.8 million and seven simplex channels in the fire radio service, is below the projected requirement as defined by Table VIII. Thus, the individual cities smaller than Los Angeles within the county have, on the whole, more fire channels per capita

VIII. Channels required in the fire radio service as a function of population

	Channels			
Population, thousands	Fire Fighting*	Dispatch	Total	Total by Estimation (normal needs)
0–120	1	1	2	1
120–360	2	1	3	1–2
360–370	3	1	4	2
370–680	3	2	5	2–3
680–1100	4	2	6	3–5
1100–1500	5	3	8	5–7
1500–1900	6	3	9	7–10
1900–2000	7	3	10	10
2000–2400	7	4	11	10–12
2400–2900	8	4	12	12–15
2900–3200	9	4	13	15–16
3200–3400	9	5	14	16–17
3400–4000	10	5	15	17–20
4000–4400	11	5	16	20–22
4400–4500	11	6	17	22–23
4500–5100	12	6	18	23–25
5100–5600	13	6	19	25–28
5600–5800	14	6	20	28–29
5800–6200	14	7	21	29–31
6200–6800	15	7	22	31–34
6800–7300	16	7	23	34–36
7300–7400	16	8	24	36–37
7400–8000	17	8	25	37–40

* It has been pointed out by one reviewer that for larger communities the figures here are probably larger than required, since a low-power personal radio network can use the same channel at two fires simultaneously if the fires are not too close.

---

*This material has benefited from review and comments received from the National Fire Protection Association. Review does not imply endorsement of the views expressed.

IX. Los Angeles area fire channels

Agency	Population Responsibility	Channels	Comment
Los Angeles County (excluding Los Angeles City and independent municipalities)	1.8 million	3	Table XIII suggests use of nine channels
Los Angeles City	2.8 million	7	Table XIII suggests use of 12 channels
Total independent municipalities surveyed	2.5 million	14	Table XIII suggests about 21 channels (assuming about seven groups of municipalities with 350 000 people per grouping), and 12 channels if all under one jurisdiction

than the larger jurisdictions. This fits quantitatively with the requirements stated in Table VIII, which shows that more populous jurisdictions can make more efficient use of channels than a linear per-capita extrapolation would indicate. The discussion of Los Angeles County is summarized in Table IX.

Channel requirements somewhat higher than the present practice are also found for the area of Chicago. Within 50 miles (80 km) of the Loop in Chicago there are 20 frequencies licensed in the fire radio service, and the population of this area is estimated at six to seven million, which requires about 22 channels (Table VIII).

The eight million residents of New York City experienced 90 000 fires in 1966, or about 250 per day. It is conjectured that it must have been common experience to have more than ten fires simultaneously. Yet, this jurisdiction has licenses for just ten frequencies. According to Table VIII, eight channels would be required for dispatch and 17 channels for fire fighting.

Apparently, the City of New York derives as many fire-fighting channels as needed from single frequencies by using portable, low-power units at the scene.

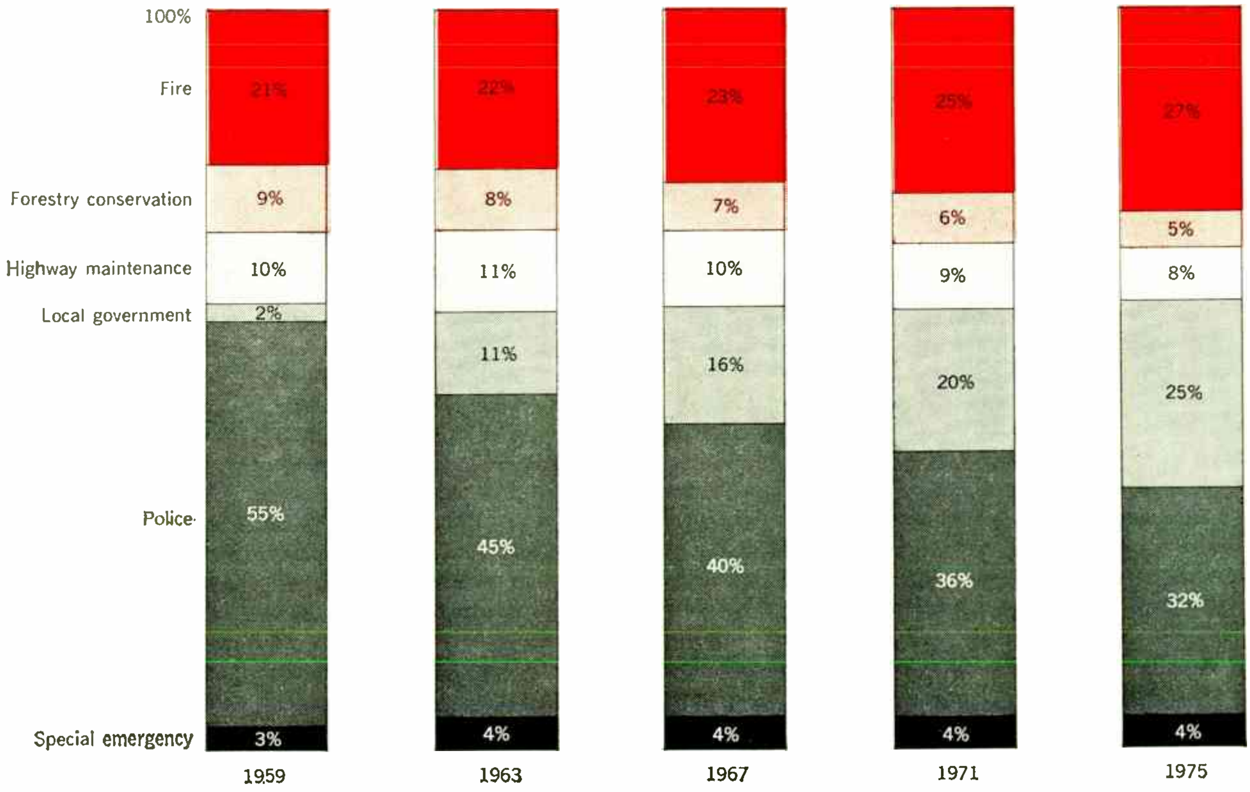
Although an accurate estimate of the fire channel requirements for a local government could be obtained by use of Table VIII, for the large amount of data handled in the report, a simpler rule was desirable; the following was used:

1. Assign fire channels for normal use in the ratio of one channel to 200 000 population.

2. Assign fire channels for emergencies (riots or major disaster) in the ratio of one channel to 100 000 population.

The foregoing is in general agreement with the more accurate estimates (see Table VIII) and provides a 2:1 ratio of police to fire channels, which is close to the usage

FIGURE 2. Trends in Public Safety Radio Service spectrum usage. Percentages are based on the number of licenses issued for mobile transmitters; data beyond 1967 are extrapolated.



ratios in 1963–1967 (see Fig. 2). The agreement diverges only for the largest population bases, and then in the direction of overstating the requirement, which, in the present context, is the desirable direction of divergence. There is, of course, only one local government agency in this analysis with a population over 3.5 million.

At the present time, fire and police have about 63 percent of the available PSRS radio spectrum (see Fig. 2). Since police and fire communications needs will continue somewhat more rapidly than that of other Public Safety users as mentioned earlier, 67½ percent of the Public Safety channel requirements will be devoted to police and fire in a 2:1 ratio, respectively.

The foregoing assumption may appear to ignore variations in local conditions, but actually it is more flexible than it appears. Police in different areas assume responsibility for additional functions—such as dispatching of ambulances—and this flexibility of interpretation of functions makes the 67½ percent rule a much less rigid one than it first appeared to be. The remaining 32½ percent is available for Public Safety users and the overall ratios are approximately 6:3:4 for police, fire, and other users, respectively.

The preceding are the ratios to be used for normal requirements, including normal peak loads. As previously stated, the estimates for police needs are doubled to allow for major emergencies. It seems reasonable also to double the fire requirements for these needs. There is less justification, however, for doing this for the remaining users; hence, the ratio used for the upper limit of requirements is 6:3:2.

It was found, when the rules of thumb just described were tentatively applied to specific areas reasonably uniformly, that the suburban areas now have many more channels than the estimates would predict. Closer examination revealed that this was due to the many small municipal agencies in the surburban areas. Note, for example, that the rules of thumb are essentially derived from the channel-to-population ratios for police of 1:50 000 and 1:100 000 for the upper and lower limits, respectively. If an area of 100 000 population is made up of ten small municipalities with independent police forces, unless there is some sharing, there will obviously be ten channels in use—one for each municipality. The estimates based solely on population ignore the existence of these independent municipalities. A multiplication factor is required to take their needs into account. This requirement is taken into consideration by means of a local-conditions multiplier applied to the lower-limit rule of thumb (13 channels in the ratio 6:3:4 for 600 000 people). For 100 000 people in ten municipalities, the multiplicative factor is 10. The approach is open to criticism because it rewards those counties whose police activities are uncoordinated by providing a projection to accommodate the wasteful usage. The approach here, however, is entirely pragmatic and does not consider the deeper implications, but simply seeks to project user needs based on present user practices.

For clarity, a review of the use of the multiplicative factor is appropriate. To determine the estimated total number of Public Safety channels required for normal use by a large municipality, the rule is

$$\text{Number of channels} \approx \frac{13P}{600\,000}$$

where P = total population. The number of channels obtained is then assigned to police, fire, and other users in the rates 6:3:4, respectively.

If, however, the average municipality size in an area is less than 100 000, a multiplying factor is used, and it is derived as follows:

$$M = \left(\frac{100\,000}{\frac{P}{N}} \right) = \frac{100\,000N}{P}$$

where N = number of independent municipalities, provided $P/N < 100\,000$.

X. Summary of projected requirements (present practices assumed)

Major Metropolitan Area	Additional Channels Required
Chicago	309–362
Los Angeles	125–173
New York	546–668

* See Tables XI, XII, and XIII for definitions of these metropolitan areas.

XI. Summary of projected requirements for Chicago area (present practices assumed)

Area	Additional Channels Required
City of Chicago	30–83
Cook County (excluding Chicago)	147
DuPage County	24
Lake County	58
McHenry County	29
Will County	21
Total	309–362

XII. Summary of projected requirements for Los Angeles County (present practices assumed)

Agency	Additional Channels Required
Los Angeles County	33–43
Los Angeles City	8–46
Independent municipalities	84
Total	125–173

XIII. Summary of projected requirements for New York metropolitan area (present practices assumed)

Agency	Additional Channels Required
New York City	79–196
Nassau County	21
Suffolk County	3–8
Westchester County	20
New Jersey	423
Total	546–668

XIV. Comparison of projected and stated user needs

	Municipalities						
	Number	Total Population	Average Population per Municipality	Additional Total Requirements	Police	Fire	Other
Projected estimate for all cities* in Los Angeles County	75	3 150 000	42 000	84	49	21	14
User survey of cities* in Los Angeles Basin	46	2 500 000	55 000	Stated: 57	31	15	11
				Estimated: 52	23	12	17

Comparison: Since the total group of municipalities and the group surveyed have average populations of less than 100 000, the requirements should be proportional to the number of municipalities rather than to population. Hence, the estimation for the surveyed group would be 46/75 × 84 or 52 additional channels. This compares well with the stated requirements of 57 channels. Note that, although projected police needs are lower than the needs stated by users, some of the additional channels in the "other" category can be used by police under present FCC rules.

* Not including the City of Los Angeles.

XV. Emergency resources under sharing

Source	Comment
30 per cent of police channels	This is based on expected practices in the newer networks. Older police networks have a much smaller ratio of special-purpose to patrol channels.
Half of PSRS channels other than police or fire	It is assumed that this source of spectrum can augment police channels by one third.

In this latter case, by substitution, the rule becomes

$$\text{Number of channels} = \frac{13}{6} N$$

Thus, for groups of small municipalities, the channel requirements projection is based on the number of those municipalities. For large municipalities, the projection is based on population.

It is apparent that the multiplying factor may be eliminated in the calculations. It has the advantage, however, that it provides, in a single number, a quick insight into the extent that additional spectra will be required in order to accommodate small municipalities. Thus, although the multiplying factor for the cities in Los Angeles County —excluding Los Angeles City—is only 3, in some New Jersey and Illinois counties it is between 10 and 20. Further, many of the additional channel requirements of these latter users disappear, since sharing has the effect of grouping populations into larger user groups and so drastically reduces these large multiplicative factors.

Estimates of user needs

Table X summarizes the estimated additional channels required in the three major metropolitan areas, assuming no changes in present usage or administration of the Public Safety radio spectrum. The approximate methods of projecting requirements that were just described are used. The details are presented in the following.

Chicago area. Table XI summarizes the results obtained for the Chicago area.

Los Angeles area. Table XII summarizes the results obtained for Los Angeles County. Since the user survey was conducted mainly inside the county, details are given for independent municipalities. Requirements of the county include channels for services provided under contract to some of these municipalities.

The detailed calculation is provided in the basic report. Account was not taken of the sharing already in existence in the police networks of the Los Angeles area. It is estimated that if these two factors had been taken into account, the needs of the independent municipalities would be reduced by about 20 channels.

New York area. Table XIII summarizes the results obtained for the New York metropolitan area.

Although New Jersey has a population of six million, as compared with New York City's eight million, the projected total requirement for New Jersey is 848 channels as compared with 173–290 for New York City. Communications needs are a reflection of organizational structure. New York City is a single agency, whereas, in the State of New Jersey, 21 counties and approximately 360 separate cities must each have their communications needs met. The sharing of facilities, which is already practiced to some extent in New Jersey, was not considered in deriving the projection. Had it been, however, the projected additional needs would have been reduced by no more than an estimated ten percent.

Comparison of estimated with stated user requirements

The user requirements developed by approximation techniques were compared with user needs as developed in the direct surveys. Table XIV is such a comparison for the independent municipalities in the Los Angeles Basin, excluding the City of Los Angeles.

The estimation techniques show that the New York area has the largest requirements of the three areas studied. In particular, the total additional requirements for New York City are 79–196 channels. New York City has developed rather complete projections of its radio requirements and they amount to 112 additional channels, which is within the range of the estimation. Thus, these checks indicate that the approximation techniques are generally consistent with the needs as defined by users.

Projections under sharing

Table XV lists the assumptions that were used in arriving at the projected user requirements under the conditions of sharing. There are reasons to believe, from the observations of riot situations, that, rather than hampering the police, interagency sharing would result in improved police coordination and improved ability to respond to wide-area emergencies.

For the suburban Chicago area, the projections under sharing indicate no additional requirements. It is known, however, that the police municipalities in that area already share facilities to a large extent. It was conjectured that, since these shared networks do not use a single dispatcher, they are not dispatcher-dominated and, hence, do not fit the assumptions developed here. Thus, the problem may be looked upon as one either of local government coordination to be solved locally, or of radio spectrum shortage to be solved in conjunction with the FCC.

The basic report goes into considerable detail on the potential for and the technical problems associated with coordination of communications and the greater use of sharing.

A review was made of the several large Federal government programs that can logically be expected to impact upon PSRS mobile-radio communications. These include the Omnibus Crime Bill of 1968, Civil Defense activities, the Federal-Aid Highways Program, the Urban Mass Transportation Act of 1964, and the Highway Safety Act of 1966. In total, these programs involve annual Federal expenditures of over $4 billion, although not all of this is directly earmarked for communications. The impact of all these activities is expected to create a need for increasing coordination of Public Safety communications. Notice has already been taken of the close relationship between problems of local government organization and coordination and the related problems in Public Safety mobile-radio communications.

Comments by reviewers

The investigation was reviewed by representatives of Federal and local government agencies, as well as by representatives of industry. By and large, the reviewers have been cooperative and favorable in their reaction to the basic approach to the task. Criticisms, when received, have been concerned with details of the presentation. In general, the comments are notable as further evidence that the complexity of the subject and the present lack of effective administration are indeed creating a sense of dissatisfaction among users. The work undertaken for the President's Task Force, if it attempted to pose as the final answer, would undoubtedly only add to the atmosphere of discontent rather than contribute to the development of solutions. Emphasis thus was placed upon the clarification of the nature of the problem, the development of facts, and the presentation of possible alternate actions that might be taken.

A number of topics not discussed here were presented in the report. The potential for common-user systems was considered, as previously mentioned. Such systems with intensive use of the geographic resource show promise of providing for the requirements with no need for additional radio spectrum. The results were considered to be quite preliminary, however, and more work is required.

Conclusion

The results of this study were concisely summarized in an early section of this article. There are a number of courses of possible action that were pointed out to the President's Task Force. It is understood that the effort was only part of a total complex of investigations, and hence that the possibility existed that any decisive actions taken might be influenced by the results obtained from these other activities.

A brief review was made of those major programs of the Federal government that will have an impact on local government radio communications. These include Civil Defense activities, programs growing out of the Omnibus Crime bill, and several programs in both the Department of Transportation and the Department of Housing and Urban Development. The impact of all these programs was seen to be that of encouraging the greater coordination of local government mobile-radio communications, entirely independent of any shortage or availability of mobile-radio spectrum.

REFERENCES

1. "Public Safety Radio Spectrum requirements," report submitted to President's Task Force on Communications Policy and U.S. Department of Justice under Office of Law Enforcement Assistance Contract LEA 68-42, June 1968.

2. Larson, R. C., "Operational study of the police response system," Tech. Rept. No. 26, Operations Research Center, M.I.T., Dec. 1967.

3. Kelly, P. M., "Communications, the police, and the Crime Commission," *IEEE Spectrum*, vol. 4, pp. 83-92, May 1967.

4. "In the matter of—an inquiry into the present and future requirements of the Public Safety Radio Service for the allocation of radio frequencies," Docket 11997, Federal Communications Commission.

5. "Estimates of police radio requirements for control of civil disturbances for 100 largest cities," internal document of Office of Law Enforcement Assistance, U.S. Department of Justice, undated.

6. Cox, D. R., and Smith, W. L., *Queues*. New York: Wiley, 1961, sect. 2.4(ii), p. 45.

Robert P. Eckert (M) received the bachelor's degree in physics from Georgetown University in 1951, and did graduate work at the University of Chicago and Johns Hopkins University. He was a professional staff member of the Research Analysis Corporation from 1957 to 1961, and then joined Booz-Allen Research, Inc., for a year. He was a systems engineer for Philco-Ford from 1962 to 1967. His work at the Kelly Scientific Corporation has included advanced techniques for satellite communications and analysis of mobile-radio communications systems.

Peter M. Kelly (M) received the B.S. in physics from Union College in 1950, and the M.S. and doctoral degrees from California Institute of Technology in 1952 and 1960, respectively. Dr. Kelly was with the Ford Motor Company, including its Aeronutronic and Philco Divisions, from 1956 to 1966, prior to founding the Kelly Scientific Corporation, which provided contractual support to the President's Crime Commission as well as the President's Task Force on Communications Policy.

Synthesizing active filters

Extensive research on the theory and practice of linear active networks has made this area one of the most attractive and promising branches of circuit theory

Sanjit K. Mitra *University of California, Davis*

The primary purpose of this article is to present a unified, comprehensive survey of accomplishments in the field of network synthesis for linear, lumped, and finite networks from 1965 through early 1968.* Because of the large number of contributions made during this period, only a sampling of the many significant works is presented here, with much of the material in the form of results. It is the author's hope that the extensive list of references included will be useful in providing the additional information necessary.

The increasing interest in the area of linear, lumped, and finite networks (more commonly known as linear active networks) can be attributed to numerous factors. Modern active-network theory is usually considered to have started with the publication of a synthesis method of inductorless active filters by Linvill in 1954.[3] In addition to the research efforts carried out since that time, a significant stimulus has been provided in recent years by the very rapidly developing field of microminiaturization.

Initial efforts concentrated on the use of a single active element for synthesis purposes because of size and cost limitations, and also as a problem of academic interest. As a result, the associated problem of sensitivity minimization was based on the same assumption, that the variable quantity was a single parameter characterizing the active element. Recently, the availability of inexpensive integrated active and passive components has changed the direction of research to a more practical and realistic course. Present work concentrates on the use of several active elements per filter and the optimization is considered as a multiparameter problem by taking into account the variation of both passive and active components.

Active-network elements

To formulate a systematic procedure for the realization of a specified network function, it is necessary to "postulate" a linear active element having simple mathematical relationships between its input and output variables. Synthesis techniques are next developed based on such idealized elements. As a result, as associated problem of active-network synthesis is the design of electronic circuits whose small-signal behavior approximates that of idealized active elements over some operating range.

The commonly used active-network elements are the negative-impedance converter, the controlled sources, the gyrator, and the operational amplifier.† The characteristics of these idealized devices have been summarized for convenience in Table I.

In this section of the article, some of the recent major contributions on the circuit realizations of active-network elements will be examined. Most pertinent works in this direction can be divided into three categories: (1) the design of operational amplifiers; (2) the design of active gyrators; and (3) the simulation of inductance.

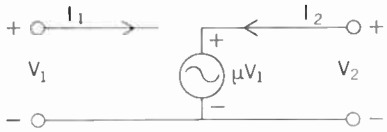
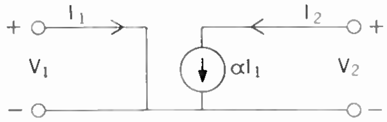
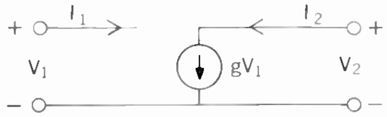
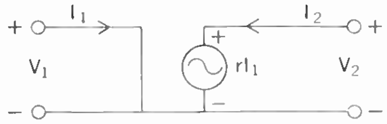
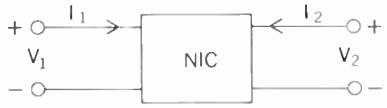
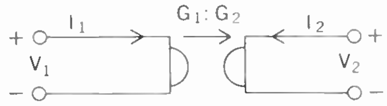
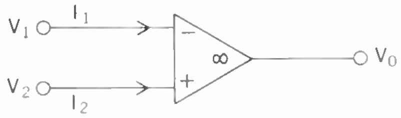
The operational amplifier. One of the most versatile active-network elements is the operational amplifier, which finds applications in such diverse areas as control systems, communications, and analog computers. This element is already making a significant impact in the design of linear circuits because of its easy availability in the monolithic integrated form as an off-the-shelf

---

*An excellent discussion of contributions to active networks made prior to 1965 will be found in Ref. 1. An elementary introduction to some selected methods of active *RC* synthesis is given in Ref. 2.

†A unified study of various active elements was recently provided by Keen.[4] Relations between these elements were investigated by Braun[5] in a companion article.

I. Some commonly used linear active elements

Active Network Element	Input–Output Relations	Network Symbol
Voltage-controlled voltage source, VCVS or voltage amplifier	$V_2 = \mu V_1$ $I_1 = 0$	
Current-controlled current source, CCCS or current amplifier	$I_2 = \alpha I_1$ $V_1 = 0$	
Voltage-controlled current source, VCCS	$I_2 = gV_1$ $I_1 = 0$	
Current-controlled voltage source, CCVS	$V_2 = rI_1$ $V_1 = 0$	
Negative-impedance converter, NIC	$V_1 = k_1 V$ $I_2 = k_2 I_1$	
Positive-impedance inverter, PIV (active gyrator)	$I_1 = G_2 V_2$ $I_2 = -G_1 V_1$	
Operational amplifier	$V_0 = A(V_2 - V_1)$ $A \to \infty$ $I_1 = I_2 = 0$	

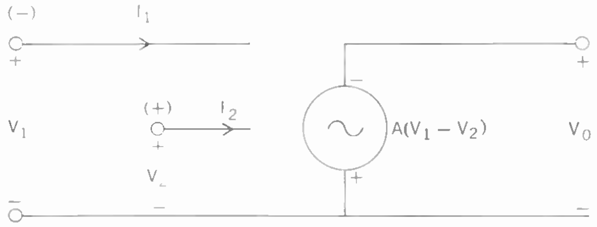
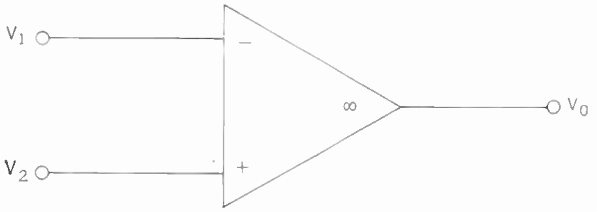


FIGURE 1. Controlled-source model of a differential-input operational amplifier.

FIGURE 2. Symbolic representation of a differential-input operational amplifier.



item.[6-8] Probably the most famous commercially available amplifier is the Fairchild μA 709, designed by Widlar.

Definition. The ideal operational amplifier is basically a two-input, infinite-gain, voltage-controlled voltage source. The output voltage is always of the same polarity as the voltage at one of the input terminals and is of opposite polarity to that of the other input terminal. The model of an ideal operational amplifier is shown in Fig. 1, with reference to which the idealized amplifier can be precisely defined as

$$V_0 = A(V_2 - V_1) \qquad A \to \infty \qquad I_1 = I_2 = 0 \quad (1)$$

$$V_0 \to 0 \text{ when } (V_2 - V_1) \to 0 \quad (2)$$

where A is assumed to be independent of frequency, temperature, and input levels. It is seen from the representation of Fig. 1, that the input impedance is infinite and the output impedance is zero. Equation (2) is more commonly called the "zero offset" requirement. The symbolic representation of the operational amplifier is shown in Fig. 2.

Design considerations.[9–13] The integrated monolithic construction of circuits offers many advantages to the circuit designer. Typical advantages are increased system reliability and reduction of cost, size, weight, and power consumption. In addition, an integrated circuit provides closer matching and tracking of active and passive components over a wide temperature range, little restriction on the number and geometry of the active devices, and excellent thermal coupling through the circuit. However, associated with these advantages there are at present several practical limitations that limit full integration of all linear circuitry. For example, integrated high-quality inductors are not available and integrated resistors and capacitors are restricted to moderate values and have wide tolerances. Circuit adjustments are difficult to achieve economically. It is also difficult with present-day technology to manufacture complementary transistors on a single silicon substrate economically.

Several novel circuit-design techniques, which take advantage of many inherent features of integrated-circuit fabrication, have been developed in recent years to overcome some of these limitations. It is useful, for example, to replace large resistances by transistor current sources and to let differential emitter-coupled amplifiers take the place of large bypass capacitors. These techniques have made possible the commercial production of inexpensive integrated monolithic operational amplifiers.

An integrated operational amplifier is a complex circuit, containing more active components than a comparable discrete amplifier, but in general providing performance features exceeding those of its discrete counterpart. The μA 709, for example, contains 14 transistors and 15 resistors. A comparable discrete circuit can be designed with less than a third of this number of components.

The primary building block in the design of integrated operational amplifiers is the versatile differential-amplifier stage shown in Fig. 3,[12] which amplifies differential-mode input voltages while suppressing interfering common-mode signals. Several attractive features of this circuit from the point of view of monolithic construction are: (1) there is inherent matching of the base-to-emitter voltages and the short-circuit-current gains of the two transistors, providing excellent balance between the differential-amplifier inputs in the face of changes in signal levels and temperature; (2) the use of large-value resistors is avoided, and the gain of the circuit can be made to depend only on resistance ratios; (3) there are no capacitive elements; (4) there is a provision for dc coupling. The constant current sink that provides the emitter currents is usually a transistor, which effectively acts as a common-emitter resistance having a high ac-to-dc impedance ratio, and thus common-mode rejection is improved. The circuit configurations of most of the commercial operational amplifiers are essentially a cascade combination of differential-amplifier stages with an appropriate output stage.

Various aspects of design of integrated operational amplifiers are detailed in Refs. 10, 11, 13, and 14.

Practical operational amplifier characteristics. The practical operational amplifier is a nonideal device, characterized by a frequency-dependent voltage gain monotonically decreasing with frequency from a high value at dc, and having finite input and output impedances. There is usually an upper limit on the input and output signal levels. Other sources of nonidealness include, typically, the finite input offset voltage and current (and their "drift"), common-mode rejection error, and finite unequal common-mode impedances of the two input terminals.

Even though the characteristics of a practical operational amplifier are far from its idealized counterpart, for most linear applications, these practical amplifiers are always used with negative feedback, thus improving their actual performances and yielding satisfactory results.

The gain characteristic in almost all integrated operational amplifiers rolls off faster than 12 dB/octave, which inevitably creates a stability problem. For stable operation, external compensating networks must be added to these amplifiers to modify their gain characteristics. In using the μA 709, it is usually necessary to use a resistor and two capacitors for compensation purposes.

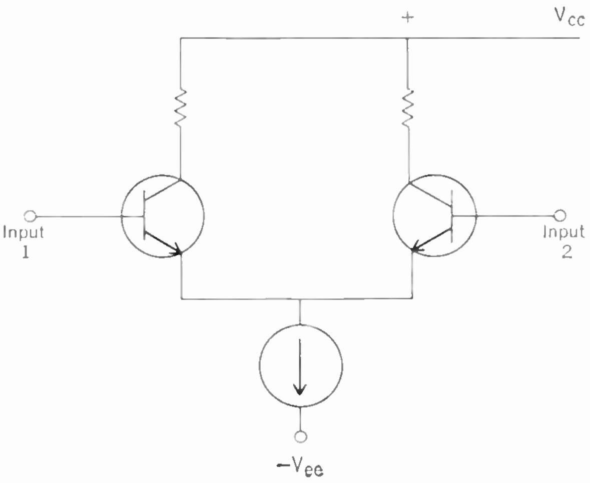
A comparative study of many practical operational amplifiers, made by Stata,[15] may be helpful in selecting the right amplifier for a particular use. If these devices are to be utilized efficiently, it is absolutely essential to understand their potentials and limitations. For a discussion on these aspects, many excellent publications are available from several manufacturers. (See, for example, Refs. 10, 11, and 16–18.)

Active gyrator. Unlike the operational amplifier, other active elements—such as negative-impedance converters or controlled sources—are not yet commercially available in integrated form. However, circuit realizations suitable for the integration of some of these elements have been proposed in recent years. One such element is the active gyrator, which has evoked considerable interest during the last three years from the point of view of design and application in inductorless filter realizations. In this section, the properties and methods of design of active gyrators are briefly reviewed.

Definition.[19] The active gyrator is a two-port described by the following short-circuit admittance matrix:

$$[y] = \begin{bmatrix} 0 & g_1 \\ -g_2 & 0 \end{bmatrix} \qquad g_1 \neq g_2 \qquad (3)$$

FIGURE 3. The differential-amplifier stage.



where g_1 and g_2 are real positive numbers, and are known as the *gyration admittances*. From the definition it is apparent that the active gyrator is essentially a positive-impedance inverter. When $g_1 = g_2$, Eq. (3) defines the more familiar passive lossless gyrator of Tellegen.

One simple but useful application of the active gyrator is in realizing an inductance. This will be elaborated later. Another attractive feature of the gyrator is that it is (theoretically) absolutely stable when designed with equal gyration admittances.

Design methods. Consider a few of the proposed design methods.

The most popular approach is based on the controlled-source representation of the active gyrator on the y basis, shown in Fig. 4, which indicates that an active gyrator can be constructed by paralleling two voltage-controlled current sources of opposite gain. Several transistorized circuits based on this idea have been advanced. An example of such a gyrator realization is the basic circuit, shown in Fig. 5(A), which has been investigated in detail by Shenoi.[20] An alternate three-transistor realization based on Fig. 4 was advanced by New and Newcomb[21]; its circuit is shown in Fig. 5(B). The same configuration was also independently proposed by Mitra and Howard.[22] A dc-coupled version of Fig. 5(A), which is suitable for monolithic construction, was suggested recently.[23] In this modification, the load resistors were replaced by constant-current sources and the base-biasing resistors were eliminated. This modified gyrator circuit realized inductances up to 200 henrys and with a Q of up to 200 at frequencies of about 10 to 20 Hz. Several modified versions (dc-coupled and self-biased) of the circuits of Fig. 5(A) and (B) were recently tested by Bach and Carlson,[24] and compared with respect to their effectiveness in simulating inductances. Resonant circuits having a Q in excess of 400 at medium frequencies have been obtained.

As will be shown, the quality of the simulated inductor using a practical gyrator is inversely proportional to the actual value of the short-circuit admittance parameters y_{11} and y_{22} of the gyrator. As a result, to obtain high-Q inductors, it is necessary to design gyrators having extremely small y_{11} and y_{22}. Basically, two paths have been followed to achieve this end.

In one approach, efforts were made to improve the design of the voltage-controlled current sources (VCCS) so that very large input and output impedances might be obtained. Sheahan and Orchard have reported two such circuits. For low-frequency applications, they built a gyrator circuit employing VCCS designed with bipolar transistors.[25] At frequencies above a few kilohertz this type of gyrator circuit is useless, owing to the excessive and very temperature-dependent phase shift of the VCCS. This problem can be avoided by using instead MOSFET's (metal-oxide semiconductor field-effect transistors) at the input of each VCCS. Sheahan and Orchard[26] report gyrator circuits using MOSFET's and bipolar transistors simulating inductances of Q above 400 at 1–50 kHz, with very low temperature sensitivity at frequencies below 1 kHz. Another gyrator design based on Fig. 4 was also reported in Ref. 23. Here both the VCCS use a differential stage at the input, which offers several advantages from the integrated-circuit point of view. Practical gyrator circuits based on Fig. 4 have been reported by several other authors. The circuit of Chua and Newcomb[27] is the only gyrator circuit that has been built in a mono-

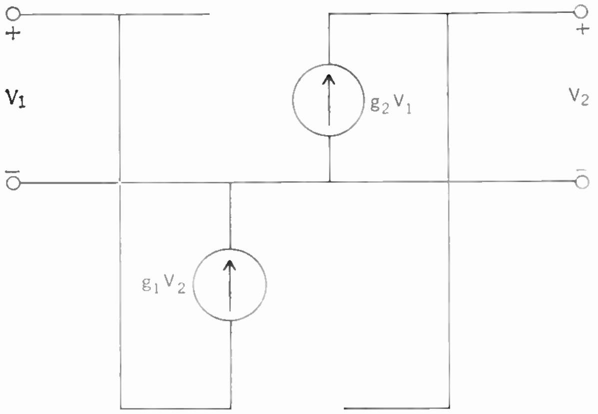
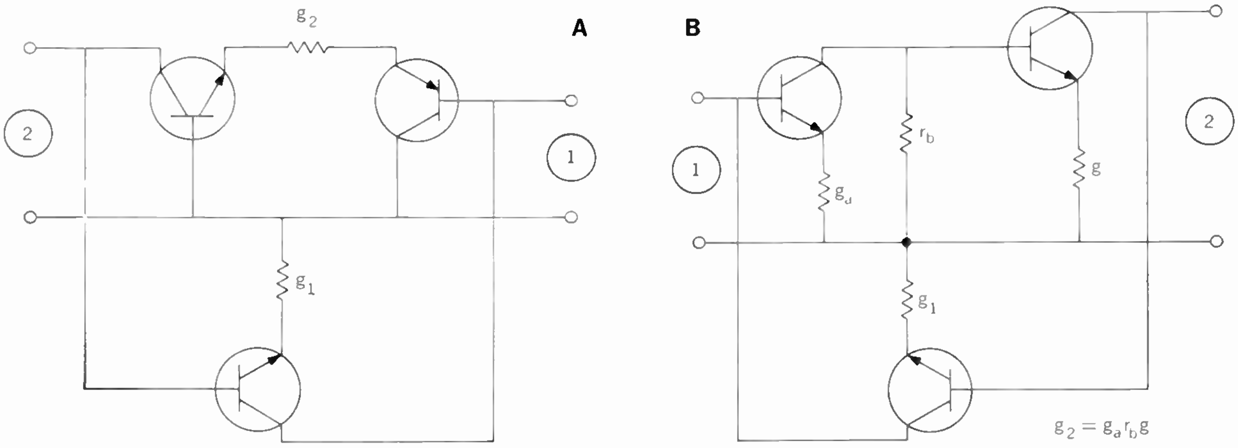


FIGURE 4 (left). Controlled-source model of the active gyrator.

FIGURE 5. Three-transistor gyrator realizations based on controlled-source model shown in Fig. 4.



lithic integrated form; the integrated gyrator-simulated inductors have stable Q's of the order of 50 at frequencies below 65 kHz. Their circuit is a slight modification of the circuit of Rao and Newcomb proposed earlier.[28] Additional gyrator circuits are described in Ref. 29.

In the second approach, circuits are designed such that the values of y_{11} and y_{22} can be individually controlled and adjusted to zero after the circuit has been built. This is usually achieved by using negative impedances to cancel the undesirable terms. As we shall show later, gyrators of the second type are not satisfactory in practice for sensitivity and stability reasons, and can be used only to simulate low-Q inductances.

Yanagisawa advanced two techniques for the design of lossless gyrators that essentially make use of the negative-impedance approach.[30,31] One of his proposed circuits, shown in Fig. 6, is essentially a modification of the basic circuit of Fig. 4. The y-parameters of this active two-port are given as

$$\begin{aligned} y_{11} &= y_{22} = \frac{1}{R} - \frac{g_bR_1 - 1}{R_2 + R_1(2 - g_bR_1)} \\ y_{12} &= g_a - \frac{1}{R_2 + R_1(2 - g_bR_1)} \\ y_{21} &= -\frac{1 + g_bR_2}{R_2 + R_1(2 - g_bR_1)} \end{aligned} \tag{4}$$

These equations imply that y_{11} and y_{22} can be made equal to zero by controlling the two shunt resistors R. Moreover, g_a appears only in y_{12}; hence, y_{12} can be controlled independently of y_{21} and made equal to $-y_{21}$, if required. Design equations can also be derived from Eq. (4). Details of this procedure and experimental results will be found in Ref. 30.

Another novel design procedure, which also allows the designer to make y_{11} and y_{22} equal to zero, was proposed by Yanagisawa and Kawashima.[19] This method is based on the use of a nonreciprocal negative-impedance inverter, as indicated in Fig. 7, and is restricted to the design of active gyrators only ($y_{12} \neq -y_{21}$).

Gyrators constructed by cascading a negative-impedance converter with a negative-impedance inverter are well known and have been re-examined recently.[32] A description of another gyrator circuit of this type is given in Ref. 33.

Several authors have considered the design of three- and four-terminal gyrators using operational amplifiers.[34-36]

Simulation of inductance. The primary reason for recent interest in gyrator design is its usefulness in simulating an inductor. Consider the active gyrator defined by Eq. (3). If port 2 is terminated by a capacitor of C farads, the input impedance seen at port 1 is equal to sC/g_1g_2 henrys. Thus a simple way of realizing an inductorless filter would be to realize first a conventional RLC filter having the prescribed transfer function and then replace the inductances by simulated inductors. Such an approach has many advantages, which will be discussed subsequently.

At this point the quality of the simulated inductor will be considered, and several other methods proposed for inductance simulation will be reviewed.

Q of gyrator-simulated inductance.[37] A practical gyrator is naturally a nonideal device and can be characterized by a y-matrix as follows:

$$[y] = \begin{bmatrix} G_1 & g_a \\ -g_b & G_2 \end{bmatrix} \tag{5}$$

where G_1 and G_2 are small numbers accounting for nonzero y_{11} and y_{22}. When this gyrator is terminated at port 2 by a capacitor C, the input impedance seen at the other port is identical to that of a lossy inductor, whose equivalent inductance is

$$L_{eq} = \frac{\omega C g_a g_b}{(G_1G_2 + g_ag_b)^2 + \omega^2C^2G_1^2} \tag{6}$$

and the Q is given by

$$Q = \frac{\omega C g_a g_b}{G_1G_2^2 + G_2g_ag_b + \omega^2C^2G_1} \tag{7}$$

Thus a gyrator-simulated inductance behaves like a coiled-wire inductor.

The maximum value of Q is

$$Q_{\max} = \frac{g_ag_b}{2[g_ag_b + G_1G_2]}\sqrt{1 + \frac{g_ag_b}{G_1G_2}} \tag{8}$$

from which it can be seen that to obtain high-Q inductors, the gyrator must be designed with values of G_1 and G_2 that are very small in comparison to the gyration admittances.

An alternate measure of the quality of the simulated inductor can be the maximum value of the Q of the

FIGURE 6. A controlled-source model suitable for realizing a lossless gyrator.

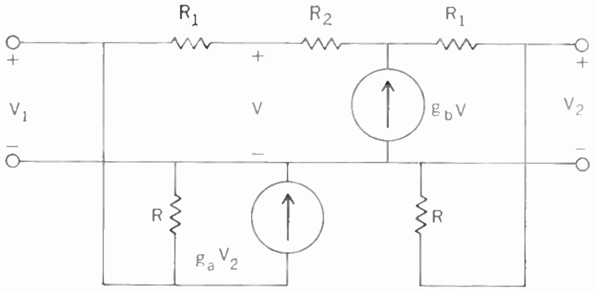
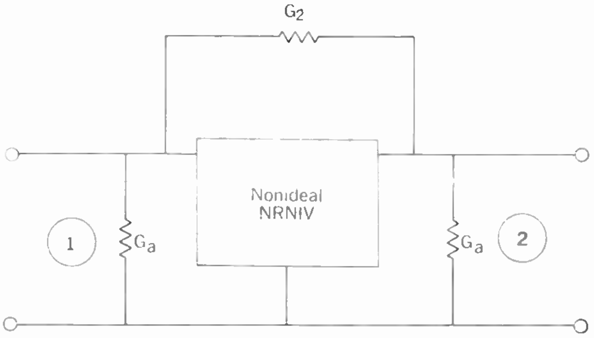


FIGURE 7. Active gyrator using nonreciprocal negative-impedance inverter (NRNIV).



resonant circuit obtained by connecting a capacitor across the simulated inductor.[38] This factor is given as

$$Q_{\text{resonance}} = \frac{1}{2}\sqrt{\frac{g_a g_b}{G_1 G_2}} \tag{9}$$

and was used by Bach and Carlson[21] to evaluate various practical integrable gyrator circuits.

Gensel[39] explains in detail why a gyrator circuit based on the use of negative impedances for cancellation purposes is unsuitable for high-Q induction simulation. A simple explanation of this is as follows. An inductance simulated by using a gyrator of the second type will have an effective impedance given as

$$Z_{\text{eff}} = j\omega L_{\text{eq}} + R_A - R_B \tag{10}$$

where R_B (or R_A) is adjusted later to make $R_B \to R_A$. The Q sensitivity with respect to R_B is

$$S^Q_{R_B} = \frac{R_B}{R_A - R_B} \tag{11}$$

and hence is extremely large as R_B approaches R_A; that is, for high-Q inductors.

It appears that a gyrator of the first type is the only possible way of getting high-Q simulated inductors. However, at high frequencies, the phase shifts introduced by the transistors create a problem. The effect of the phase shift is to introduce a negative resistance in series with the simulated inductor and as a result makes it highly sensitive if the original Q is very large.[38, 40]

Additional methods of inductance simulation. In addition to the gyrators, the active transformer (also known as the positive-impedance converter[41]) can also be used to simulate inductors. The active transformer is a two-port described by the following h-matrix:

$$[h] = \begin{bmatrix} 0 & \pm\alpha \\ \mp\beta & 0 \end{bmatrix} \qquad \alpha \neq \beta \tag{12}$$

where α and β are real and nonnegative numbers. A generalized active transformer is obtained by making α and β complex numbers. To illustrate their use in simulating inductances, consider the generalized active transformer of Fig. 8, which is characterized by

$$[h] = \begin{bmatrix} 0 & 1 \\ -\dfrac{Z_1 Z_3}{Z_2 Z_4} & 0 \end{bmatrix} \tag{13}$$

The input impedance seen at port 1, when port 2 is terminated by a resistance R_5, is

$$Z_i = \frac{Z_1 Z_3 R_5}{Z_2 Z_4} \tag{14}$$

It can be seen from Eq. (14) that if we make either Z_2 or Z_4 a capacitance and the remaining Z's resistances, the resistively terminated two-port of Fig. 8 will simulate an inductance. The circuit of Fig. 8 was originally advanced by Riordan[42] and appears promising for several reasons. The active elements used are operational amplifiers. Equation (14) does not have any difference terms, hence sensitivity should not be any problem. A circuit having similar properties was also advanced by Lampard and Rigby.[43]

Realization of grounded inductance has also been considered by Su[44, 45] and by Keen and Peters.[46]

The circuits given earlier for inductance simulation, including the gyrators of the previous section, are useful only for simulating grounded inductors, and hence cannot be used to replace "floating" inductors.

One way to circumvent this problem would be to use two grounded gyrators per floating inductor, as shown in Fig. 9.[47, 48] But, as pointed out by Sheahan,[49] this scheme is not satisfactory from a practical point of view because precise matching of the gyration admittances of the two gyrators is required. Assuming nonidentical active gyrators, he has shown that the equivalent circuit of the structure of Fig. 9 has additional parasitics to ground whose effect would be serious in sharp-cutoff filters. Sheahan has proposed a novel circuit to convert a grounded gyrator into a floating gyrator and thus avoid the sensitivity and adjustment problems that can result from the two-gyrator method of Fig. 9.

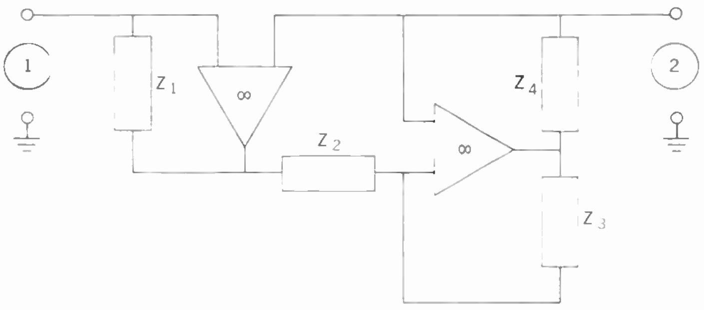
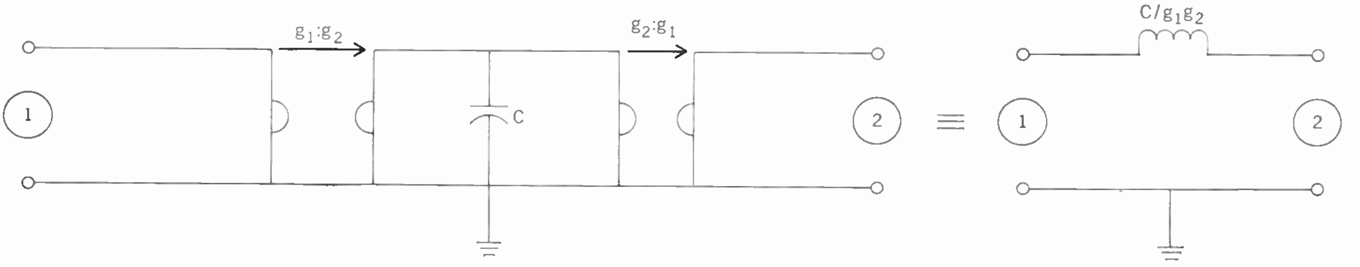


FIGURE 8 (left). Active-transformer realization employing operational amplifiers.

FIGURE 9. Simulation of a floating inductor through the use of two active gyrators.



Another feasible method would be to use gyrator circuits with floating ports, as suggested in Ref. 50. Riordan[49] suggests the use of two grounded inductors in a "back-to-back" fashion to realize ungrounded inductance. However, the effective inductances of the two simulated inductors must be exactly equal, otherwise any unbalance would appear as a parasitic inductance to ground. Realization of the floating inductor has also been considered by Keen and Peters.[51]

A novel method was advanced recently by Deboo[34] for realizing ungrounded inductance. His proposed circuit, shown in Fig. 10, has a y-matrix given by

$$[y] = \begin{bmatrix} \dfrac{1}{sCR^2} & -\dfrac{1}{sCR^2} \\ -\dfrac{1}{sCR^2} & \dfrac{1}{sCR^2} \end{bmatrix} \tag{15}$$

which is identical to that of a floating inductance of value CR^2 henrys.

Active-network synthesis methods

Considering some of the recently developed active-network synthesis methods, it is again convenient to classify them into various groups. There are several ways in which this might be done. One such classification is based on the single-loop feedback representation of the active network configuration with the active parameter in the forward path of the feedback loop.[52] Thus, active RC synthesis procedures are classified as positive-feedback or negative-feedback types, depending on the sign of the active parameter. In addition, the form of the return difference function has been used to aid the classification.

Here a different basis for classification is employed. Synthesis methods are grouped according to the way in which the passive and the active parameters are identified. Accordingly, there are three possible design approaches: (1) the conventional active-synthesis approach; (2) the coefficient-matching approach; and (3) the simulated-inductor approach. In the first two, the synthesis techniques can be further grouped according to the type of active element used.

Conventional active-network synthesis approach. In the conventional approach, the active synthesis is based on a general active-network configuration containing one or more idealized active elements. The parameters characterizing the passive subnetwork are obtained by suitably decomposing and partitioning the specified network function. Active realization is completed by realizing the passive subnetwork following standard passive-synthesis procedures, and then connecting the active elements to the realized passive subnetwork in the way indicated in the original active-network configuration. In this category fall the well-known methods of Linvill,[3] Yanagisawa,[53] Kinariwala,[54] Sandberg,[55] Kuh,[56] Horowitz,[57] and Sipress.[58]

The key to almost all of the conventional active-synthesis methods is the $RC{:}-RC$ decomposition reviewed next.

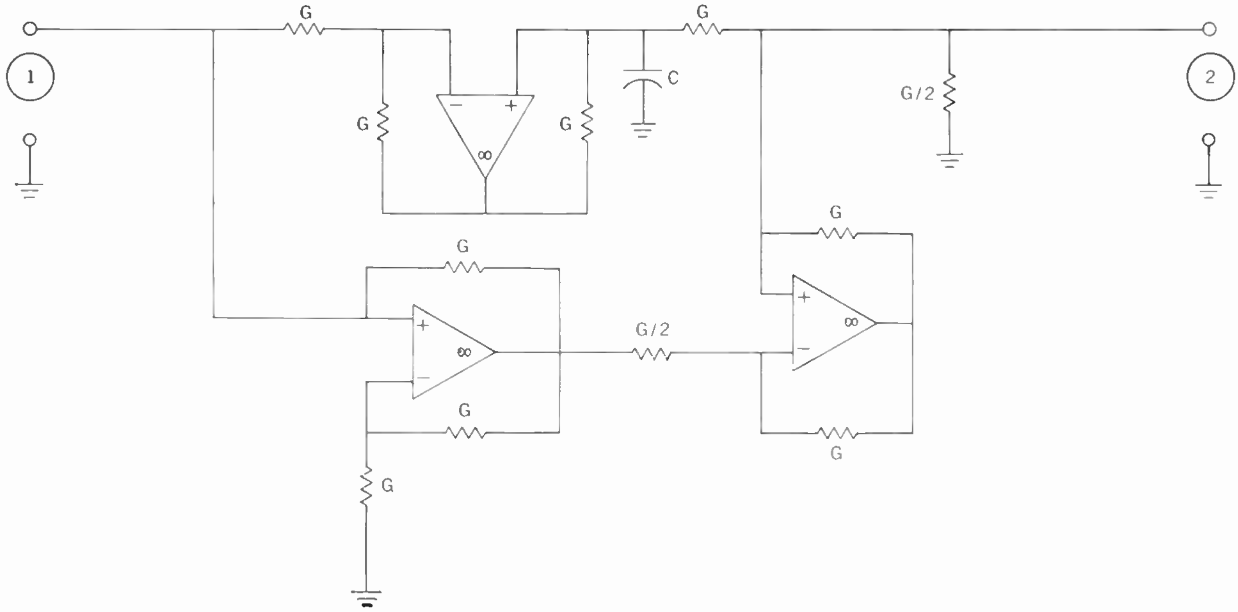
$RC{:}-RC$ decomposition. According to the $RC{:}-RC$ decomposition, a real rational function $D(s)/Q(s)$ having unrestricted zeros anywhere in the s-plane and simple poles restricted only to the negative real axis can be expressed in either of the following two forms:

$$\frac{D(s)}{Q(s)} = Z^A_{RC} - Z^B_{RC} \qquad \text{if } Q(s)^0 > D(s)^0 \tag{16}$$

$$\frac{D(s)}{Q(s)} = Y^a_{RC} - Y^b_{RC} \qquad \text{if } Q(s)^0 > D(s)^0 + 1 \tag{17}$$

where Z^A_{RC} and Z^B_{RC} (Y^a_{RC} and Y^b_{RC}) are passive driving-point impedances (admittances). The degree of a polynomial $D(s)$ is denoted here by $D(s)^0$. The decompositions given in Eqs. (16) and (17) will be referred to as

FIGURE 10. Realization of floating inductance using three operational amplifiers.



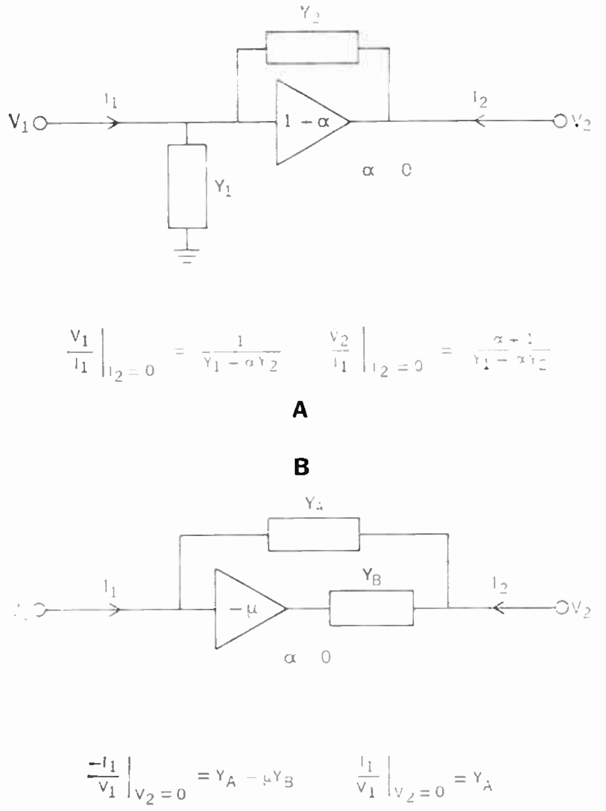


FIGURE 11. A—Basic "pole" section, suitable for realizing complex poles. B—Basic "zero" section, suitable for realizing complex zeros.

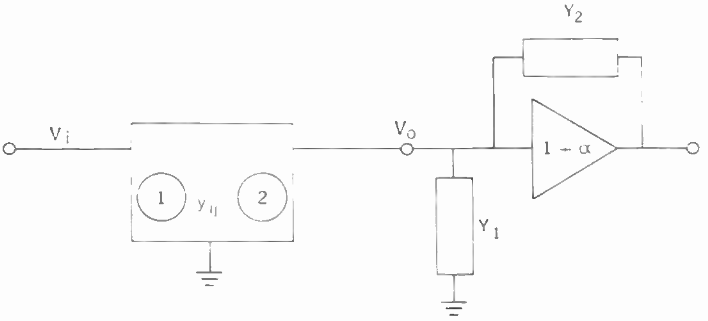
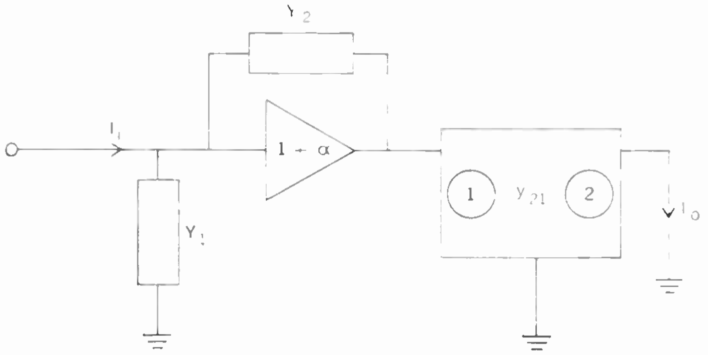


FIGURE 12. Kuh's active RC configuration.

FIGURE 13. Hakim's active RC configuration.



RC: −*RC* decompositions; they are obtained by making a partial fraction expansion of $D(s)/Q(s)$, or $D(s)/sQ(s)$ as appropriate, and grouping positive and negative terms separately.

Synthesis using voltage amplifiers. Only the finite-gain voltage amplifier (VCVS) is considered here. The following discussion treats the operational amplifier as an active element.

A unified and systematic development of active *RC* network synthesis procedures using voltage amplifiers as the active elements was recently provided by Mitra.[59] The development depends on the use of basic active "pole" and "zero" sections. Various connections of these circuits with each other and with passive *RC* two-ports have led to some well-known active-synthesis techniques in addition to several new ones. The two basic circuits and their pertinent network functions are shown in Fig. 11.

The active *RC* configuration proposed by Kuh[56] is obtained by cascading an *RC* two-port with the basic pole section as indicated in Fig. 12. This circuit has a voltage-transfer ratio given as

$$t_v = \frac{V_o}{V_i} = \frac{-y_{21}}{y_{22} + Y_1 - \alpha Y_2} \qquad (18)$$

where y_{21} and y_{22} are the short-circuit parameters of the *RC* two-port. Synthesis is achieved by dividing the numerator and denominator polynomials by a suitable polynomial having distinct negative real roots, expressing the denominator function by *RC*: −*RC* decomposition, and grouping the terms properly. Since the transmission zeros are realized by the *RC* two-port, this method cannot realize positive-real-axis transmission zeros in an unbalanced configuration. Moreover, t_v can be realized only within a multiplicative constant. In the case of transfer functions having complex transmission zeros, the *RC* two-port has to be realized by following either Guillemin's parallel-ladder method or the Fialkow–Gerst method or Dasher's method, which are somewhat difficult to use. An interesting feature of Kuh's method is that, by suitably choosing the *RC* one-ports Y_1 and Y_2, the input, output, and feedback impedances of the voltage amplifier can be accounted for, thus making possible the use of a nonideal amplifier. If the output impedance of the voltage amplifier is negligible, the output of the amplifier can be taken as the output of the filter, allowing the cascading of the filter stage without an additional buffer amplifier. Kuh's method was adapted for analog-computer simulation by McVey.[60] Several second-order realizations based on Kuh's active *RC* configuration and method will be found in Ref. 61.

The active *RC* structure of Hakim,[62] shown in Fig. 13, is obtained by cascading the basic "pole" section with an *RC* two-port in the reverse order. This structure has a current-transfer ratio given as

$$t_I = \frac{I_o}{I} = \frac{(1 + \alpha)y_{21}}{Y_1 - \alpha Y_2} \qquad (19)$$

where y_{21} is the short-circuit transfer admittance of the *RC* two-port. The synthesis method is similar to that of Kuh, except that Y_1 and Y_2 are identified by straightforward *RC*: −*RC* decomposition of the denominator polynomial. Hakim's method has all the drawbacks of Kuh's method except that it does not impose any restriction on the multiplicative constant of the transfer function. Only the effect of the input and feedback impedances of the amplifier can be precorrected here. In general, Hakim's structures will have fewer elements than Kuh's structure for identical transfer-function realization.

A new active *RC* configuration (Fig. 14) is obtained by connecting the input port of an *RC* two-port in parallel with the input terminals of the basic "pole" section. The pertinent network function of this structure is

$$t_I = \frac{I_o}{I_i} = \frac{-y_{21}}{y_{11} + Y_1 - \alpha Y_2} \tag{20}$$

Note that the form of the transfer function is identical to that of Kuh's configuration [Eq. (18)] and hence similar synthesis steps can be followed. This structure therefore has the same features and drawbacks as Kuh's structure.

In each of the previous three methods, the complex transmission zeros are realized by an *RC* two-port. This makes the overall synthesis difficult by requiring the realization of the *RC* two-port as an intermediate step. The problem can be eliminated if the basic "zero" section is used instead of the *RC* two-port in all the previous three structures. The resultant configurations are shown in Fig. 15 and their corresponding network functions are as follows:

$$t_v = \frac{V_o}{V_i} = \frac{Y_A - \mu Y_B}{Y_A + Y_B + Y_1 - \alpha Y_2} \qquad \text{for Fig. 15(A)} \tag{21}$$

$$t_I = \frac{I_o}{I_i} = \frac{Y_A - \mu Y_B}{Y_1 + Y_A - \alpha Y_2} \qquad \text{for Fig. 15(B)} \tag{22}$$

$$t_I = \frac{I_o}{I_i} = \frac{(1 + \alpha)(Y_A - \mu Y_B)}{Y_1 - \alpha Y_2} \qquad \text{for Fig. 15(C)} \tag{23}$$

Note that these transfer functions are similar in form to Yanagisawa's configuration.[53] Synthesis is achieved in a similar manner. The procedure may be illustrated by considering the realization of a transfer-current ratio $N(s)/D(s)$ in the form of Fig. 15(B). By making use of the $RC:-RC$ decomposition, the specified network function can be expressed as

$$\frac{N(s)}{D(s)} = \frac{N(s)/Q(s)}{D(s)/Q(s)} = \frac{Y_{RC}^{(1)} - Y_{RC}^{(2)}}{Y_{RC}^{(3)} - Y_{RC}^{(4)}} \tag{24}$$

where $Q(s)$ is a suitably chosen polynomial having only distinct negative real roots. In Eq. (24), $Y_{RC}^{(j)}$ is a passive *RC* driving-point admittance. Comparing Eqs. (22) and (24), one possible identification is

$$Y_A = Y_{RC}^{(1)} \qquad Y_B = \frac{1}{\mu} Y_{RC}^{(2)}$$

$$Y_1 = Y_{RC}^{(3)} \qquad Y_2 = \frac{1}{\alpha}[Y_{RC}^{(4)} + Y_{RC}^{(1)}] \tag{25}$$

It should be pointed out here that the configuration of Fig. 15(A) has also been independently advanced by Paul[63] and by Cooper and Harbourt.[64] The amplifiers in Fig. 15(A) and (B) can be made nonideal if necessary. With slight modification, the two amplifiers of Fig. 15(C) can be combined into a single-input differential-output amplifier. The resulting configuration is the one suggested by Bobrow.[65] A similar modification of Fig. 15(B) also can be carried out. The synthesis methods based on the active configurations of Fig. 15 do not impose any restrictions on the realizability of the transfer function other than that it be a real rational function; in addition, these methods allow independent adjustment of the poles and the zeros by simply controlling the gains of the amplifiers.

A completely new approach to the problem of transfer-voltage-ratio realization has been proposed by Hazony and Joseph.[66] They begin by considering a general active *RC* configuration employing an ideal finite-gain VCVS as shown in Fig. 16. The transfer-voltage ratio of this structure is given as

$$t_v = \frac{V_0}{V_1} = \frac{kt_{31}}{1 - kt_{32}} \tag{26}$$

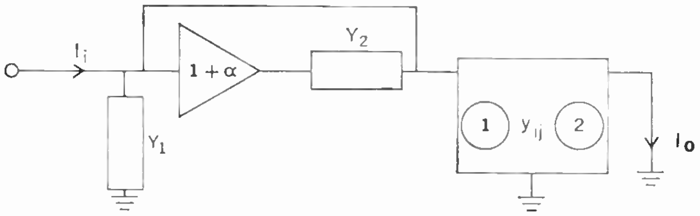
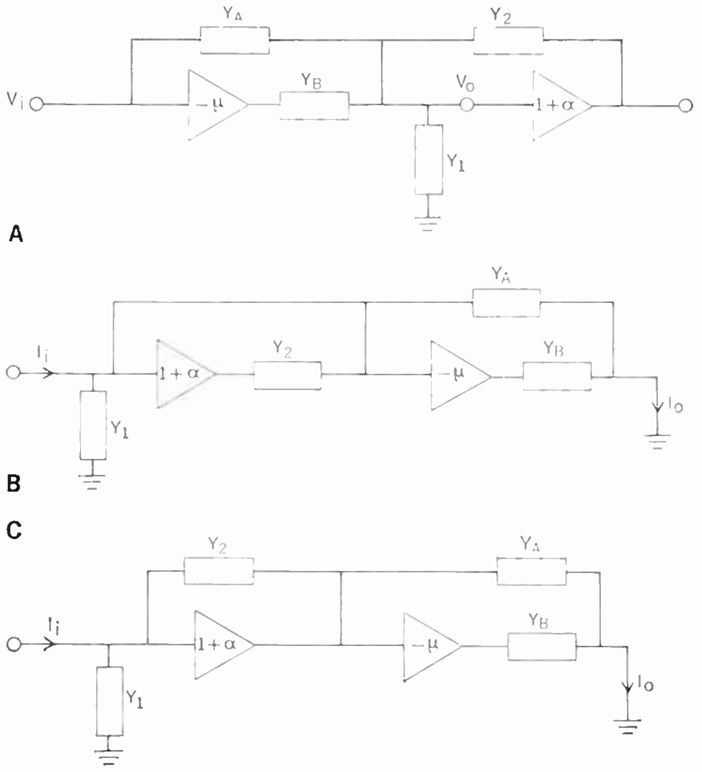


FIGURE 14. A new active RC configuration.

FIGURE 15. Active RC configurations employing two voltage amplifiers suitable for use in obtaining arbitrary transfer-voltage-ratio realizations.



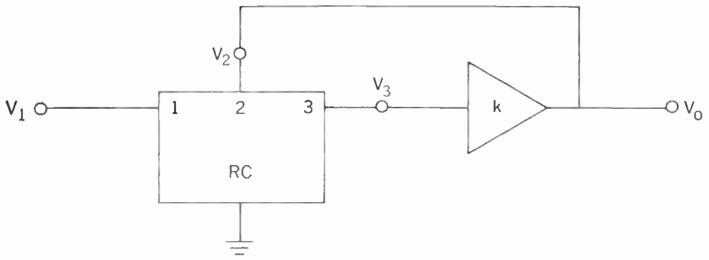
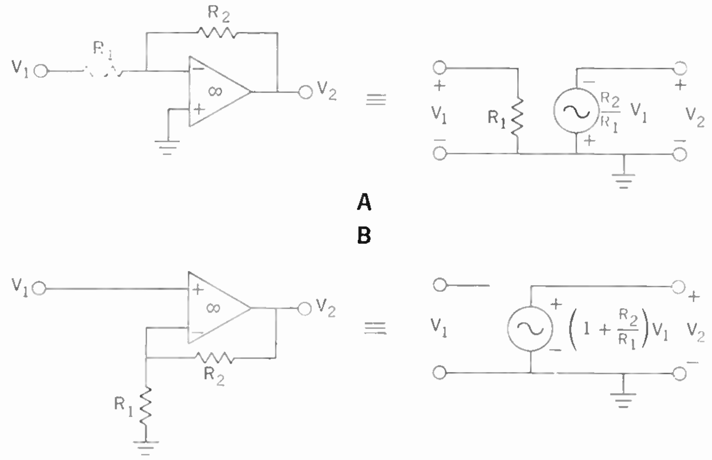


FIGURE 16. A general active RC configuration employing a single voltage amplifier.

FIGURE 17. Operational-amplifier realizations of finite-gain voltage amplifiers.



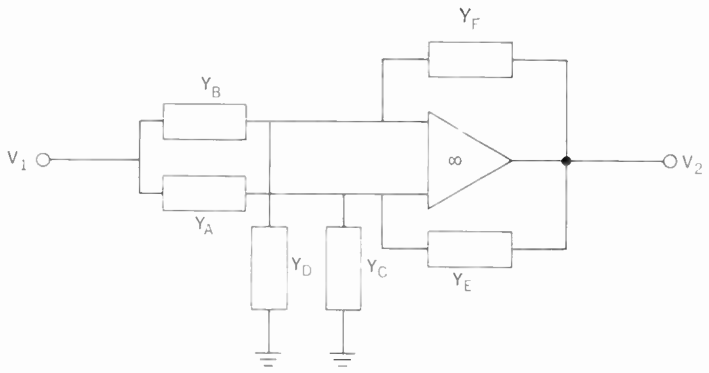
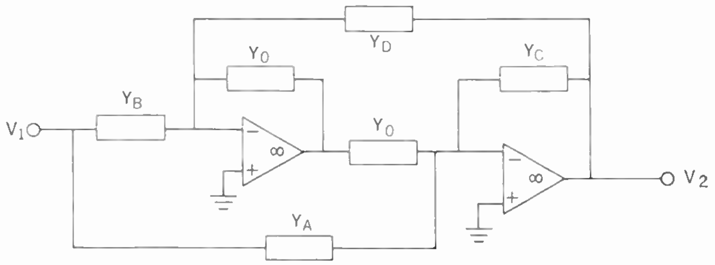


FIGURE 18. An RC operational-amplifier two-port suitable for arbitrary transfer-voltage-ratio realization.

FIGURE 19. RC operational-amplifier configuration developed by Lovering.



where

$$t_{31} = \left.\frac{V_3}{V_1}\right|_{V_2=0} \quad \text{and} \quad t_{32} = \left.\frac{V_3}{V_2}\right|_{V_1=0} \tag{27}$$

A specified voltage-transfer function $N(s)/D(s)$ can be put into the form of Eq. (26) by decomposing $D(s)$ as

$$D(s) = D_1(s) - D_2(s) \tag{28}$$

and then expressing

$$t_v = \frac{N(s)}{D(s)} = \frac{N(s)/D_1(s)}{1 - \dfrac{D_2(s)}{D_1(s)}} \tag{29}$$

Comparing Eqs. (26) and (29), we identify

$$kt_{31} = \frac{N(s)}{D_1(s)} \quad \text{and} \quad kt_{32} = \frac{D_2(s)}{D_1(s)} \tag{30}$$

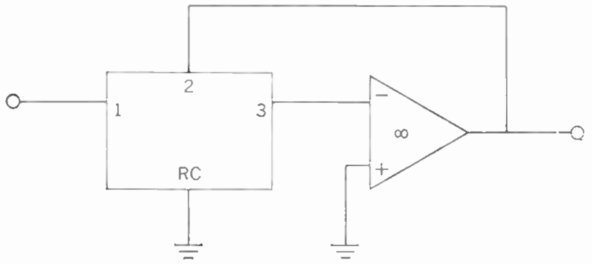
Note that if we short-circuit terminals 2 and 1 together, the voltage-transfer ratio is given by $(t_{31} + t_{32})$. Therefore, t_{31}, t_{32}, and $(t_{31} + t_{32})$ must each satisfy the Fialkow–Gerst conditions for an *RC* transformerless three-terminal network. This condition then determines the way in which the denominator polynomial $D(s)$ must be decomposed. Once a suitable decomposition has been found, the *RC* four-terminal network is realized, following the multiport synthesis technique of Fialkow and others.[67] It should be noted that this method is restricted to transfer functions having no positive-real-axis transmission zeros. Hazony also has considered the minimization of passive components in the realization.[68]

Another general, active *RC* synthesis procedure utilizing unity-gain voltage amplifiers is attributable to Hogin,[69] who has also considered optimization of the active filter using digital computers for a class of low-pass structures. Related work in this field is described in Refs. 70–80.

Synthesis using operational amplifiers. It should be noted that the finite-gain voltage amplifiers can be realized with the aid of operational amplifiers, as illustrated in Fig. 17. Thus, in a way, the methods previously discussed can also be considered as synthesis using operational amplifiers. Here we shall consider several additional methods.

A completely general method employing a single differential-input operational amplifier is available.[81,82] The method is based on the active-network configuration

FIGURE 20. A general active RC configuration employing a single-ended operational amplifier.



of Fig. 18, which is characterized by the following transfer-voltage ratio:

$$t_v = \frac{Y_A - Y_B}{Y_F - Y_E} \tag{31}$$

provided that

$$Y_B + Y_D + Y_F = Y_A + Y_C + Y_E \tag{32}$$

Realization of the RC one-ports follows standard procedures. Y_A, Y_B, Y_F, and Y_E are identified by expressing the specified transfer function in the form of Eq. (24). On this basis, Y_C and Y_D are obtained from the constraint equation (32). It is evident that any real rational transfer function can be realized this way. Note that the filter configuration allows cascading without additional buffer amplifiers.

A two-amplifier configuration, suggested by Lovering,[83] is particularly suitable for analog-computer simulation because of its use of single-ended operational amplifiers. Lovering's circuit, shown in Fig. 19, has a transfer-voltage ratio given by:

$$t_v = \frac{Y_B - Y_A}{Y_C - Y_D} \tag{33}$$

The synthesis method should be evident. Since Y_0 does not appear in Eq. (33), any suitable preselected nonzero value can be used. Incidentally, the synthesis method that is suggested in Ref. 76 is an extension of Lovering's method.

A general active RC synthesis method using a single-ended operational amplifier[84] is easily obtained by following an approach similar to the Hazony–Joseph method discussed earlier. The proposed configuration, which is indicated in Fig. 20, is characterized by the following transfer-voltage ratio:

$$t_v = -\frac{t_{31}}{t_{32}} \tag{34}$$

As can be seen from Eq. (34), this method is restricted to transfer functions having no poles and zeros on the positive real axis.

Related work will be found in Refs. 60 and 85–87.

Other methods. Several other synthesis techniques that have recently appeared do not fall directly within the previous two classes. A few of these methods will now be considered briefly.

Continued fraction expansion was used by Gorski-Popiel and Drew to prove that any real rational function can be realized in a ladder structure with negative RC impedances in the shunt branches.[88] Holt and Canning[89] advanced a technique based on an n-port transformerless realization scheme to realize a voltage-transfer ratio. Another synthesis method using an inverting-type voltage amplifier and a negative-impedance converter was developed by Antoniou,[90] whose method also leads to the two-amplifier configuration of Fig. 15(A).

The use of active elements characterized by frequency-dependent parameters was considered by Mitra and Howard,[22] who extended the RC gyrator cascade synthesis method to enable the use of gyrators having gyration admittances that are passive RC driving-point admittance functions.

Sensitivity considerations. Most of the synthesis methods discussed so far have a common drawback from a practical point of view—that is, high pole sensitivity with respect to the parameter of the active element. This arises because of the realization of complex poles by subtracting two polynomials. A simple proof of this well-known fact was provided by Bown,[91] who showed that in an active RC realization of a second-order transfer function employing a single positive-gain VCVS or a negative-impedance converter, the Q-sensitivity with respect to the active parameter k is

$$S_k^Q = \frac{dQ/Q}{dk/k} > 2Q - 1 \tag{35}$$

This implies that for a Q of 100, the variation of k must be kept within at least 0.025 percent to keep the Q variation within 5 percent.

Because of this situation, the majority of the active synthesis procedures discussed here are not satisfactory for transfer functions having Q-values greater than 10, except possibly when the unity-gain VCVS is used (due to the very high stability of this active element). To ensure low sensitivity, two avenues have been explored. In one approach, the problem of approximation has been reexamined in the light of sensitivity requirements. In the second approach, interest is focused on the use of multiloop feedback active networks. Briefly reviewed next is some recent work done in the former direction. The multiloop synthesis approach is considered later.

Lee[92] has investigated the relation between system sensitivity and the pole and zero locations for high-Q active networks. Specifically, he has shown that for a second-order transfer function having a pair of high-Q poles, the magnitude of the transfer-function sensitivity with respect to any passive or active parameter at the resonant frequency is directly proportional to the system Q and inversely proportional to the distance between the poles and the zeros. He thus asserts that the aim of the approximation should be to obtain a transfer function having poles with the smallest allowable system Q with the zeros located as far as possible from the poles.

Use of low-Q poles in the approximation procedure was also proposed by Vlach and Bendik.[93] They verified the low-sensitivity feature by constructing a fourth-degree and a sixth-degree active filter for pulse transmission by using Linvill's method.[3] Gorski-Popiel[94] suggests the use of a higher-order approximation, which increases the complexity of the active filter and thus may not be desirable in many applications.

An interesting contribution to the sensitivity aspects of RC-amplifier-type filters has been made by Saraga.[95] He specifically considers the sensitivity of a second-order RC voltage-amplifier circuit where the voltage amplifier has been built using an operational amplifier with an open-loop gain of μ. He has shown that the Q-sensitivity with respect to μ can be decreased from a value of $2Q^2/\mu$ to a value of $3Q/\mu$ by increasing the closed-loop gain from unity to 4/3. However, this decrease is associated with an increase of the Q-sensitivities with respect to the passive components. Saraga proposes the minimization of $\Delta Q/Q$ for maximum component variations as a criterion for optimum design.

Coefficient-matching approach. In the second approach, followed by many authors, the active-network configuration is selected a priori, and its network function determined by analysis. A specified network function is then compared with the computed network function and the element values determined by equating like coefficients. Usually the number of equations obtained by comparison is more than the number of unknowns. This gives the designer some freedom in selecting suitable element values according to various design criteria.

For convenience, the coefficient-matching approach is usually restricted to biquadratic transfer functions. The active-network configurations are of the types shown in Figs. 16 and 20. This allows the realization of higher-order functions by cascading second-order stages without additional buffer amplifiers.

The well-known Sallen and Key realization methods[96] and the analog-computer simulation methods[97] are classic examples of the coefficient-matching approach. Some of the recent methods based on this approach will now be reviewed.

The voltage amplifier as an active element. One significant result is attributable to Moschytz.[98] He considers the active realization of a second-order transfer-impedance function of the form:

$$Z_{21}^{A}(s) = K_A \frac{s^2 + \dfrac{\omega_p}{g_R} s + \omega_p^2}{s^2 + \dfrac{\omega_p}{g_p} s + \omega_p^2} \tag{36}$$

where $Z_{21}^{A}(s)$ has negative-real-axis zeros (that is, $g_R < 0.5$). The realization of an arbitrary transfer function

$$T(s) = K \frac{s^2 + \dfrac{\omega_z}{g_z} s + \omega_z^2}{s^2 + \dfrac{\omega_p}{g_p} s + \omega_p^2} \tag{37}$$

is obtained by cascading a passive *RC* two-port characterized by a transfer admittance:

$$Y_{21}^{R}(s) = K_R \frac{s^2 + \dfrac{\omega_z}{g_z} s + \omega_z^2}{s^2 + \dfrac{\omega_p}{g_R} s + \omega_p^2} \tag{38}$$

with the active two-port realizing $Z_{21}^{A}(s)$ given by Eq. (36).

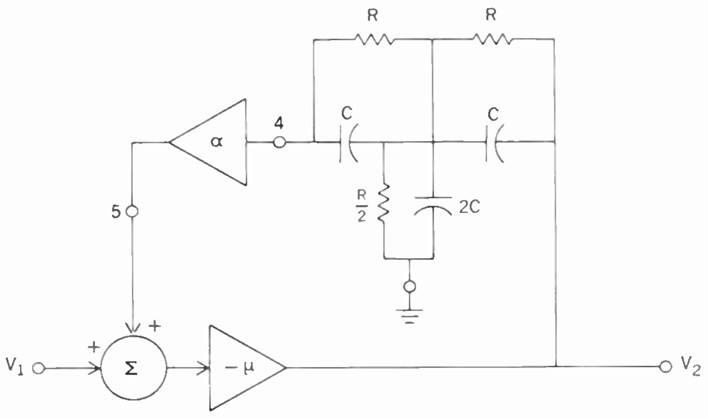


FIGURE 21. Moschytz's RC voltage-amplifier filter section (Type I) for medium-selectivity realization.

FIGURE 22. Moschytz's RC voltage-amplifier filter section (Type II) for high-selectivity realization.

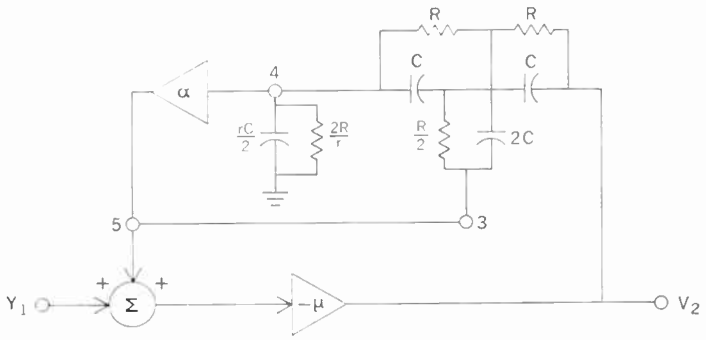
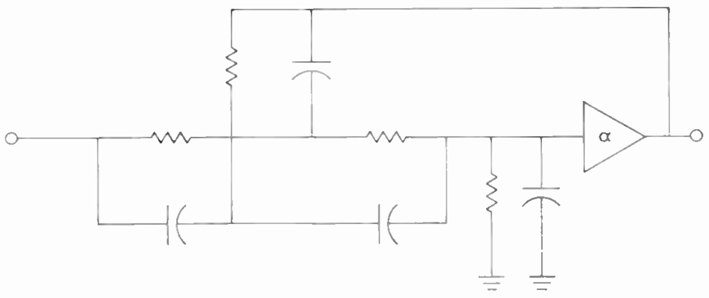


FIGURE 23. An RC voltage-amplifier filter section for realizing elliptic transfer functions.



Two different active *RC* configurations have been proposed for the realization of $Z_{21}^{A}(s)$. The choice is dictated by the value of g_p. For medium-selectivity transfer functions, the active-network configuration of Fig. 21 is used. For high-selectivity realization, the proposed active-network structure (Fig. 22) is obtained by connecting terminal 3 to terminal 5 (after disconnecting terminal 3 from ground) and by adding a resistor and a capacitor in parallel from terminal 4 to ground. In the second case, the actual transfer impedance becomes different from that specified by Eq. (36), thus introducing some error in the realization of $T(s)$. However, for many high-selectivity applications, this error can be ignored. One particularly attractive feature of Moschytz's approach is easily seen by comparing the structures of Figs. 21 and 22, which indicates the possibility of standardizing the structure for economic production in integrated form. An integrated bandpass filter of Q exceeding 300 has been built following this approach. Many details concerning the choice of components for optimum performance will be found in the original paper.[98] A number of possible realizations of $Y_{21}^{R}(s)$ have been catalogued for rapid design.[99]

The design of an *RC* amplifier section for the realization of the so-called "elliptic" transfer function

$$t_e(s) = K \frac{s^2 + \omega^2}{s^2 + as + b} \tag{39}$$

was considered by several authors. To this end, Piercey[100] advanced the active configuration of Fig. 23, along with design equations. He derived a table of sensitivity constants to enable the user to predict the change in response for a change in amplifier gain and in the values of several passive components. He also included a method for choosing the temperature coefficient of elements to keep the response temperature-invariant. The same configuration was independently proposed by Moschytz.[101]

Another design method for elliptic filter sections, based on Fig. 24, is given in Ref. 102, which includes a complete design example on the realization of a sixth-order transfer function having two pairs of $j\omega$-axis transmission zeros. A similar filter design was reported elsewhere by Kerwin.[103]

Second-order filter design, using Sallen and Key type filter sections and modified filter sections based on Fig. 24, was considered in great detail by Boyce.[104] An interesting second-order low-pass filter section employing an active feedback to allow independent adjustment of the bandwidth was advanced by Bloodworth and Nesbitt.[105]

The use of multiloop active feedback structures to improve sensitivity figures has also been considered by Kerwin.[106] He has modified two Sallen and Key configurations by using both a positive-gain VCVS and a negative-gain VCVS in each configuration and has shown that the resulting structures have a transfer-function sensitivity with respect to the gain of the noninverting-type VCVS that is considerably smaller than that which would have been obtained by using the noninverting-type VCVS alone.

Related efforts are described in Refs. 107 and 108.

The operational amplifier as an active element. One of the most significant contributions in which operational amplifiers were employed was advanced recently by Kerwin, Huelsman, and Newcomb.[109] Their proposed configuration, sketched in Fig. 25, employs three operational amplifiers and provides low-pass, bandpass, or high-pass responses depending on the choice of output terminals. For example, for a low-pass response, the output is taken from terminal 4, and the pertinent transfer function is then given as

$$\frac{V_4}{V_1} = \frac{R_2(R+1)}{R_2+1}\left[\frac{1}{s^2 + \dfrac{1+R}{1+R_2}s + R}\right] \qquad (40)$$

assuming that

$$R_1 = R_3 = R_8C_1 = R_9C_2 = 1 \qquad (41)$$

It can be seen that terminals 3 and 2 will provide, respectively, a bandpass and a high-pass response. Transfer functions having complex transmission zeros can be realized by adding V_2, V_3, and V_4 by means of an additional summing amplifier. Note from Eq. (40) that for a high-Q pole pair the resonant frequency is effectively equal to $\sqrt{R}$.

The most important feature of this new method is its extremely low sensitivity with respect to passive and active components. It can be shown that various sensitivities of interest are always equal to or less than unity, and are essentially independent of the Q of the pole pair. Since operational amplifiers are very-high-gain devices, the sensitivities with respect to the gains of the amplifiers are very low, provided the gains of the amplifiers are larger than the system Q.

Filter configurations similar to that of Fig. 25 and having low sensitivity figures have been proposed independently by Geffe[110] and Sutcliffe.[111]

Active-*RC*-filter design using a single-ended operational amplifier has been investigated by Holt and Sewell,[112] who have catalogued 11 networks for biquadratic realizations.

Multiloop active-network design using reactive amplifiers. The synthesis of arbitrary transfer functions by a special class of multiple-loop feedback configurations, containing amplifiers as the active elements, has been investigated recently by Biswas and Kuh.[113] Their approach allows the use of amplifiers, characterized by a frequency-dependent model. To illustrate, we consider here the design of a bandpass transfer function:

$$T(s) = \frac{2\alpha s}{s^2 + 2\beta s + \omega_0^2} \qquad (42)$$

The desired active-network configuration can be conveniently represented by a signal-flow graph shown in Fig. 26, in which $A_1(s)$ and $A_2(s)$ are the gains of the amplifiers. To obtain a realization of Eq. (42), the gain of the signal-flow graph is compared with this equation and the desired result is obtained by equating like coefficients. One such design is given by

$$\begin{aligned} a_{11} &= a_{22} = 1 - \beta \\ b_1 &= c_2 = 1 \\ b_2 &= c_1 = \alpha \\ a_{12} &= \alpha(\omega_0 - \beta) \\ a_{21} &= \frac{1}{\alpha}(-\omega_0 - \beta) \end{aligned} \qquad (43)$$

for

$$A_1(s) = A_2(s) = \frac{1}{s+1} \qquad (44)$$

FIGURE 24. Alternative RC voltage-amplifier filter sections for realizing elliptic transfer functions.

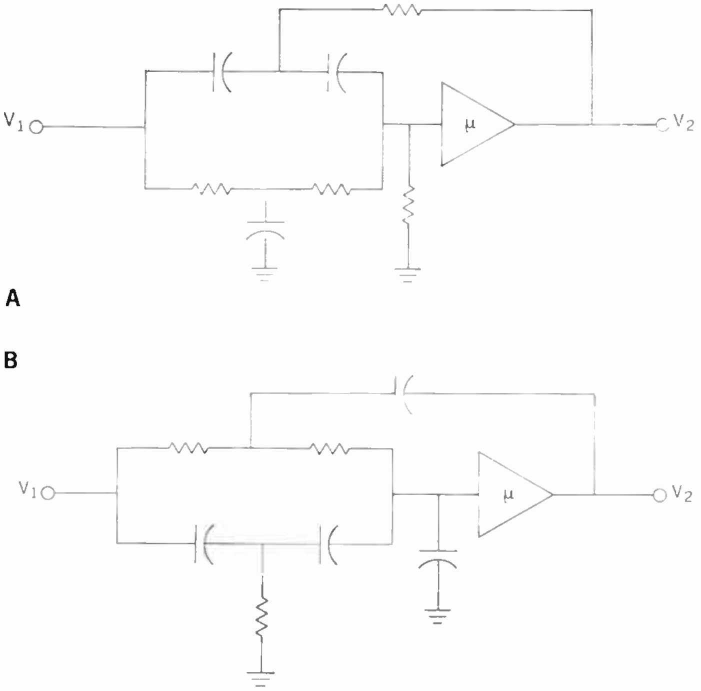


FIGURE 25. A minimum sensitive RC operational-amplifier filter configuration providing low-pass, high-pass, and bandpass response depending on the choice of output terminals.

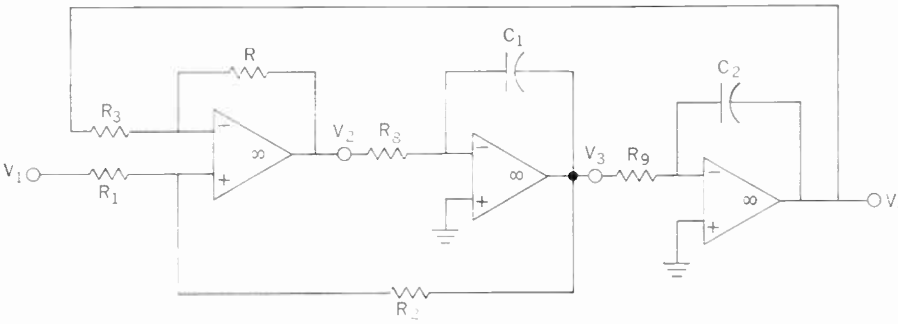


FIGURE 26. Signal-flow-graph representation of an active network configuration using two amplifiers.

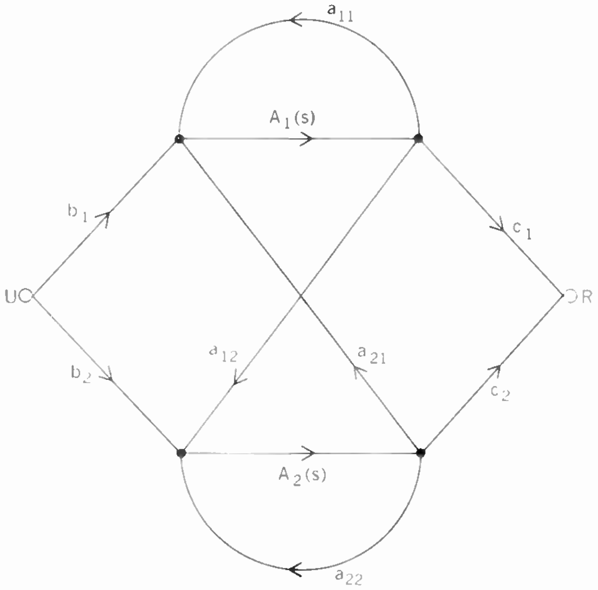
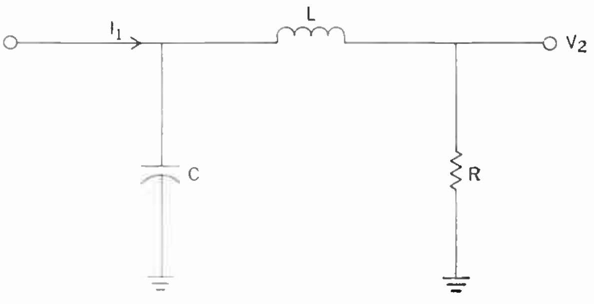


FIGURE 27. A second-order low-pass RLC filter section.



They have shown that the foregoing solution realizes a minimum-sensitivity active network and is thus the "optimum" solution. In fact, for second-order systems, the optimal design can be obtained analytically. For higher-order systems, however, a computer-based optimization must be followed.

Filter design employing simulated inductors. Recent interest in active filters using simulated inductors started with Orchard's observation on the inherent low-sensitivity feature of resistively terminated LC ladder filters.[111] To illustrate this, consider the second-order low-pass ladder of Fig. 27. Its transfer impedance is given as

$$Z_{21}(s) = \frac{V_2}{I_1} = \frac{R/LC}{s^2 + \frac{R}{L}s + \frac{1}{LC}} \tag{45}$$

The Q of the pole pair is

$$Q = \frac{1}{R}\sqrt{\frac{L}{C}} \tag{46}$$

from which may be obtained

$$S_R^Q = -1 \qquad S_L^Q = \tfrac{1}{2} \qquad S_C^Q = -\tfrac{1}{2} \tag{47}$$

Note that the Q-sensitivities are independent of the system Q, and are much smaller than Eq. (35) for moderate-Q to high-Q realizations. Gorski-Popiel[115] and Woodward and Newcomb[116] independently have quantitatively compared second-order, doubly loaded LC filters with an RC negative-impedance-converter filter, and have shown that LC filters are much superior from the point of view of sensitivity. A similar conclusion has been reached by Gensel.[117]

Thus, a simple way of designing an inductorless filter is to design a conventional LC filter first, and then replace the inductances by simulated inductors. In addition to the low-sensitivity feature, this approach offers the additional advantage of tunability simply by the use of trimmer capacitors. Relatively complex active filters built using simulated inductors have been reported in Refs. 38, 118, and 119. The experimental results are encouraging.

The outlook for the future

Attention now may be directed toward the future of active-filter networks. Some new and unanswered questions will be asked and discussed, in the hope that their solutions may be found in the years ahead.

An examination of past and present contributions reveals a welcome and growing trend toward the practical implementation of active filters, away from the original emphasis on academic problems. Practical active filters are now a reality, corroborated by the fact that there are presently several industrial concerns in the United States involved in their commercial production. Frequency-selective active filters are now employed in precision instruments, in satellites, in telephone and data communication systems, and in many other applications too numerous to mention. However, most of the presently available active filters are limited to low-Q response and to low-frequency applications.

The emphasis on practical implementation is always governed and directed by technological contributions and progress in the manufacture of components and systems. Integrated-circuit technology certainly will play a dominant role in active filters, with the consequence that the design of economic miniaturized, integrated, and hybrid active filters will become a major problem, with probable emphasis on high-Q bandpass networks.

Integrated monolithic construction offers many advantages that are particularly attractive to the designer of active filters. One of these is the reduction in cost, particularly if the filter is manufactured in large quantities. Unfortunately, the design of transmission networks until now has been more or less on a custom-job basis—with each network designed according to the specified signal transmission requirements.

One solution to this problem may be to design a limited number of standardized filter building blocks for a given frequency range. Then, for a specific application, certain combinations of some of them may be used to produce a filter having the desired characteristic. To minimize the total number of building blocks, two approaches can be followed: (1) certain characteristics of each block must be controllable by a single (perhaps external) element; and (2) for some applications, it may be necessary to "overdesign"—that is, to design a filter satisfying excessively stringent requirements.

Without doubt, the introduction of inexpensive integrated monolithic designs has enhanced the attractiveness of the operational amplifier as the active element. There is a strong indication that the commercial operational amplifier will continue to decline in cost, and eventually may be priced comparably with present-day quality transistors. (For example, the Fairchild μA 709, introduced in November 1965 at a price of $50, is now available for $5.)

An important problem, which has received scant attention, is that of "tuning" or adjusting the performance of a practical active filter. Because integrated components, particularly those that are passive, are available only with very wide tolerances, the performance of a practical active filter may be far from the desired one; thus, postdesign adjustment of circuit parameters becomes a necessity. However, in the case of integrated passive components, only the resistors may be trimmed economically and only to a limited extent. Again, to minimize production costs the number of adjustments should be held to a minimum. A solution of this tuning problem will be welcome.

To increase the yield and to minimize the number of adjustments, an integrated active filter must also have the lowest possible multiparameter sensitivity. Even though some recent contributions in this respect appear very encouraging, many practical problems remain unsolved. Future attempts to design optimum active filters must also take into account some inherent features of integrated circuits, such as closer matching of active and passive components through a wide temperature range and excellent thermal coupling.

Another research area that will receive more attention in coming years is directly related to monolithic production of linear circuits. To minimize the cost of a circuit, more circuits must be contained in a single wafer; thus, it is implied that the size of a circuit must be minimized. Basically, in an integrated circuit, a capacitor requires more area than a resistor, which, in turn, uses more space than a transistor. As a result, for the economic production of an integrated active filter, the circuit must be designed with a minimum number of passive components, particularly capacitors. In addition, an attempt should be made to minimize the total capacitance and resistance in the circuit. A partial solution to these problems was given recently by Horowitz and Branner[120] and by Horowitz.[121] An alternative approach, which minimizes the number of capacitors, is that of state-variable synthesis.[109,122]

With integrated-circuit technology it is much easier to manufacture composite blocks of interconnected passive and/or active components than to manufacture individual passive and active components. It therefore appears reasonable to develop synthesis techniques that employ composite circuit blocks as components. Incidentally, some contributions in this direction have already been made in passive-network synthesis, in which the circuit blocks are lossy inductors, quartz crystals, or microwave resonators. Some results in the design of active RC filters based on the use of RC resonators have been advanced by O'Neill and Ghausi.[78,79]

There are many high-frequency applications in which active integrated filters would be extremely useful. However, at high frequencies the phase shift introduced in the active elements creates serious problems that make it almost impossible to use direct synthesis methods found suitable for low-frequency applications. It is expected that, during the next few years, much effort will be devoted to extending the useful frequency range of linear active circuits.

The author wishes to thank Drs. E. S. Kuh, W. G. Howard, Jr., J. Watson, G. C. Temes, R. W. Newcomb, and H. J. Orchard for their help and advice in this work.

This article is based on a lecture presented at the 1968 Summer School on Circuit Theory, Prague, Czechoslovakia, June 28–July 12, 1968.

REFERENCES

1. Su, K. L., *Active Network Synthesis.* New York: McGraw-Hill, 1965.

2. *Handbook of Operational Amplifier Active RC Networks*, Burr-Brown Research Corp., Tucson, Ariz., 1966.

3. Linvill, J. G., "RC active filters," *Proc. IRE*, vol. 42, pp. 555–564, Mar. 1954.

4. Keen, A. W., "Active and nonreciprocal networks," *Acta Polytech. (Prague)*, vol. 14, no. 1, pp. 47–76, 1966.

5. Braun, J., "Analytical methods in active network theory," *Acta Polytech. (Prague)*, vol. 14, no. 1, pp. 5–28, 1966.

6. Field, R. K., "The tiny, exploding world of linear microcircuits," *Electron. Design*, vol. 15, pp. 49–66, 68–72, July 19, 1967.

7. Electro-Technology Staff Report, "Packaged operational amplifiers," *Electro-Technol.*, vol. 79, pp. 73–77, Jan. 1967.

8. Fowler, W. J., "A look at linear integrated circuits," *Electron. Ind.*, vol. 24, pp. 64–69, 194, Sept. 1965.

9. Leeds, M. B., "Linear IC's: Part 4—Inside the operational amplifier," *Electronics*, pp. 86–91, Oct. 16, 1967.

10. Giles, J. N., ed., *Linear Integrated Circuits Applications Handbook*, Fairchild Semiconductor, Mountain View, Calif., 1967.

11. *Linear Integrated Circuit Fundamentals*, Tech. Series IC-40, RCA, Harrison, N.J., 1966.

12. Keller, J. P., "Linear IC's: Part 3—Differential amplifiers at work," *Electronics*, pp. 96–105, Sept. 18, 1967.

13. Widlar, R. J., "A unique circuit design for a high performance operational amplifier especially suited to monolithic construction," *Proc. Nat'l Electron. Conf.*, vol. 21, pp. 169–174, 1965.

14. Jespers, P., "Etude theroique du taux de rejection du mode commun des amplificateurs differentiels," *Revue MBLE (Belgium)*, vol. 9, no. 2, 1966.

15. Stata, R., "Operational amplifiers, Part III—Survey of commercially available operational amplifiers," Applications Notes, Analog Devices, Inc., Cambridge, Mass.

16. *Applications Manual for Computing Amplifiers for Modelling, Measuring, Manipulating and Much Else*, 2nd ed., George A. Philbrick Researchers, Inc., Dedham, Mass., 1966.

17. *Integrated Operational Amplifiers*, 1st ed., Technical Information and Applications, Microelectronic Div., Radiation, Inc., Melbourne, Fla., 1966.

18. *Handbook of Operational Amplifier Applications*, Burr-Brown Research Corp., Tucson, Ariz.

19. Yanagisawa, T., and Kawashima, Y., "Active gyrator," *Electron. Letters*, vol. 3, pp. 105–107, Mar. 1967.

20. Shenoi, B. A., "Practical realization of a gyrator circuit and RC gyrator filters," *IEEE Trans. Circuit Theory*, vol. CT-12, pp. 374–380, Sept. 1965.

21. New, W., and Newcomb, R. W., "An integrable time-variable gyrator," *Proc. IEEE*, vol. 53, pp. 2161–2162, Dec. 1965.

22. Mitra, S. K., and Howard, W. G., Jr., "The reactive gyrator—a new concept and its application to active network synthesis," *1966 IEEE Internat'l Conv. Record*, vol. 14, pt. 7, pp. 319–326.

23. Holmes, W. H., Gruetzman, S., and Heinlein, H. E., "High performance direct-coupled gyrators," *Electron. Letters*, vol. 3, pp. 45–46, Feb. 1967.

24. Bach, R. E., Jr., and Carlson, A. W., "Practical realization of direct-coupled gyrator circuits," *Proc. Asilomar Conf. on Circuits and Systems*, Monterey, Calif., pp. 877–890, 1967.

25. Sheahan, D. F., and Orchard, H. J., "High-quality transistorized gyrator," *Electron. Letters*, vol. 2, pp. 274–275, 1966.

26. Sheahan, D. F., and Orchard, H. J., "Integrable gyrator using MOS and bipolar transistors," *Electron. Letters*, vol. 2, pp. 390–391, Oct. 1966.

27. Chua, H. T., and Newcomb, R. W., "Integrated direct-coupled gyrator," *Electron. Letters*, vol. 3, pp. 182–184, May 1967.

28. Rao, T. N., and Newcomb, R. W., "A direct coupled gyrator suitable for integrated circuits and time variation," *Electron. Letters*, vol. 2, pp. 250–251, July 1966.

29. Mitra, S. K., "Equivalent circuits of gyrators," *Electron. Letters*, vol. 3, p. 388, Aug. 1967.

30. Yanagisawa, T., "Realization of a lossless transistor gyrator," *Electron. Letters*, vol. 3, pp. 167–168, Apr. 1967.

31. Yanagisawa, T., "Direct-coupled transistor lossless gyrators," *Electron. Letters*, vol. 3, pp. 311–312, July 1967.

32. Tsirel'son, D. A., "Methods of making gyrators," *Telecommun. Radio Eng.*, pp. 12–15, Apr. 1966.

33. Vasudeva, M. S., "Gyrator realization," *Electron. Letters*, vol. 2, pp. 201–202, 1966.

34. Deboo, G. J., "Application of a gyrator type circuit to realize ungrounded inductors," *IEEE Trans. Circuit Theory*, vol. CT-14, pp. 101–102, Mar. 1967.

35. Butler, E., "The operational amplifier as a network element," Res. Rept. EERL-40, Cornell University, Ithaca, N.Y., June 1965.

36. Antoniou, A., "Gyrators using operational amplifiers," *Electron. Letters*, vol. 3, pp. 350–352, Aug. 1967.

37. Rao, T. N., Gary, P., and Newcomb, R. W., "Equivalent inductance and *Q* of a capacitor-loaded gyrator," *IEEE J. Solid-State Circuits*, vol. SC-2, pp. 32–33, Mar. 1967.

38. Sheahan, D. F., "Inductorless filters," Tech. Rept. 6560-15, SEL-67-086, Stanford Electronics Lab., Stanford University, Calif., Sept. 1967.

39. Gensel, J., "Notes on the structure of gyrator circuits," *Proc. IEEE*, vol. 55, pp. 1735–1736, Oct. 1967.

40. Bialko, M., "On Q factor and Q stability of an inductor simulated by a practical gyrator," *Electron. Letters*, vol. 3, pp. 168–169, Apr. 1967.

41. Gorski-Popiel, J., "RC-active synthesis using positive-immittance converters," *Electron. Letters*, vol. 3, pp. 381–382, Aug. 1967.

42. Riordan, R. S. H., "Simulated inductors using differential amplifiers," *Electron. Letters*, vol. 3, pp. 50–51, Feb. 1967.

43. Lampard, D. G., and Rigby, G. A., "The application of a positive immittance inverter to the design of integrated selective filters," *Proc. IEEE*, vol. 55, pp. 1101–1102, June 1967.

44. Su, K. L., "FET-circuit realization of the inductance," *Electron. Letters*, vol. 2, pp. 469–470, Dec. 1966.

45. Su, K. L., "A transistor-circuit realization of the inductance," *Proc. IEEE*, vol. 54, pp. 2025–2027, Dec. 1966.

46. Keen, A. W., and Peters, J. L., "Inductance simulation with a single differential-input operational amplifier," *Electron. Letters*, vol. 3, pp. 136–137, Apr. 1967.

47. Holt, A. G. J., and Taylor, J., "Method of replacing ungrounded inductors by grounded gyrators," *Electron. Letters*, vol. 1, p. 105, 1965.

48. Anderson, B. D., New, W., and Newcomb, R., "Proposed adjustable tuned circuits for microelectronic structures," *Proc. IEEE*, vol. 54, p. 411, Mar. 1966.

49. Sheahan, D. F., "Gyrator floating circuit," *Electron. Letters*, vol. 3, pp. 39–40, Jan. 1967.

50. Holmes, W. H., Gruetzman, S., and Heinlein, H. E., "Direct coupled gyrators with floating ports," *Electron. Letters*, vol. 3, pp. 46–47, Feb. 1967.

51. Keen, A. W., and Peters, J. L., "Nonreciprocal representation of the floating inductor with grounded amplifier realizations," *Electron. Letters*, vol. 3, pp. 369–371, Aug. 1967.

52. Horowitz, I. M., and Branner, G. R., "A unified survey of active RC synthesis techniques," *Proc. Nat'l Electron. Conf.*, vol. 23, 1967.

53. Yanagisawa, T., "RC active networks using current inversion type negative impedance converters," *IRE Trans. Circuit Theory*, vol. CT-4, pp. 140–144, Sept. 1957.

54. Kinariwala, B. K., "Synthesis of active RC networks," *Bell System Tech. J.*, vol. 38, pp. 1269–1316, Sept. 1959.

55. Sandberg, I. W., "Synthesis of driving-point impedance with active RC networks," *Bell System Tech. J.*, vol. 39, pp. 947–962, July 1960.

56. Kuh, E. S., "Transfer function synthesis of active RC networks," *1960 IRE Internat'l Conv. Record*, vol. 8, pt. 2, pp. 134–138.

57. Horowitz, I. M., "Synthesis of active RC transfer functions," MRI Res. Rept. R-507-56, PIB-437, Polytechnic Institute of Brooklyn, N.Y., Nov. 1956.

58. Sipress, J. M., "Synthesis of active RC networks," *IRE Trans. Circuit Theory*, vol. CT-8, pp. 260–269, Sept. 1961.

59. Mitra, S. K., "Transfer function realization using RC one-ports and two grounded voltage amplifiers," *Proc. 1st Ann. Princeton Conf. on Information Sciences and Systems*, Princeton University, N.J., pp. 18–23, Mar. 1967.

60. McVey, P. J., "Synthesis of transfer function by RC networks with two or three computing amplifiers," *Internat'l J. Control*, vol. 2, pp. 125–133, Aug. 1965.

61. Bogdanov, M., Boite, R., and Leich, H., "Synthèse des filtres actifs a caracteristique elliptique," *Rev. Haut. Freq. (Belgium)*, vol. 6, pp. 245–258, 1965.

62. Hakim, S. S., "RC active filters using an amplifier as the active element," *Proc. IEE (London)*, vol. 112, pp. 901–912, May 1965.

63. Paul, R. J., "Active network synthesis using one-port RC networks," *Proc. IEE (London)*, vol. 113, pp. 83–88, Jan. 1966.

64. Cooper, R. E., and Harbourt, C. O., "Sensitivity reduction in amplifier-RC synthesis," *Proc. IEEE*, vol. 54, pp. 1578–1579, Nov. 1966.

65. Bobrow, L. S., "On active RC synthesis using an operational amplifier," *Proc. IEEE*, vol. 53, pp. 1648–1649, Oct. 1965.

66. Hazony, D., and Joseph, R. D., "Transfer matrix synthesis with active RC networks," *SIAM J. Appl. Math.*, vol. 14, pp. 739–761, July 1966.

67. Fialkow, A. D., Hazony, D., Kodali, P., and Zeren, T., "Transfer-function realizability of multiport *RLC* transformerless grounded networks," *IEEE Trans. Circuit Theory*, vol. CT-11, pp. 73–79, Mar. 1964.

68. Hazony, D., "Grounded RC–unity gain amplifier transfer vector synthesis," *IEEE Trans. Circuit Theory*, vol. CT-14, pp. 75–76, Mar. 1967.

69. Hogin, J. L., "Active RC networks utilizing the voltage follower," Ph.D. dissertation, Montana State University, Bozeman, Mont., Mar. 1966.

70. Holt, A. G. J., and Linggard, R., "RC active synthesis procedure for polynomial filters," *Proc. IEE (London)*, vol. 113, pp. 777–782, May 1966.

71. Balabanian, N., "Active network synthesis with a non-ideal device," *J. Franklin Inst.*, vol. 279, pp. 334–346, May 1965.

72. Hakim, S. S., "A synthesis procedure for RC-active filters using unity-gain current and voltage amplifiers," *Internat'l J. Control*, vol. 3, pp. 553–564, 1966.

73. Zai, Y. F., "RC-active filters using unity-gain amplifiers," *Electron. Letters*, vol. 3, p. 461, Oct. 1967.

74. Horowitz, I. M., "Active RC synthesis," *IEEE Trans. Circuit Theory*, vol. CT-13, pp. 101–102, Mar. 1966.

75. Joseph, R. D., and Hilberman, D., "Immittance matrix synthesis with active networks," *IEEE Trans. Circuit Theory*, vol. CT-13, p. 324, Sept. 1966.

76. Richards, P. I., "Synthesizing transfer functions with two grounded pentodes," *Proc. IEEE*, vol. 55, pp. 552–553, Apr. 1967.

77. Rajasekharan, P. K., and Rao, S. N., "Synthesis of active RC networks using a differential amplifier," *Internat'l J. Electron.*, vol. 22, pp. 297–305, Apr. 1967.

78. O'Neill, J. F., and Ghausi, M. S., "Transfer function synthesis using resonators and some active RC resonators," *Proc. 1st Ann. Princeton Conf. on Information Sciences and Systems*, pp. 150–154, 1967.

79. O'Neill, J. F., and Ghausi, M. S., "Approximation in gain and phase of two-ports using lattice of paralleled resonators," *Proc. Hawaii Internat'l Conf. on System Sciences*, 1968.

80. Hazony, D., and Lagerlöf, R., "Canonical active RC networks," Tech. Rept., Chalmers Institute of Technology, Gothenburg, Sweden, Mar. 1967.

81. Mitra, S. K., "Active RC filters employing a single operational amplifier as the active element," *Proc. Hawaii Internat'l Conf. on System Sciences*, pp. 433–436, 1968.

82. Brugler, J. S., "RC synthesis with differential-input operational amplifiers," in Tech. Rept. 6560-4, SEL 66-058, Stanford Electronics Lab., Calif., June 1966.

83. Lovering, W. F., "Analog computer simulation of transfer functions," *Proc. IEEE*, vol. 53, pp. 306–307, Mar. 1965.

84. Mitra, S. K., "Transfer function synthesis using a single operational amplifier," *Electron. Letters*, vol. 3, pp. 333–334, July 1967.

85. Hakim, S. S., "RC-active synthesis using an operational amplifier," *Internat'l J. Control*, vol. 1, pp. 433–445, 1965.

86. Pande, H. C., and Shukla, R. S., "Synthesis of the transfer functions using an operational amplifier," *Proc. IEE (London)*, vol. 112, pp. 2208–2212, Dec. 1965.

87. Deliyanis, T., "Sensitivity study of five RC active networks using the method of single inversion," *Internat'l J. Electron.*, vol. 22, pp. 197–213, Mar. 1967.

88. Gorski-Popiel, J., and Drew, A. J., "RC-active ladder networks," *Proc. IEE (London)*, vol. 112, pp. 2213–2219, Dec. 1965.

89. Holt, A. G. J., and Canning, M. L., "Synthesis of an active 2-port network using *n*-port synthesis," *Proc. IEE (London)*, vol. 112, pp. 1689–1694, Sept. 1965.

90. Antoniou, A., "New RC-active network synthesis procedures using negative-impedance converter," *Proc. IEE (London)*, vol. 114, pp. 894–902, July 1967.

91. Bown, G. C., "Sensitivity in active RC filters," *Electron. Letters*, vol. 3, pp. 298–299, July 1967.

92. Lee, S. C., "Sensitivity minimization in active RC integrated-circuit design," *Proc. 4th Allerton Conf. Circuit and System Theory*, pp. 269–281, 1966.

93. Vlach, J., and Bendik, J., "Active filters with low sensitivity to element changes," *Radio Electron. Engr.*, vol. 33, pp. 305–316, May 1967.

94. Gorski-Popiel, J., "Reduction of network sensitivity through the use of higher-order approximating functions," *Electron. Letters*, vol. 3, pp. 365–366, Aug. 1967.

95. Saraga, W., "Sensitivity of 2nd-order Sallen–Key active RC filters," *Electron. Letters*, vol. 3, pp. 442–443, Oct. 1967.

96. Sallen, R. P., and Key, E. L., "A practical method of designing RC active filters," *IRE Trans. Circuit Theory*, vol. CT-2, pp. 74–85, Mar. 1955.

97. Bradley, F. R., and McCoy, R., "Driftless dc amplifiers," *Electronics*, pp. 144–148, Apr. 1952.

98. Moschytz, G. S., "Miniaturized filter building blocks using frequency emphasizing networks," *Proc. Nat'l Electron. Conf.*, pp. 364–369, 1967.

99. Moschytz, G. S., "Active RC filter building blocks using frequency emphasizing networks," *IEEE J. Solid-State Circuits*, vol. SC-2, pp. 59–62, June 1967.

100. Piercey, R. N. G., "Synthesis of active RC filter networks," *ATE J.*, vol. 21, pp. 61–75, Apr. 1965.

101. Moschytz, G. S., "Sallen and Key filter networks with amplifier gain larger than or equal to unity," *IEEE J. Solid-State Circuits*, vol. SC-2, pp. 114–116, Sept. 1967.

102. Kerwin, W. J., and Huelsman, L. P., "The design of high-performance active RC bandpass filters," *1966 IEEE Internat'l Conv. Record*, vol. 14, pt. 10, pp. 74–80.

103. Kerwin, W. J., "An active RC elliptic function filter," *1966 IEEE Region VI Conf. Record*, pp. 647–654, Apr. 1966.

104. Boyce, A. H., "A theoretical and practical study of active filters," *Marconi Rev.*, vol. 30, pp. 68–97, 2nd qtr., 1967.

105. Bloodworth, G. G., and Nesbitt, N. R. S., "Active filters operating below 1 Hz," *IEEE Trans. Instrumentation and Measurement*, vol. IM-16, pp. 115–120, June 1967.

106. Kerwin, W. J., "RC active networks," in Tech. Rept. 6560-4, SEL-66-058, Stanford Electronics Lab., Calif., June 1966.

107. Foster, E. J., "Active low-pass filter design," *IEEE Trans. Audio*, vol. AU-13, pp. 104–111, Sept.–Oct. 1965.

108. Blackburn, H., Campbell, D. S., and Muir, A. J., "Low-pass active filters," *Radio Electron. Engr.*, vol. 34, pp. 90–96, Aug. 1967.

109. Kerwin, W. J., Huelsman, L. P., and Newcomb, R. W., "State-variable synthesis for insensitive integrated circuit transfer functions," *IEEE J. Solid-State Circuits*, vol. SC-2, pp. 87–92, Sept. 1967.

110. Geffe, P., "Make a filter out of an oscillator," *Electron. Design*, pp. 56–58, May 10, 1967.

111. Sutcliffe, H., "Tunable filter for low frequencies using operational amplifiers," *Electron. Eng.*, pp. 399–403, June 1964.

112. Holt, A. G. J., and Sewell, J. I., "Active RC filters employing a single operational amplifier to obtain a bi-quadratic response," *Proc. IEE (London)*, vol. 112, pp. 2227–2234, Dec. 1965.

113. Biswas, R. N., and Kuh, E. S., "Optimum synthesis of a class of multiloop feedback systems," Memo. ERL-M218, Electronics Research Lab., University of California, Berkeley, July 11, 1967.

114. Orchard, H. J., "Inductorless filters," *Electron. Letters*, vol. 2, p. 224, Sept. 1966.

115. Gorski-Popiel, J., "Horowitz minimum sensitivity decomposition," *Electron. Letters*, vol. 2, pp. 334–335, Sept. 1966.

116. Woodward, J., and Newcomb, R. W., "Sensitivity improvement of inductorless filters," *Electron. Letters*, vol. 2, pp. 349–350, Sept. 1966.

117. Gensel, J., "Ein quantitativer Vergleich zwischen Negativ-Immittanz-Konverter- und Gyratorschaltungen," *Electron. Letters*, vol. 2, pp. 361–362, Oct. 1966.

118. Sheahan, D. F., and Orchard, H. J., "Bandpass filter realization using gyrators," *Electron. Letters*, vol. 3, pp. 40–42, Jan. 1967.

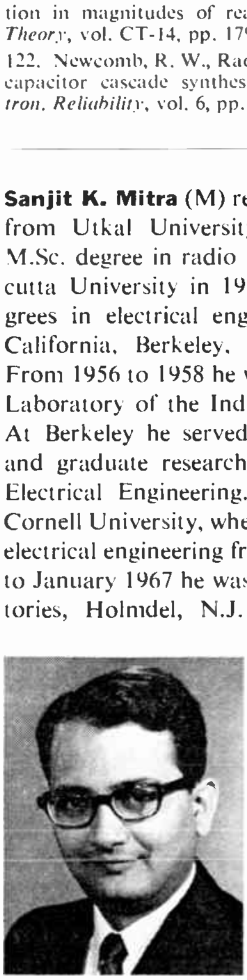
119. Holmes, W. H., Gruetzmann, S., and Heinlein, W. E., "High quality active C filters using gyrators," *Internat'l Solid-State Circuit Conf. Dig.*, 1967.

120. Horowitz, I. M., and Branner, G. R., "Active RC synthesis technique for significant reactance reduction," *Conf. Record 10th Midwest Symp. Circuit Theory*, 1967.

121. Horowitz, I. M., "Active RC synthesis for significant reduction in magnitudes of reactive elements," *IEEE Trans. Circuit Theory*, vol. CT-14, pp. 179–187, June 1967.

122. Newcomb, R. W., Rao, T. N., and Woodward, J., "A minimal capacitor cascade synthesis for integrated circuits," *Microelectron. Reliability*, vol. 6, pp. 113–124, 1967.

Sanjit K. Mitra (M) received the B.Sc. degree in physics from Utkal University, Cuttack, India, in 1953, the M.Sc. degree in radio physics and electronics from Calcutta University in 1956, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1960 and 1962, respectively. From 1956 to 1958 he was with the Electronic Computer Laboratory of the Indian Statistical Institute, Calcutta. At Berkeley he served as teaching assistant, associate, and graduate research engineer in the Department of Electrical Engineering. He then joined the faculty of Cornell University, where he was an assistant professor of electrical engineering from 1962 to 1965. From June 1965 to January 1967 he was with the Bell Telephone Laboratories, Holmdel, N.J. At present he is an associate professor of electrical engineering at the University of California, Davis. He is also a consultant to the Ampex Corporation. He has written a number of papers on active and passive networks and a forthcoming book, *Analysis and Synthesis of Linear Active Networks* (New York, Wiley). He is a member of Sigma Xi and Eta Kappa Nu.



Stimulating change in engineering education

The increasing involvement of the engineer with society, and the desire of students in general to take a more active role in operating their universities, are two of the major reasons calling for a basic change in engineering education

J. O. Kopplin *The University of Texas at El Paso*

It is generally agreed that major changes are necessary in engineering education, mainly for three reasons. First, the role of the engineer is changing markedly. Second, advances in technology are altering educational methods. Third, today's students are seeking active roles in the operation of the university. What is needed is a program incorporating experimentation, flexibility, and feedback, such as an honors program of the type now functioning at the University of Illinois. The Illinois program, which is described in this article, by definition exists for honor students. However, it has proved to be a stimulus for change in the entire undergraduate program of the college.

A continuous upgrading of the scientific and technical content of engineering education has taken place during the past two decades. Recent increased interest in engineering education,[1-4] however, stems not only from recognition of the rapid technological changes but also from recognition that the role of the engineer with respect to society is changing. This latter change is presenting the greater challenge to engineering colleges in that it requires corresponding change in the very nature and goals of engineering education—fundamental change that will require careful exploration.

Various authors have pointed out the need for revisions; some have indicated the type of change that should take place in undergraduate engineering education. However, this awareness in itself will not necessarily bring about change. Two reasons for the difficulty are that (1) for economic reasons, engineering schools tend to be rather large and, therefore, develop considerable inertia; and (2) because engineering schools are professional schools, engineering faculty establish their professional reputation through the practice of the profession. The latter is usually accomplished either by consulting or research activities, both of which tend to center the faculty member's interest on graduate education. It must be acknowledged that professional engineering education is steadily moving from the undergraduate to the graduate level. There is also the fact that change is sought only by the minority of any group, including educators.

It must be recognized that many of the advocated changes have been stated in terms of the desired results rather than in terms of methods for achieving the results. There is much more agreement concerning the capabilities engineers should have than there is concerning the educational processes that will produce engineers with the desired capabilities. There are, of course, many processes that will produce equal results. Because of the breadth of the engineering profession today, engineers play various roles throughout all society. Many engineering students are looking forward to careers in law, business, medicine, and politics.

Professor Shepherd[2] has stated, "One cannot speak of the education of engineers as a single track along which all must follow." This diversity, together with the real need to liberalize engineering education better to prepare engineers to compete and to participate fully in the highly educated world of today, is pushing undergraduate engineering education into a preprofessional role. Fundamental change is needed. To effect this degree of change, engineering schools need a pilot-plant operation, a program to explore new approaches and to promote continuous evolutionary development in engineering education to match the dynamism of the profession. What engineering education needs is flexibility, experimentation, and feedback,[3] because the solution to its problems lies not in "curriculum tinkering" but in the development of wholly new approaches—the development of a pluralism of educational patterns.

There are three major reasons for revision. The first reason is that the role of the engineer is changing markedly. Second, advances in technology, principally in communication technology and computers, are altering educational methods and procedures in all fields. Third, today's students are seeking active roles in the operation of the university—particularly with regard to curricula and student life. They expect their education to involve them in the search for solutions to today's problems.

The role of the engineer

The role of the engineer is changing because of the increasing importance of the interface of science and society—the interface at which the engineer finds his employment. The engineer today must be concerned not only with the development of practical technology but also with the application of that technology. This latter role involves the engineer with human problems in a manner that is not traditional. Today, society is not only looking for technologists—people who can apply scientific principles to the development of new devices and systems, engineers who can work with "things"—but also for engineers who can innovate in the sociotechnological arena, engineers who can work with people. Such engineers will need to be familiar with economics, human organization, and politics. If an engineer today designs and builds a factory that discharges waste into the atmosphere or a passing waterway, he is not acting as a professional. A professional attempts to consider all the consequences of his work; and because of the tremendous ability the engineer has today to change the surface of the earth and the processes that take place there, including those of human interaction, he must be extremely sensitive to the world in which he works.

Because of the engineer's new role with respect to society, there is a real need to liberalize engineering education.[5,6] An engineer without a liberal education is a mere technician unaware of the cultural and social forces of his age. All definitions of engineering include the phrase "for the benefit of mankind." If the needs of mankind change, then the education of engineers should be revised. Little is needed in the way of new technology to provide an abundance of food and shelter, to restore purity to our air and water, or to make our cities enjoyable places in which to live. If engineers are to serve mankind, then engineering schools should have the courage to analyze closely the needs of mankind today, and upon that analysis base their educational program. Engineers need empathy. They need to be humanized, not simply to "round them out" but as part of their professional training.

Educational technology

Advances in communication technology and computers are altering educational methods and procedures in all areas. This technology was created by engineers, and engineering schools should be providing leadership in its application to the educational process. The need to bring a group of students together at a specific place and time in order to lecture to them is vanishing. The unidirectional transfer of information accomplished by a lecture can be made available at an almost infinite number of locations and at any desired time through the application of technology. Most classroom demonstrations can be more intimately viewed through the utilization of television. The operation of many instruments and demonstrations can be programmed, allowing self-instruction and self-demonstration exercises. Computer-based teaching machines can often provide more individual attention and better utilized instruction. It appears likely that in the future teachers and students will meet in person only in the laboratory, around the conference table, and over coffee.

A complete restructuring of the practice of engineering and wholly new approaches to executing the engineering process are taking place because of computers; consequently, basic changes are needed in engineering education. Students, however, are not computers and their education in no way should resemble the programming of a computer. Excessive curriculum tinkering takes on aspects of computer programming—an attempt to "program the student" with a specific amount of data in a specified time. It is not simply that engineers should be trained to use a computer but that they need to be trained to work with a computer. Because of the availability of computers, increased emphasis should be given to educating engineers to make evaluations and judgments based on incomplete or nonquantitative information, and to judge the relevance of data and the appropriateness of a solution.

The concept—flexibility, experimentation, feedback

Undergraduate engineering students, in common with other students, are demanding that their education be relevant to today's world. They are seeking personal involvement in the major problems of society. Students fully realize that the development of improved refrigerator and vacuum-cleaner motors is not one of the major needs of mankind. They will accept such problems as part of their professional training but will not be content with spending four years involved with trivia. Students are keenly aware of the current needs of society and expect their education to consider those needs. Engineering schools cannot allow their students to sidestep human factors in the development of technology.

The dynamism of the engineering profession demands a constant search for new concepts and methods. Engineering education must also seek new techniques.

It is apparent that an experimental approach to new engineering education concepts is needed. In order to have experimentation, flexibility is required; and in order that the results of the experimentation be evaluated, feedback is necessary. All this requires a program—a pilot operation. A program that offers flexibility, experimentation, and feedback—the only solution to the dilemma of engineering education, as stated by Harvey Brooks[3]—can be an honors program. Excellent honors programs are developing across the United States in all areas of education. The Honors Program in Engineering at the University of Illinois, called the James Scholars Program in Engineering, was created to challenge the very top students of the college. The program exists for honor students; however, it has proved to be a stimulus for change in the entire undergraduate program.

The Illinois program

The first objective of the Honors Program is to fully challenge each student individually. This requires flexibility. It was determined early in the program that even the select group of honors students could not be fully challenged en masse; they can be fully challenged only as individuals. Each honors student is assigned to a special honors advisor as soon as he enters the program and he remains with that advisor until he is graduated. Because the program operates with the very minimum number of rules and regulations, all major decisions concerning the student's plan of study and his participation in the program rest with the advisor. Using the published curricula as a guide, each student is expected to develop his own plan of study. Such an individualized plan is created through the use of the substitution privilege. It is not necessary that an individual course be substituted for another course, only that the total number of hours substituted be equal to the total number of hours replaced in the curriculum. The criterion for judging the appropriateness of a substitution is simply that the substitution result in a more suitable and stronger program for the student. The decision is made by the honors advisor.

The second objective of the program is to provide varied academic opportunities. This requires experimentation. In doing this, the program offers several engineering honors courses each semester. The program, however, avoids establishing any permanent courses and, therefore, each course is set up primarily on a once-only basis. The majority of these courses are taught under a variable credit number with a blanket description of "Honors Seminar" or "Honors Project." Courses have been offered in space vehicle design, science of engineering materials, ocean engineering, dynamic systems, and systems design. Other courses have published titles such as "The Engineer and Society" and "The Engineer and His Profession." These courses have no formal outline; each semester they vary widely in content and approach and are always the creation of the instructor. With essentially a new and different program each semester, continuous experimentation with regard to course content and classroom approach is achieved. The lecture approach rarely is used. Many courses use a seminar approach with the class meeting around a conference table; some employ a continuous flow of outside experts for dialogue; for others, all class meetings are scheduled and are conducted by the students themselves.

The program has encouraged students to engage in independent and project-type studies. Many students have received academic credit for the study of subjects not considered by formal courses, and for study and work in research laboratories.

Various activities of the program have brought about the liberalization and humanizing of the student's education. This has been achieved through special activities and student involvement. A recent two-day conference, held at the university's off-campus center where meals and lodging are provided, considered "Engineering and Urban Development." Following a discussion entitled "Attitudinal Incompatibility and Survival in the Urban Complex," the speaker was engaged in informal discussion of inner-city enclaves until almost dawn!

Students in a year-long systems design course tackled the problems of developing programmed instruction for the blind. They became deeply involved both with programmed instruction as an educational method and with the very real problems of the blind student. They designed, adapted, and tested both computer-based and non-computer-based teaching devices and became enmeshed not only in difficult technical problems but also in sticky educational and human problems. The "human factors" were the most challenging and frustrating and could not be set aside.

Feedback is provided through the close advisor–advisee relationships established by the program and through the instructor–student contacts established in the small classes and seminars employed throughout the program. Feedback from the engineering profession itself is provided through the many outside contacts. The various activities of the program are continually subjected to evaluation analysis. Several of the advisors and some of the instructors have invited their advisees and students to their homes. This is sometimes a bit of a shock for the students as they have generally not had the opportunity to appreciate that their professors are human! They learn that their professors have families, that they are concerned with family matters, that they are interested in the affairs of their community, attend lectures, concerts, and plays, play musical instruments, and collect stamps. Of course, information passes both ways, and this enables the advisor to truly serve as the student's mentor.

A pilot operation is a proved engineering technique. Although it is true that many of the activities of the Honors Program are successful simply because of the relatively small number of students involved, it is also true that the program has stimulated change in undergraduate programs throughout the college. Increased flexibility has been incorporated into several curricula. Consideration is being given to stating certain degree requirements in terms of hours of study in various areas rather than in terms of specific courses. Petition committees are accepting more petitions from regular students as a means of achieving certain desired results rather than accepting petitions only as a means of granting exceptions for specific extenuating circumstances. Increased importance is being given to the advisor–advisee relationship. A new three-course circuits sequence, first explored in the Honors Program, has been adopted for all electrical engineering students. A sequence of laboratory courses designed to teach the theory and practice of experimentation has stimulated changes in several regular laboratory courses. This sequence was developed to replace all the traditional subject matter laboratories (circuits, electronics, energy conversion, etc.) in the curriculum and is now being explored as an interdisciplinary laboratory program. The Honors Program has also stimulated renewed interest and consideration of the nontechnical content of all curricula.

Conclusions

Fundamental change in engineering education is needed. This will come about through the application of flexibility, experimentation, and feedback, which can be provided by an honors program. Such change will enable engineering schools to approach potential students not simply with four years of rigorous training, but with an exciting opportunity to become involved with the world. Such programs will also produce a few "super-engineers" who will have the sensitivity, the competence, and the political acumen to change the world!

REFERENCES

1. Johnson, H. W., "Educating for the new technology," *IEEE Spectrum*, vol. 4, pp. 99–101, Aug. 1967.
2. Shepherd, W. G., "Education for the engineering mission," *IEEE Spectrum*, vol. 5, pp. 94–100, Feb. 1968.
3. Brooks, Harvey, "Dilemmas of engineering education," *IEEE Spectrum*, vol. 4, pp. 89–91, Feb. 1967.
4. Kincaid, J. F., "Engineering education for an age of change," *J. Eng. Educ.*, vol. 58, pp. 505–508, Feb. 1968.
5. Florman, S. C., "Liberal learning and the engineer," *J. Eng. Educ.*, vol. 57, pp. 629–630, May 1967.
6. Turner, E. M., "The human side of engineering," *IEEE Spectrum*, vol. 4, pp. 70–71, July 1967.

J. O. Kopplin (M) was chairman of the Honors Program in the College of Engineering at the University of Illinois for three years beginning in the fall of 1965; his association with the program dated from its inception in 1960. He was born in Appleton, Wis., and received the B.S.E.E. degree from the University of Wisconsin in 1949. He was then employed as a corrosion engineer with the Northern Indiana Public Service Company, from 1949 through 1953. He earned the M.S.E.E. degree at Purdue University, which granted him the Ph.D. degree in 1958. In the fall of 1958 Dr. Kopplin joined the faculty of the Electrical Engineering Department at the University of Illinois, and during his tenure spent four years with the Coordinated Science Laboratory investigating high magnetic field superconductors. He held visiting appointments at M.I.T. in 1961 and The University of Colorado in 1967 and has been active in development of programs designed to teach the theory and practice of experimentation. Since September 1968, he has been head of the Electrical Engineering Department, University of Texas at El Paso.

The compatibility of materials

When the concept of compatibility is applied to materials, some modification of its meaning becomes necessary—two different materials must be compatible not only with each other but with their environment as well

Walter H. Kohl

National Aeronautics and Space Administration Electronics Research Center

The success of space-age endeavors depends greatly upon the materials used to construct a certain module, part, or subsystem, and the interaction between these materials and their immediate environment. A metal, for example, must retain its physical integrity in shape, ductility, and tensile strength over stringent stress and temperature ranges; it must be able to withstand proximity with other materials, metallic and nonmetallic, while maintaining chemical stability. Nonmetallic materials must be compounded to prevent such reactions as outgassing or decomposition at high altitudes; others, such as ablative ceramics, must disintegrate in a certain manner. In short, the demand for more specialized, more durable substances has created a new materials technology.

The importance of selecting the right kind of material for the design of any device, equipment, or engineering structure has frequently been emphasized.[1] It is quite obvious that a number of questions must be answered before a design concept can be reduced successfully to practice, because the intended use will impose certain desired materials characteristics. Should the material be hard or soft, stiff or flexible, conductive to heat and electricity, or not? How long is the device expected to be in service, and what abuse is it likely to get? What price can it demand in the marketplace to sell in sufficient volume? Can the material be readily fabricated into the desired shape, and can it be joined by available techniques to other components that are part of the structure? Is the material strong enough to carry an expected load and to withstand the extra stresses that may arise under unusual conditions? These are some of the questions that need to be answered. One should also consider the interaction of the material with the user. Thus it would not be very practical to make a snow shovel from a corrosion-resistant alloy, such as stainless steel, and choose a glass rod for a handle, for such a combination would have high cost and short life.

As suggested by this example, materials must be chosen so that they can economically and safely achieve their intended function in a given environment. One must consider compatibility, the main environmental parameters, and the interaction of materials with the environment and with each other.

The meaning of compatibility

In its more general use, the term "compatibility" applies to people. Two persons are referred to as being compatible when they live together harmoniously and are able to face, without undue stress to their relationship, the varying conditions that life presents. When this concept of compatibility is extended to materials, some modification of its meaning becomes necessary; two different materials must not only be compatible with each other but also with their environment if they are to survive without harmful effects. One given material may not be compatible with its environment and suffer drastic changes of its physical characteristics as a result of such exposure.

In some cases materials are compatible with each other in a given environment up to a certain temperature limit and incompatible beyond, regardless of whether they are in contact or not. Tungsten and alumina behave in this manner in a vacuum. The familiar heater for an oxide-coated cathode in an electron tube operates near 1500°C and gives satisfactory service for a long time. Above 1900°C, however, a chemical reaction sets in that reduces the alumina and causes the deposition of a film of aluminum on the wall of the container. This reaction also occurs in the vapor phase, when no direct contact between the reactants exists.[2]

It may happen that materials indirectly interfere with the intended function of the device of which they are a part, or that they have adverse effects on the operator or present a hazard to his safety, or that prescribed test procedures present constraints that must be satisfied.

In a spacecraft, for example, gaskets used for the

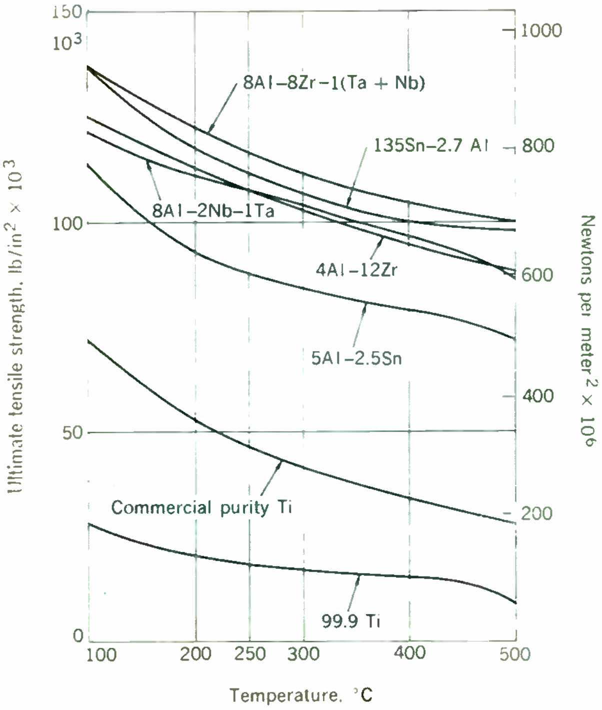


FIGURE 1. Temperature dependence of the ultimate tensile strength of certain titanium alloys compared with that of unalloyed titanium. (Courtesy George Newnes, Ltd., London)

mounting of windows should not give off vapors that cause fogging when deposited on the windowpanes or on camera lenses. Similarly, structural elements made of magnesium alloys containing traces of thorium would be harmful in the proximity of sensitive photographic film or when a gamma-ray experiment is to be performed.

In missions that are aimed at the exploration of extraterrestrial life the landing capsule must be rigorously sterilized so as to prevent the deposition of life forms carried aloft from the earth. This is achieved either by a thermal or a chemical treatment, or both. A heat sterilization procedure consists of three 36-hour exposure cycles in a nitrogen atmosphere at 150°C, and a typical chemical treatment specifies six 30-hour cycles at 50°C in a gas mixture consisting of 12 percent ethylene oxide (ETO) + "refrigerant 12." Components and structures must be able to withstand such treatment; otherwise, the materials from which they are made are not compatible with the mission. The different compatibility categories that a designer should bear in mind therefore include: (1) compatibility of one material with its environment; (2) compatibility of several materials with each other and with their environment; (3) compatibility of materials with a desired function of adjacent components; (4) compatibility of materials with the health and comfort of the operator; and (5) compatibility of materials with prescribed processing and test procedures.

Environmental parameters

Ambient conditions must be known before the suitability of a given material for a chosen design can be established. In some cases an artificial environment for materials is created by enclosing them in containers that

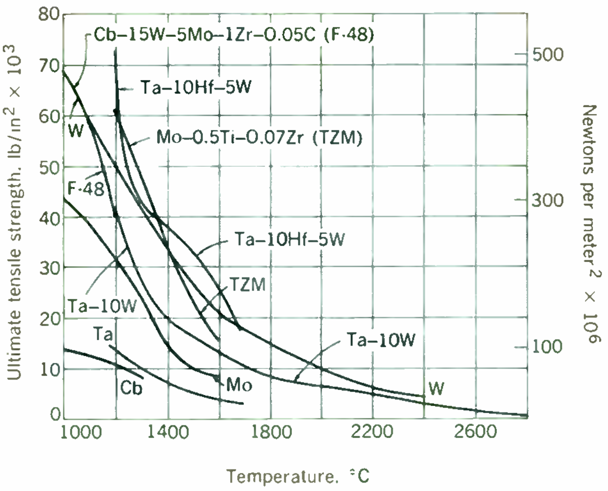


FIGURE 2. Effect of temperature on the ultimate strength of selected refractory metals and alloys. (Courtesy Battelle Memorial Institute)

are either evacuated or filled with desired gases or vapors. Vacuum tubes and mercury-vapor rectifiers are examples of this category. On the other hand, a device may be immersed in a fluid to safeguard its reliable operation. Thus transformers and switchgear are put into oil-filled containers to reduce the danger of high-voltage breakdown, and components of fuel cells and liquid-cooled reactors are in contact with highly corrosive fluids.

The environment encountered by space vehicles varies widely, and judicious selection of materials for these structures is of critical importance.[3-8] Severe shock and vibration exist during launch of the vehicle, intense radiation and particle bombardment are met at high altitudes, and the re-entering capsule must endure extremely high temperatures. These problems have been the subject of many conferences and have been extensively treated in the literature.[9-11]

The interaction of materials

Although materials are often regarded as solid, their transformation into the liquid and gaseous state at high temperatures requires inclusion of all three phases in a consideration of possible interactions, since solid materials may have to exist compatibly in a liquid or gaseous environment as previous examples have shown.

It is also essential that the materials themselves are clearly defined,[12] not only in terms of their chemical composition and impurity content but also by a quantitative description of surface texture and the presence or absence of dislocations and strain. The extent to which such characterization is required will vary from case to case, but a duplication of test data will not be possible unless the full story is told.

The interactions encountered may be of chemical or physical nature, leading to oxidation, reduction, compound formation, to diffusion and bonding, or to evaporation and film formation, to name a few possible effects. When the compatibility of materials is to be evaluated under a given set of conditions, it is important that these interactions are predictable.

Most of the basic property data of materials are avail-

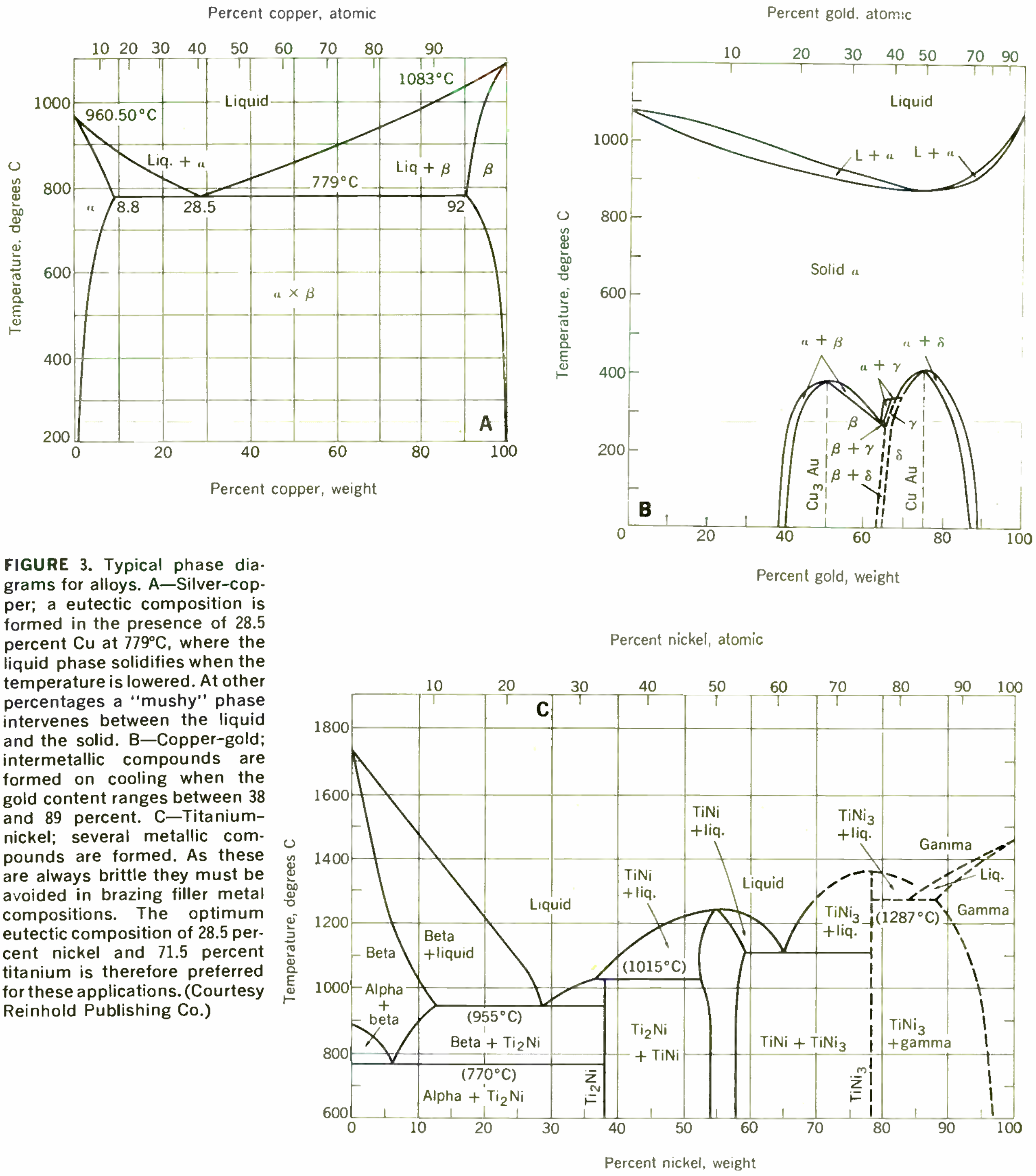


FIGURE 3. Typical phase diagrams for alloys. A—Silver-copper; a eutectic composition is formed in the presence of 28.5 percent Cu at 779°C, where the liquid phase solidifies when the temperature is lowered. At other percentages a "mushy" phase intervenes between the liquid and the solid. B—Copper-gold; intermetallic compounds are formed on cooling when the gold content ranges between 38 and 89 percent. C—Titanium-nickel; several metallic compounds are formed. As these are always brittle they must be avoided in brazing filler metal compositions. The optimum eutectic composition of 28.5 percent nickel and 71.5 percent titanium is therefore preferred for these applications. (Courtesy Reinhold Publishing Co.)

able from handbooks, although their reliability leaves much to be desired when one searches for information that is valid at high temperatures. The melting point of a given refractory oxide, for example, may be indicated by numbers that widely differ from each other, depending on the measurement technique and on the purity and microstructure of the material. Even for metals, data may differ by large amounts. Thus, the melting point quoted for titanium (1660°C) has been going down by several hundred degrees during the past decade as material of higher purity became available. Figure 1 shows, on the other hand, how judiciously chosen alloy additions to Ti markedly increase the ultimate tensile strength (UTS) of this material, which has attained such importance in the construction of supersonic aircraft and for space vehicles. Figure 2 portrays the adverse effect of temperature on UTS, already apparent in Fig. 1, for a number of other refractory metals and alloys when these are exposed to much higher temperatures.

The formation of alloys of different metals is readily appraised from a study of their phase, or equilibrium, diagrams,[13] which indicate the relative solid solubility of the components and the formation of intermetallic compounds (Fig. 3). Phase diagrams for ceramics are also available, and serve to predict the interaction between these materials.[14] As the name implies, these diagrams are based on measurements made when the two components were in chemical and physical equilibrium. During processing and in operation of a device, time may not be available to establish equilibrium, so that deviations from the predicted behavior will occur. The same comment applies when predictions are made on the basis of independent thermodynamical calculations that aim at determining the "free energy" of a reaction at constant temperature and pressure.

Such thermodynamical considerations are a valuable guide for the appraisal of reactions that are likely to occur in a closed system. Four long-established basic laws, i.e., the zeroth, the first, the second, and the third laws of thermodynamics, have general validity on a macroscopic scale, but do not assume or imply any specific mechanism of interaction on the atomic scale. As the behavior of materials, and their properties, are dictated by atomic and electronic structure, an understanding of this behavior requires more careful investigation.

The free energy of a reaction at constant temperature and pressure, mentioned above, is also called the *Gibbs function* and is expressed by the symbol ΔF. It is the fraction of internal energy that is available to do external work, and may thus be looked upon as the driving force in any physical or chemical change. Under equilibrium conditions, when no such change occurs, the free energy of the system is minimal. By definition, the larger the calculated value of $-\Delta F$, the more likely a reaction between two components or a change of phase of one component will take place.

In some cases an "activation energy" may be required before the free energy of the system is released. The reader may have to consult texts on thermodynamics to refresh his memory on these matters, but the following equations may serve this purpose. In general, the relation between free energy and activation energy may be represented by

$$\Delta E = Q - W$$

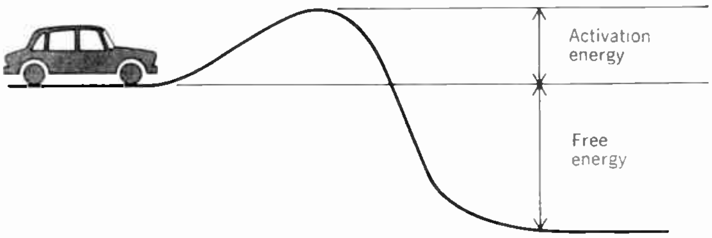
$$\Delta S = \frac{Q}{T}$$

$$\Delta E = T\Delta S - P\Delta V$$

$$H = E + PV$$

$$\Delta F = \Delta H - T\Delta S$$

FIGURE 4. An example of the relationship between activation energy and free energy, where an "activation energy" must be applied to release the free energy of a system. Thus energy must be applied to get the car to the top of the hill before it can coast to the bottom.



where E is equal to internal energy, Q equals heat added, W is external work, S equals entropy, T is absolute temperature, V is equal to volume, P represents pressure, F is free energy, and H equals enthalpy.

Figure 4 illustrates the relation between free energy and activation energy. Thus, energy must be supplied to get the car to the top of the hill before it can coast to the bottom.

Thermodynamic calculations may allow one to predict the likelihood of a reaction, but they cannot foretell how fast a possible reaction will proceed. The reaction rate is an important parameter in both physical and chemical events, and it depends on the arrival of atoms at the reaction site, once a reaction has been nucleated. Such considerations apply in oxidation and corrosion, absorption, grain growth, sintering, and phase transformation, to name a few examples.

Nucleation is assumed to take place at point defects, dislocations, or grain boundaries. Neighboring atoms are then fed to the nucleation site either by surface migration or by diffusion through the lattice. Diffusion may take the form of vacancy diffusion or interstitial diffusion, and it may also proceed by more complex mechanisms. It is a difficult experimental task to distinguish between the different types of reaction.

Imperfections such as point defects, dislocations, and grain boundaries may be looked upon as internal surfaces. Just as external surfaces are an interruption of extended order and the seat of surface free energy, so is an internal surface free energy associated with imperfections. For this reason impurities that are present in a crystal preferentially precipitate at grain boundaries and contribute to internal friction and oxidation. These subtle effects can have very practical implications concerning materials compatibility.

Stress corrosion

Many metals and alloys exhibit stress corrosion when they are exposed to certain hostile environments in the presence of internal tensile stress.[15-19] Such stresses may be induced by nonuniform heat treatment, by differential thermal expansion, or by nonuniform distribution of an external load, to name a few possible causes. After a given lapse of time, which varies with existing conditions, cracks develop that may lead to brittle fracture in metals that ordinarily show only ductile fracture when stressed to the ultimate limit of strength. The cracks may occur along the grain boundaries (intergranular) or across the grains (transgranular, or intragranular), and they may be observed in environments where ordinary corrosion is practically absent. The reason for such behavior is not fully understood to date; a complete theory of stress corrosion remains to be established, but would be of obvious value.

Although pure metals are not subject to this effect, there are enough impurities present in commercial metals to make them fall into the category of alloys. A great many acid and alkaline solutions, including ammonia and the halogens, are harmful environments. High-alloy steels and aluminum experience stress corrosion in hydrogen, as do many body-centered cubic metals, as well as titanium and its alloys, α-zirconium, and α-uranium. Copper is adversely affected in the presence of stress in nitrogen and ammonia atmospheres when the metal contains as little as 0.004 percent phosphorus.

Stress corrosion can be avoided by a number of preventive measures, but a free choice is not always available to the designer. Assuming that the environment is determined by the task at hand, the selection of a material that is not subject to stress corrosion would be an obvious approach. One might also strive to keep the stress level below a safe limit if the substitution of a different metal is not feasible. On the other hand, the surface of the metal may be protected by suitable coatings so that the harmful medium has no access to it. But then another potential hazard arises in the form of cracks in the coating that may develop in service. Shotpeening, a process that puts the surface of the metal into compression, has been a useful remedy in some cases.[20] Sometimes it is possible to add an inhibitor either to the metal or to the fluid with which it is in contact. Thus, workers at the George C. Marshall Space Flight Center of NASA were able to use the latter approach when the storage of nitrogen tetroxide (N_2O_4) in titanium (Ti-6Al-4V) pressure tanks led to stress-corrosion failure.[15] It was found that the addition of 0.25 percent (or more) of nitric oxide (NO) to the N_2O_4 acted as an effective inhibitor of stress corrosion without adversely affecting the characteristics of the rocket engine propellant. Incidentally, most of these propellants, of which liquid hydrogen, liquid oxygen, and hydrazine are other examples, are also highly corrosive in the ordinary sense, and very reactive; the container metal may even promote their decomposition and subsequent failure. A report issued by the Defense Metals Information Center discusses the problem in greater detail.[21]

Other environmental failures

Although a number of specific applications have been mentioned as examples of material interaction, not much has been said about mechanical structures and the effect of temperature upon their stability. Such structures are intended to bear a load, and they provide the framework for the distribution of stresses. The mechanical, civil, aeronautical, and aerospace engineers select materials to be strong enough for this purpose; effects such as vibration, impact loads, creep under prolonged stress, and fatigue in the presence of cyclic loading also must be considered. Expansion and contraction under the influence of temperature variations are additional parameters that must be taken into account, particularly where tolerances are critical.

When different materials are joined, differential thermal expansion can become a serious problem unless special precautions are observed; an example taken from the electron tube industry will illustrate this point. Glass-to-metal and ceramic-to-metal seals may be of the "matched" or "mismatched" type, depending on the chosen materials and the geometry of the design.[2] In matched glass-to-metal seals the metal should have the same coefficient of thermal expansion as the glass, over as wide a temperature range as possible, so that tensile stresses in the glass are avoided both during manufacture and in operation.

Mismatched components can be utilized, however, when the metal is made thin enough at the joint so that thermal expansion stresses are relieved by flexing of the metal (Houskeeper seals), or when compression is purposely introduced in the glass in cylindrical assemblies where an outer metal ring of much larger expansion coefficient is used. Glass is subject to failure when stressed in tension, but it is strong in compression; the same holds true for ceramic-to-metal seals. It is apparent from these examples that certain types of stresses can be desirable in a particular design and actually contribute to the reliability of the device.

Composite materials

Glass, ceramics, metals, and plastics are frequently combined to form composites because the physical properties resulting from such a combination are often superior to those exhibited by one compound alone. Much research and development effort has been devoted to this subject during the past ten years.[22-24] The second compound is generally introduced in the form of very fine fibers. Reinforced plastics utilize high-strength glass fibers; ceramics may be combined with refractory metal fibers, or whiskers; and some metals can be strengthened by incorporating refractory metal or high-strength ceramic fibers.

An intimate bond between the two components must be established during the manufacture of such composites so that the stress is transferred from the matrix to the fibers when an external load is applied. The elastic modulus of the fibers, or filaments, must be substantially higher than that of the matrix material, and the fiber content of the composite usually comprises about 20 percent of the bulk volume.

Composite materials are being used for fishing rods, pleasure craft, tank cars, helicopter rotor blades, wings and stabilizers in aircraft, and motor cases in rockets, to name just a few examples, resulting in substantial weight reduction in comparison with previously used bulk structures. Fibers of carbon, boron, silicon carbide, and boron carbide are being investigated as additives for composites to be used in space vehicles because of their favorable high-temperature strength-to-weight ratio. Efforts are also under way to impart heat-shock resistance and improved tensile strength to the high-melting-point carbides of silicon, tantalum, titanium, hafnium, and zirconium by reinforcement with tungsten fibers.

The beneficial effect of the fibers would be forfeited, however, if chemical reactions at the matrix interface caused the consumption of the filaments, or if brittle intermetallic compounds were formed; diffusion of one of the matrix components into the fibers also could produce inferior physical properties that would endanger the soundness of the structure. Silicon carbide, for example, reacts with tungsten, and boron carbide with both tungsten and tantalum, whereas carbides of hafnium, zirconium, and titanium are free from reactions with tungsten.

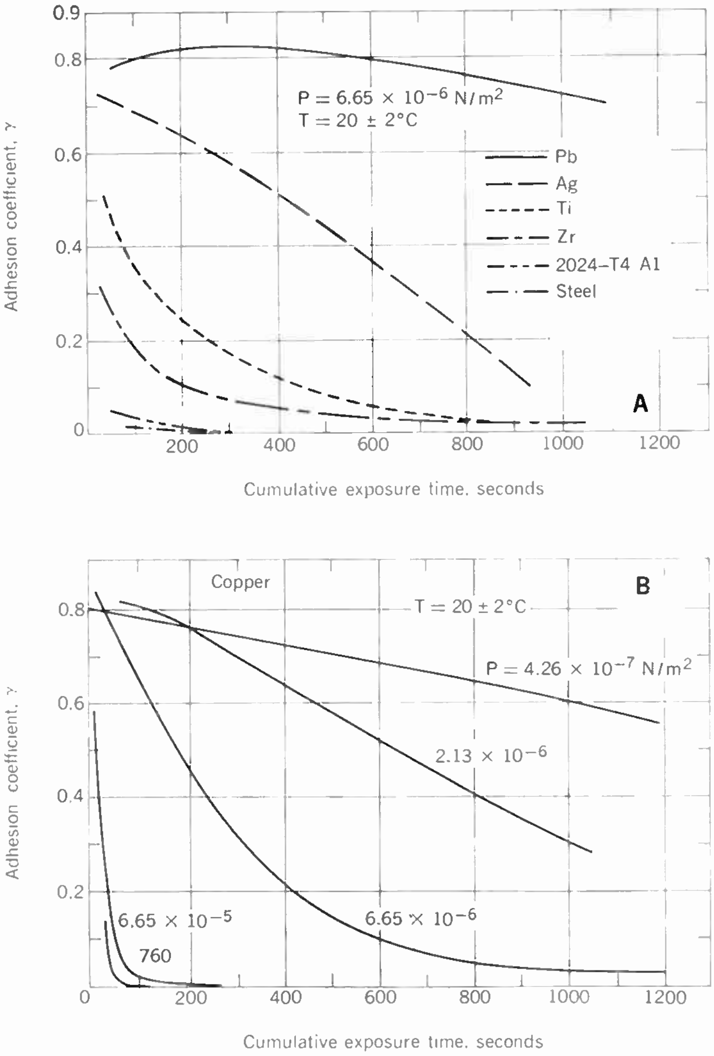
A degrading reaction would be probable in aluminum strengthened with boron filaments by either the formation of the borides AlB_2 and AlB_{12} or the eutectic reaction between aluminum and AlB_2 at 648°C. A liquid phase will form, rapidly dissolve most, if not all, of the filament, and then solidify on cooling to produce a localized, cast structure. Fabrication temperatures are therefore held at 620°C or below. The formation of AlB_2 at temperatures below 704°C can be prevented by coating the boron filament with nickel, but Al_3Ni forms if liquid aluminum is in contact with the fibers for more than a few seconds at 574°C, according to a recent report by Greenstine.[25]

The fibers, or filaments, therefore must be carefully selected for each given application.

Liquid metals and alkali vapors

Many materials are severely attacked by alkali vapors when exposed to them at high temperatures. This condition prevails in thermionic energy converters and liquid-cooled reactors, where metals and ceramics must be chosen with great care if prolonged service is to be ensured. Very small amounts of impurities contained in a given material can have serious, deleterious effects. To give an example, it has been found by workers at the Eimac Division of Varian Associates, Palo Alto, Calif.,[26, 27] during their development of "bore seals" for liquid-cooled reactors, that the silica content of alumina must be kept below 0.05 percent if this material is to survive in an alkali vapor atmosphere at 1000°C for 1000 hours. This requirement imposes severe limitations on the fabrication of large alumina cylinders. The joining of these ceramic members to the metallic housing also presents problems in material selection. Although stainless steel, zirconium, niobium, and Zircaloy® withstand operation in a cesium atmosphere at high temperature, niobium is adversely affected by brazing filler metals that contain nickel; it should be plated with iron to prevent this attack. Similarly, brazing alloys containing copper, silver, and gold are attacked by cesium. Although copper itself is resistant and suffers only surface discoloration at temperatures below 600°C, alloys of copper with silver or gold become porous after prolonged contact with cesium at 450°C, thus losing structural strength and durability.

FIGURE 5. Variation of the adhesion coefficient for differen metals and alloys as a function of accumulated exposure time (A) and variation of the adhesion coefficient as a function of accumulated exposure time for copper exposed to different ambient pressures (B). (Courtesy of American Institute for Aeronautics and Astronautics)



Thermionic energy converters present similar compatibility problems in the choice of their structural components.[28, 29] Such converters essentially are close-spaced diodes in which a refractory cathode operates at approximately 2000°C, and the anode at about 700°C; the two are separated by a space measuring a few thousandths of an inch. A cylindrical geometry is frequently adopted; and a ceramic cylinder made from high-purity alumina serves as the envelope that is filled with cesium vapor at a pressure of about 133 newtons per square meter. Electrons emitted by the hot cathode are collected at the anode, so that useful power is developed in an external load when the space charge is neutralized by cesium ions formed by contact ionization at the cathode. Thermal energy, used for heating the cathode, can be converted directly into electric energy in this manner. The efficiency of this device depends upon the difference between cathode and anode work functions, and may reach a theoretical value of 35 percent; 15 percent has been achieved in practice at power levels of several watts. Atomic fission may be used to heat the cathode; consequently, nuclear power systems for spacecraft are a potential application of thermionic energy converters. Therefore, substantial effort has been devoted to the realization of these devices, but many difficulties remain to be overcome before long life and efficient service can be achieved.

Cathodes for thermionic energy converters using nuclear fission as a heat source are commonly made of a composite of uranium and zirconium carbides (two thirds UC, by weight, and one third ZrC) because they have better thermal and electrical conductivities than the oxides. Zirconium carbide is added to counteract the tendency of unadulterated UC to crack when heat-cycled to temperatures near 2000°C. Evaporation from the surface of such composite cathodes becomes a major problem, so that their enclosure in a more refractory material is indicated. This raises the important problem of compatibility of UC + ZrC with possible sheath materials such as tungsten, tantalum, molybdenum, niobium, and rhenium at temperatures in the 2000°C range. Studies conducted at Los Alamos and Harwell[30] have shown that niobium completely disintegrates at 2000°C due to its intergranular penetration into the cathode and the simultaneous diffusion of uranium and zirconium into the niobium. Molybdenum reacts catastrophically with the cathode pellet at 1850°C, and tantalum begins to show reaction after heating to 2170°C, and more extensively at 2285°C, although the reaction does not proceed as rapidly as with molybdenum. Tungsten does not show any reaction in the 2000–2200°C range for periods up to one hour, and good compatibility is found for rhenium up to 2100°C. It would then appear that tungsten, tantalum, and rhenium are satisfactory for long-term service, up to 2000°C.

Vacuum friction

When certain clean metals are brought into contact with each other in a high vacuum, such as prevails in outer space, in a space simulator, or in a sealed-off vacuum device, they tenaciously adhere to each other and form a cold weld.[31] Advantage is taken of this effect when pinching-off copper tubulations of power tubes or in making "polyoptic seals"[32,33] on fused silica and ceramics. The application of moderate pressure, and an increase in temperature, facilitate this effect. Such mating of surfaces occurs more readily when the two materials exhibit the same crystal structure, either body-centered cubic (b.c.c.) or face-centered cubic (f.c.c.), but recently it has been demonstrated that dissimilar metals may adhere to each other. Atomic size should not differ by more than 15 percent, a condition that favors alloy formation. Other factors, such as hardness, modulus of elasticity, crystal orientation, electron affinity, and the number of valence electrons per atom, play an important role in choosing critical materials.

Gilbreth and Sumsion[34] of the NASA–Ames Research Center have conducted experiments on solid-phase welding of metals in a high vacuum and demonstrated that a "coefficient of adhesion" can be expressed by the following formula:

$$\gamma_\epsilon = \frac{C(\text{UTS})e_f}{H}$$

where UTS is expressed in kg/mm², elongation to fracture e_f in mm/mm, Vickers hardness number (H) in kg/mm², and the correlation factor C is 4 for the materials under test. Notched rods, about 1.3 cm in diameter by 5 cm long, were fractured in ultrahigh vacuum and then rejoined by the application of a compressive load. The coefficient of adhesion represents the ratio of the tensile force required to produce or separate the two surfaces to the compressive force applied to join them after initial fracture. This coefficient is a function of the time of exposure to the surrounding atmosphere, or the degree of vacuum prevailing near the test sample. Figure 5(A) shows how the adhesion coefficient for several different metals and alloys varies with accumulated exposure time, and Fig. 5(B) indicates the effect of ambient pressure on one particular metal, i.e., copper. It was found that the procedure of breaking and pressure welding could be repeated as often as desired, or until adhesion was no longer measureable. Of the various gaseous contaminants tested, such as air, oxygen, nitrogen, carbon monoxide, and carbon dioxide, only oxygen significantly affected the adhesion of copper. Chemisorption, rather than physisorption, is apparently the controlling parameter. Except for 2024-T4 aluminum, correlation of the measured and calculated data has been quite satisfactory to date.

Winslow and McIntyre,[35] at Hughes Aircraft Company, recently conducted tests of the adhesive forces that are developed on contact of various metal couples, and find that these are lower in static tests than they are in dynamic tests where the load is repetitively applied several times a second. Thus, type 304 stainless steel did not show adhesion to itself in a static test over the temperature range from 25° to 500°C at a pressure of 1.33×10^{-7} newtons per square meter, but did so in a dynamic test. On the other hand, a contact couple using 304 S.S., with type 2014 aluminum as a partner, developed static adhesion at 300°C and dynamic adhesion at 150°C.

In many practical applications, the welding of materials brought into contact in a vacuum is highly objectionable, and must be prevented. Examples include sliding and rotating contacts, where friction and wear must be held to a minimum. Rotating anodes in X-ray tubes and tuning plungers in magnetrons are examples familiar to the tube engineer. In space applications, steerable antennas and movable solar panels require shaft rotation in bearings; motor brushes sliding on rotors must conduct substantial currents, and should do this at a constant contact resistance and without creating electrical noise that would interfere with communication systems. Since conventional lubricants cannot be used in any one of these applications, compatible materials must be chosen that give satisfactory service under such conditions. Steel bearings in rotating-anode X-ray tubes are coated with silver, a soft metal that minimizes friction and wear. Silver-impregnated graphite brushes have served well in earth-bound applications, but they do not stand up in a vacuum, because it has been found that the lubricating action commonly associated with graphite depends on the presence of water vapor from the atmosphere. Workers at Lockheed[36] have developed a new brush material for space applications that consists of a composite of 82.5 to 85 percent silver, 2.5 percent copper, and 12.5 to 15 percent molybdenum disulfide; very satisfactory results have been obtained. In cases where electrical conduction is not required in a sliding contact, cermet materials have been found useful.

Electric contacts

As suggested in the preceding paragraph, electric contacts are devices in which the compatibility of materials is of critical importance.[2] Such contacts may have to operate under widely varying conditions. Relays and switches may be exposed to the atmosphere, or they may be hermetically sealed in a container that is either filled with an inert gas or evacuated. Contact closure current may be small or large, and the force exerted to close the circuit may vary greatly. The current must be interrupted within a specified small fraction of a second when the force acting on the contact is released. Even thin tarnish films can prevent current flow in a light-duty relay where contact forces are small, but they may be readily pierced in a heavy-duty switch where these forces are large. The formation of arcs must be prevented, because excessive heat may cause melting, and possible fusion, of the contact metals.

Apart from oxidation and erosion, strange effects may take place on specific metals in particular environments. Tin, zinc, and cadmium are known to grow whiskers when their surfaces are stressed; this happens even when the metals are exposed to a nitrogen atmosphere or when they are immersed in a hydrocarbon oil. Such whiskers can cause short circuits in a variety of electric devices; when they grow outside the active area, for example, on armatures and shield cans, they may be dislodged by shock and vibration and land in the gap of a relay. Sliding contacts made of precious metals may develop "frictional polymers" on their surfaces when operated in the presence of vapors released by plastics and organic coatings, or present as a contaminant. It is assumed that the polymer is formed by catalytic action of the metals in the presence of electron emission that occurs during sliding. Platinum, palladium, and rhodium readily produce this effect, and gold and silver are practically free of it. According to Chaikin,[37, 38] formation of such frictional polymers can be prevented by the application of a small amount of diluted tetraethyl lead.

If silver is used in contact with fibrous insulation, "silver migration" often occurs when appreciable voltage exists between two electrodes, especially in the presence of high humidity. Under these circumstances, little silver threads grow from one of the electrodes, ultimately producing considerable leakage. Silver and copper form sulfide films in H_2S atmospheres that readily spread over metal and plastic surfaces, increasing the contact resistance of metals and causing leakage over insulators.

Contact matches also play an important role in the fabrication of solid-state devices and integrated circuits.[39, 40] The deposition of thin films and the application of leads by thermocompression bonding bring into play a host of problems whose discussion is beyond the scope of this article.

REFERENCES

1. Ruskin, A. M., *Materials Considerations in Design.* Englewood Cliffs, N.J.: Prentice-Hall, 1967.

2. Kohl, W. H., *Handbook of Materials and Techniques for Vacuum Devices.* New York: Reinhold, 1967.

3. Benesovsky, F., ed., "Metals for the space age," *Plansee Proc. 5th Seminar,* Metallwerk Plansee, Reutte/Tirol, Austria, 1964.

4. Goetzel, C. G., Rittenhouse, J. B., and Singletary, J. B., eds., *The Space Materials Handbook.* Reading, Mass.: Addison-Wesley, 1965.

5. Kennedy, A. J., *The Materials Background to Space Technology.* London: George Newnes, 1964.

6. Parker, E. R., ed., *Materials for Missiles and Spacecraft.* New York: McGraw-Hill, 1963, chap. 5.

7. Rittenhouse, J. B., and Singletary, J. B., *Space Materials Handbook,* 2nd ed., Suppl. 1 (NASA SP-3025, 1966); Suppl. 2 (AFML-TR-64-40, 1968), U.S. Air Force Materials Lab., Wright-Patterson AFB, Ohio.

8. "The effects of space environments on materials," presented at the 11th Nat'l Symp. Society of Aerospace Materials and Process Engineers (SAMPE), North Hollywood, Calif., Apr. 19, 1967.

9. Campbell, I. E., and Sherwood, E. M., eds., *High Temperature Materials and Technology.* New York: Wiley, 1967.

10. Hiester, N. K., ed., "High temperature technology," Internat'l Symp., Stanford Research Inst., Menlo Park, Calif.

11. Schwab, G. M., ed., *Reactivity of Solids.* New York: Elsevier, 1965.

12. "Characterization of materials," (AD 649 941), Materials Advisory Board, Nat'l Acad. Sci., Washington, D.C., Mar. 1967.

13. Hansen, M., and Anderko, K., *Constitution of Binary Alloys,* 2nd ed. New York: McGraw-Hill, 1958.

14. Levin, E. M., Robbins, C. R., and McMurdie, H. F., *Phase Diagrams for Ceramists,* 3rd ed. Columbus, Ohio: The American Ceramic Society, 1964.

15. Cataldo, C. E., *Stress Corrosion* (NASA TM X-53610, 1966), George C. Marshall Space Flight Center, Huntsville, Ala.

16. Logan, H., *The Stress Corrosion of Metals.* New York: Wiley, 1966.

17. Logan, H., "The stress corrosion cracking of metals," *Metals Eng. Quart.,* vol. 5, May 1965.

18. "Stress corrosion cracking in aircraft structural materials," North Atlantic Treaty Organization Symp. (AGARD), Turin, Italy, Apr. 1967.

19. Williams, J. C., "Some observations on the stress-corrosion cracking of three commercial titanium alloys," *Am. Soc. Metals Trans. Quart.,* vol. 60, pp. 646–650, Dec. 1967.

20. Zorller, H. W., and Cohen, B., "Shot-peening for resistance to stress-corrosion cracking," *Metals Eng. Quart.,* vol. 6, pp. 16–20, Feb. 1966.

21. "The compatibility of materials with rocket propellants and oxidizers," Memorandum 201 (AD 613553), Defense Metals Information Center, Battelle Memorial Inst., Jan. 29, 1965.

22. "Fiber composite materials," American Society for Metals Seminar, Metals Park, Ohio, 1965.

23. Atkin, H. P., "Boron-filament reinforced plastic composites for aircraft structures," *Metals Eng. Quart.,* vol. 7, pp. 9–17, Feb. 1967.

24. Broutman, L. J., and Krock, R. H., eds., *Modern Composite Materials.* Reading, Pa.: Addison-Wesley, 1967.

25. Greenstine, R. B., "Status report on advanced metal composites," *Metals Eng. Quart.,* vol. 7, pp. 8–16, Feb. 1967.

26. McRae, W. C., Reed, L., Toth, J., and Lindberg, R., "Compatibility of alumina and beryllia ceramics with potassium vapor." *Am. Ceram. Soc. Bull.,* vol. 47, pp. 484–488, May 1968.

27. Reed, L., and McRae, R. C., "Ceramic-metal seals for liquid metal environments," presented at the SAMPE Symp., Seattle, Wash., Nov. 1963.

28. Allen, R. C., and Skeen, C. H., "Materials requirements for thermionic energy conversion," *Metals Eng. Quart.,* vol. 4, pp. 36–45, Feb. 1964.

29. Yang, Ling, and Carpenter, F. D., "Materials problems in cesium thermionic energy converters," *J. Electrochem. Soc.,* vol. 108, pp. 1079–1086, Nov. 1961.

30. Adam, J., "Materials for nuclear thermionic converters," *Advances in Materials.* New York: Pergamon, 1966.

31. Cameron, J. A., and Disanatis, J. A., "Frictional behavior of materials," *Metals Eng. Quart.,* vol. 7, May 1967.

32. Bleuze, J., and Dussausoy, P., "Sealing techniques by optical polish (in French)," *Le Vide,* vol. 7, pp. 1182–1190, May 1952.

33. Danzin, A., and Destois, E., "Optically polished surfaces for vacuum tube seals (in French)," *Ann. Radioelec.,* vol. 3, pp. 281 289, Oct. 1948.

34. Gilbreth, W. P., and Sumsion, H. T., "Solid-phase welding of metals under high vacuum," presented at the 6th AIAA Annual Structures and Materials Conf., Palm Springs, Calif., 1965.

35. Winslow, P. M., and McIntyre, D. V., "Adhesion of metals in the space environment," *Vacuum Sci. Technol.,* vol. 3, pp. 54–61, Mar. 1966.

36. Clauss, F. J., "Sliding electrical contacts for ultra-high vacuum," *Rev. Sci. Instr.,* vol. 36, p. 1251, Aug. 1965.

37. Chaikin, S. W., "Study of effects and control of surface contaminants of electrical materials," Contr. DA-36-039-sc-74965/89150, Stanford Res. Inst., Menlo Park, Calif.

38. Chaikin, S. W., "Inhibition of frictional polymer formation on rubbing contacts." *Proc. Internat'l Conf. Electromagnetic Relays,* Sandai, Japan, pp. 78–80, Oct. 1963.

39. Goldberg, M. F., and Vaccaro, J., eds., *Physics of Failure in Electronics.* Baltimore, Md.: Spartan, 1963.

40. Harlow, J. E., III, and Josephs, H. C., "Compatibility in monolithic integrated circuits," *Electro-Technol.,* June 1967.

Walter H. Kohl (F) is a native of Germany, where he received the Dipl. Ing. and Dr. Ing. degrees in engineering physics from the Technical University in Dresden. He emigrated to Canada in 1930 and joined the engineering staff of Rogers Electronic Tubes, Ontario, ultimately becoming a director of the company. From 1935 to 1940 he was also a special lecturer at the McLennan Institute of the University of Toronto, where he initiated work on the electron microscope. From 1952 to 1958 he was Research Associate and lecturer at the Electronics Research Laboratory, Stanford University. He joined the Electronics Research Center of NASA, Cambridge, Mass., in 1966, where he is now chief of the University Affairs Office. During the past year he has been concerned principally with information processing and retrieval in relation to exploration of earth resources via satellite.

Ceramic capacitors for hybrid integrated circuits

The high-dielectric-constant ceramic chip represents one of the most effective solutions to the problem of obtaining capacitance in the fabrication of hybrid thick-film integrated circuits

Donald W. Hamer *Erie Technological Products, Inc.*

Ruggedness, wide capacitance range, high volumetric efficiency, and relatively attractive cost have been the main reasons for the popularity of ceramic chip capacitors. Continuing improvements in most of these categories promise to keep the ceramic chip in its present position of prominence. This article considers multilayer, single-layer, and screened-on configurations. In addition, relationships between size, capacitance, and cost are covered for three common ceramic formulations (NPO, W5R, and Z5U).

Like any other material or component, the high-dielectric-constant ceramic chip has its own set of capabilities, special characteristics, problems, pitfalls, advantages, and disadvantages. The fact that high-*K* ceramic is a nonlinear dielectric—a ferroelectric—gives it many characteristics that are not likely to be known or understood by the user. It is the purpose of this article to shed light on some of these characteristics.

What's available today

There are three basic types of ceramic capacitors available to the hybrid circuit manufacturer today. They are (1) the multilayer ceramic chip, (2) the single-layer chip, and (3) the screened-on glass/ceramic materials.

In spite of higher cost, the multilayer type is the most popular of this group because of the extremely wide capacitance range available (seven full decades now) and the attractive mounting geometry.

The single-layer chip is available primarily in an opposed-electrode configuration. This electrode geometry adds to the cost of assembly, since a jumper wire must be used to make connection to the top electrode. The presence of the extra connection also increases the reliability problem. Having only one layer limits the capacitance range (only four decades). Offsetting these limitations are low cost, extremely wide choice of formulations (up to 30 as opposed to three or four with multilayers), and great flexibility in choice of size and shape, with relatively low tooling costs.

Also available are "flip-chip" versions of the single-layer chip, in which both electrodes are on the same side. However, the gain in ease of mounting is at the expense of capacitance range and cost.

In addition to the ceramic chip, there are the glass/ceramic systems. In screened-on form, these systems are currently the subject of much interest and no doubt have a bright future. Use of the screened-on type puts many of the control problems (thickness control, area, firing cycles, etc.) into the lap of the IC fabricator. This is not necessarily an unsurmountable disadvantage, because the hybrid producer has to face similar problems in resistor making. But in effect it puts the IC fabricator into the capacitor-making business. If one has a rather high volume of circuits in which more than one or two capacitors can be screened on in one pass, this approach can be quite attractive. For low-volume producers, or for only one or two capacitors per circuit, it is probably more economical to stick to conventional chips. The screened-on capacitor also offers a possible reliability advantage because of the integrated construction features.

Dielectric pastes are just recently becoming available to hybrid IC users, ranging in *K* values from 5 or 6 to about 600. Relationships between capacitance, area, and temperature coefficient in today's screened-on capacitors are very roughly equivalent to those of a single-layer opposed-electrode chip.

The manufacturing process

The construction principle of the multilayer ceramic chip, whereby electrodes and ceramic are assembled "green" and fired in one piece, is direct and simple in its approach, but not necessarily easily realized. Properly achieved, however, this simplicity is one of the prime reasons for the multilayer ceramic chip's success in crowding out various other approaches.

Although there is considerable room for variation from manufacturer to manufacturer, the most common approach in starting the monolithic chip manufacturing cycle is to mix a slurry that can be cast in thin layers on a flat surface. A doctor-blade technique is normally used to achieve the required thinness and uniformity. Organic binders in the slurry give enough strength and flexibility to the tape after drying for removal and handling.

Pieces can be cut out of this tape and placed in a silk-screening jig, where a precious-metal electrode such as platinum or palladium is screened on. A multiple pattern is commonly used to cut down on assembly costs. A "stick" of multilayer capacitors can be built up by stacking several of these sheets one on top of the other in such a way that alternate layers have electrodes offset (slightly) in opposite directions. When a "stick" of properly stacked sheets of electroded tape is carefully cut into the proper number of individual parts, each is an unfired monolithic capacitor in which alternate electrodes are brought out in opposite directions, resulting in a monolithic stack of electrically paralleled plates. After firing, the exposed electrodes on either end are connected by applying a fired-on precious-metal conductive paste. Silver is often satisfactory for these collecting electrodes, although many users prefer other metal systems because of soldering or migration problems.

The other type of chip—the opposed-electrode type—has a similar start in its manufacturing cycle. After casting, handy-size squares are cut out of the tape, placed in a kiln, and fired to maturity. The fired squares are electroded with a fired-on composition (silver, palladium–silver, etc.) and then cut or scribed into the proper size.

Ceramic chip capacitors are used by hybrid circuit manufacturers in this "bare naked" form. Most chips sold by the industry are now sold with leads and encapsulation, but many parts will eventually be replaced by chips as more circuitry shifts to the integrated approach.

Controlling electrical characteristics

The basic material used by ceramic capacitor manufacturers is barium titanate ($BaTiO_3$). Commercial-grade $BaTiO_3$ has a dielectric constant (*K*) at room temperature of about 1500 to 3000, a characteristic that places this material in a class almost by itself, since most insulators have *K* values of less than 10. In looking at the *K* versus temperature relationship for $BaTiO_3$, the *K* remains relatively stable until 100°C is passed. At about 120°C the *K* value climbs very precipitously to something over 10 000, falling off rapidly as this temperature is passed. The reason for the behavior at this temperature (called the Curie temperature) can be traced to the transformation of the crystalline structure of a material in the perovskite ($CaTiO_3$) class, such as $BaTiO_3$, from a tetragonal to a cubic structure. At temperature below the Curie point, $BaTiO_3$ and most other perovskites are ferroelectric. Above the Curie point they are paraelectrics—that is, not ferroelectrics. It is because these materials are ferroelectric/paraelectric systems that they produce high dielectric constants (very high near their Curie point). Along with the high dielectric constant comes a host of other characteristics, to be described later, that one should understand to use these materials properly.

There are a variety of perovskite materials with Curie points at temperatures other than 125°C that are compatible with $BaTiO_3$ (compatible to the extent that they will combine in a solid-state solution with $BaTiO_3$). The Curie points of the solid-state solutions are usually somewhere between the Curie points of the two individual compounds. One might correctly suppose that combining a perovskite having a low-temperature Curie point with $BaTiO_3$ would result in a solid-state solution system whose Curie point is around room temperature. It is in this way that manufacturers can control the Curie-point temperature and thus can produce a ceramic having a very high *K* at room temperature. The heavy temperature dependence of such a system is obvious from Fig. 1, which shows how the Curie point can be put at various other temperatures by varying the percentage of shifter compounds used. Common Curie-point shifters are lead titanate (upward shift) and strontium titanate (downward shift).

The very high *K* at room temperature of two perovskites in solid-state solution carries with it too much *K* variation with temperature for many applications. To get less *K* variation with temperature, one can add to the basic formulation certain compounds that will depress the high *K* value at the Curie point. These depressor compounds act in the nature of a contaminant, surrounding the pure grains of perovskite material with a diffuse layer. The more depressor material one uses, the lower the *K* is at the Curie point. Along with this reduction in *K* comes a reduction in temperature dependence. This leads to the major tradeoff one has to face in dealing with high-*K* ferroelectric systems. High *K* and temperature dependence are more or less mutually exclusive. Figure 2 illustrates the effect of depressor compounds on a high-*K* mix whose Curie point is at room temperature.

It can thus be seen that proper combinations of depressor and shifter materials with $BaTiO_3$ could result in a family of high-*K* materials with *K*'s ranging from close to 10 000 down to 1000 with increasing temperature stability as the *K* is decreased. This is essentially the approach all manufacturers follow in tailoring dielectric characteristics.

It is worth mentioning that physical characteristics of raw material and processing parameters also have major effects on electrical properties and these can also be very

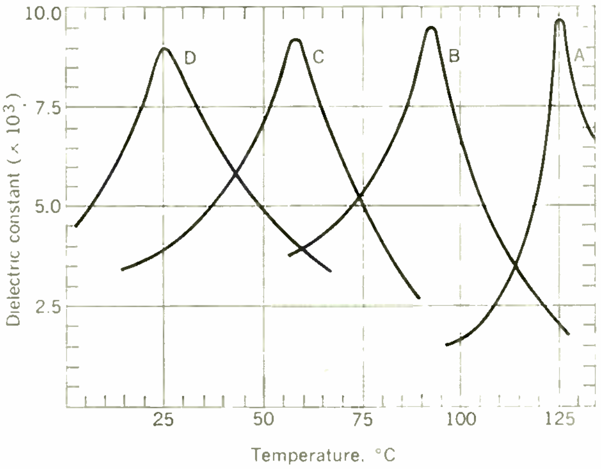
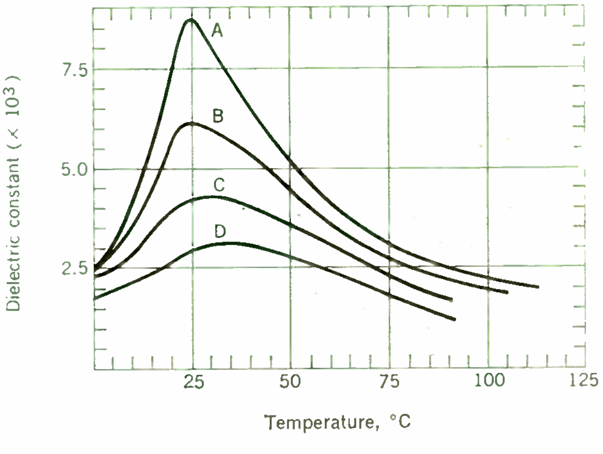


FIGURE 1. Effect of shifter materials in high-K ferroelectric ceramics. A—Pure $BaTiO_3$ (no shifter). B—5 percent shifter. C—10 percent shifter. D—15 percent shifter.

FIGURE 2. Effect of depressor materials in high-K ferroelectric ceramics. A—No depressor (15 percent shifter). B—0.5 percent depressor. C—1 percent depressor. D—2 percent depressor.



effective "handles" for controlling electrical characteristics of high-*K* ceramics.

Capacitance/size relationships

In the previous section we demonstrated that there is a tradeoff between temperature stability and *K* that can result in a family of formulations for high-*K* ceramics. If we couple this flexibility in *K* with the opportunity for varying the area thickness and number of plates, the availability of a wide range of capacitance and performance characteristics in multilayer ceramics becomes obvious. To inject a little simplicity into what could become an overwhelming variety of choices, charts have been prepared showing the capacitance available with different sizes and various number of layers for the three most common ceramic formulations in industry today.

Figure 3(A) shows a material with a *K* of about 6000. This would be pure ferroelectric material depressed just slightly (see Fig. 2). The material is quite sensitive to temperature changes, losing about half of its dielectric constant as the temperature rises to 85°C (the Electronic Industries Association Z5U characteristic). For many applications a change of capacitance with temperature is not particularly important and this material is used very extensively where "brute force" capacitance is needed over relatively small temperature ranges. Z5U chips with *K* values of 6000 are available in ±20 percent and GMV (guaranteed minimum value) tolerances. Closer tolerances than this are usually not offered.

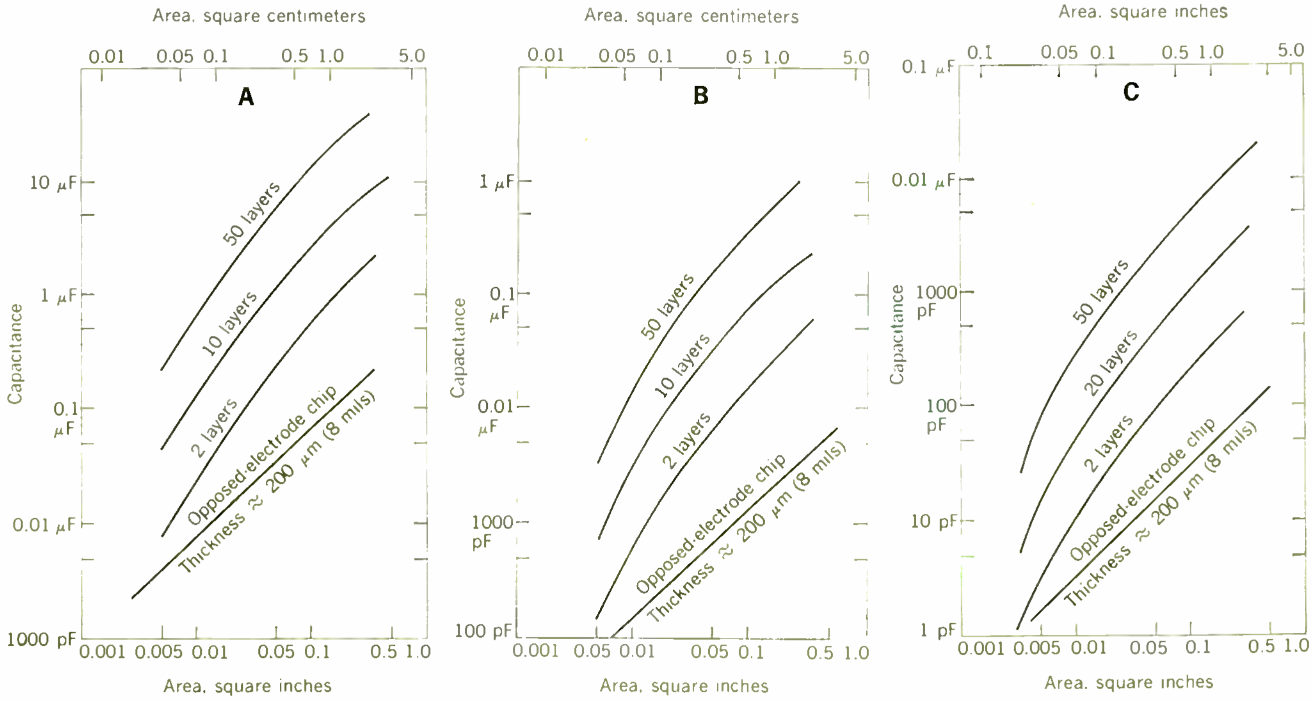
Figure 3(B) shows a stabilized material whose *K* is 1200 and whose temperature coefficient (W5R) varies less than ±15 percent between −55° and +125°C. Besides being used where stability is important, these chips, referred to as K1200, are also commonly used where temperature coefficient may not matter, but extended operating temperatures are encountered. The higher-*K* material continues to lose capacitance and as +125°C is passed, it will have even less capacitance than the K1200 material. Tolerance of W5R chips ranges from ±5 percent to ±20 percent.

Figure 3(C) shows a typical NPO class of material. This material has not been discussed in any detail up to now, but it is really no different from the others in many ways. An NPO material can result if the Curie point is pushed way down and the material throughout its operating range is far out in its paraelectric region. Its main characteristics are low dissipation factor at radio frequencies, and temperature stability. Tolerance of NPO chips is generally around 5 percent. By trimming or selection, tolerances of 1 percent can be achieved. The dielectric constant is only in the 30s, so the capacitance available is considerably lower than in the high-*K* ceramic formulations.

These charts are made up for layers about 50 μm (2 mils) thick, resulting in the various working voltages given in the figure caption. Many manufacturers are able to produce film thinner than this and thus can offer higher capacitance, although at a lowering of working voltage and often at an increase in price.

It should also be mentioned that there is a considerable variety from manufacturer to manufacturer in working-voltage ratings and in dielectric constants available; however, the three chosen here could be viewed as typical. The variety of available single-layer chips is considerably more extensive with regard to the number of formulations available, although the capacitance range is much more limited. More formula variety in multilayer capacitors is developing as demand increases. Several types of temperature-compensating formulations have recently become available.

FIGURE 3. Capacitance vs. size for multilayer ceramic chips. Thickness = 50 μm (2 mils). A—Type Z5U temperature coefficient, K = 6000, dc working voltage = 50. B—Type W5R temperature coefficient, K = 1200, dc working voltage = 100. C—Type NPO temperature coefficient, K = 30, dc working voltage = 200.



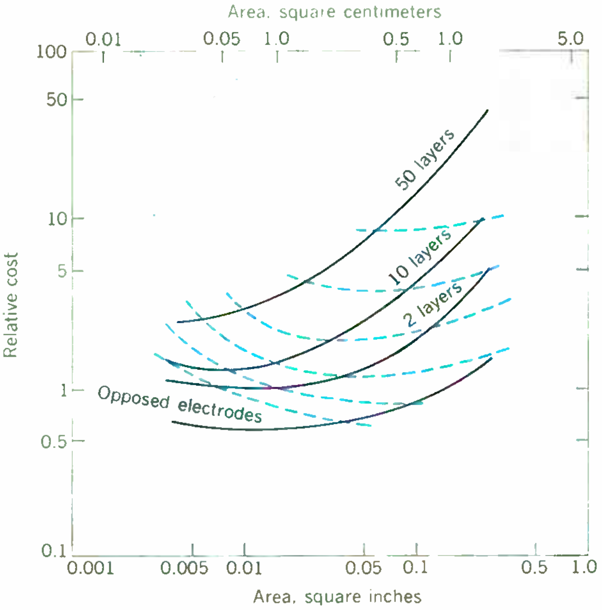


FIGURE 4. Relative costs of multilayer ceramic chips.

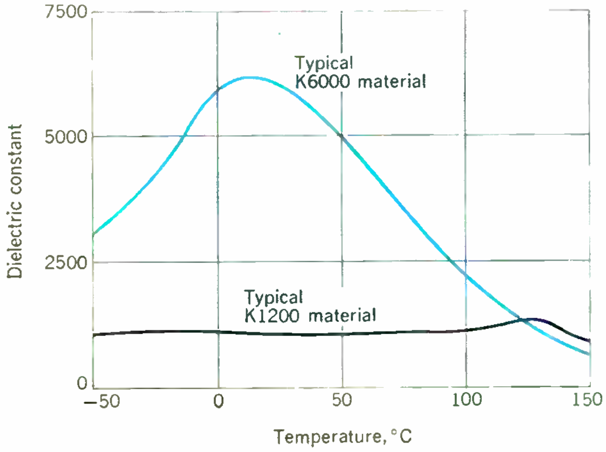
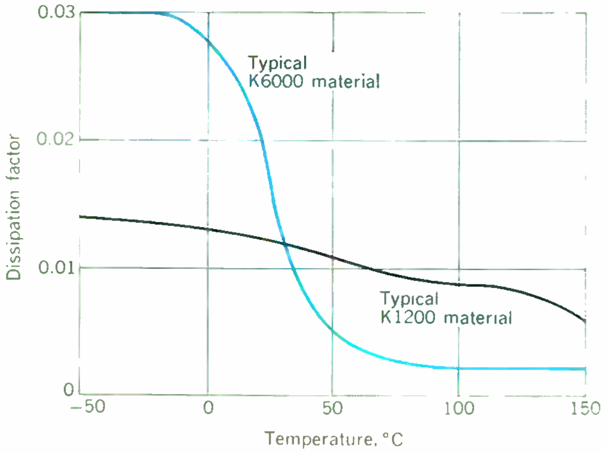


FIGURE 5. Typical variations of dielectric constant with temperature for commercial high-K ceramics.

FIGURE 6. Typical variations of dissipation factor with temperature for commercial high-K ceramics.



Cost relationships

Figure 4 shows relative costs of multilayer chips, normalized to a two-layer 0.065-cm^2 chip. These relative costs would apply to a 50-volt K6000 chip in a ±20 percent tolerance, a 100-volt K1200 chip in a ±20 percent tolerance, or a 200-volt NPO chip in a ±10 percent tolerance. In the case of K1200 chips, improving the tolerance from ±20 to ±10 percent results in a cost increase of about 20 percent. If 5 percent tolerance is desired in an NPO chip, the increase in cost over the 10 percent chip is about 25 percent.

As mentioned previously, many manufacturers can produce these materials in a thinner version—down to about 25 μm thick. Generally speaking, going thinner makes a unit more expensive, but not as much as keeping the same thickness and using more layers to get more capacitance. So, as a rule, if one can sacrifice working voltage (and reliability) it is somewhat cheaper to use thinner dielectric layers to get more capacitance.

Figure 4 also shows equicapacitance lines (dashed lines), each successive line representing about half a decade more capacitance than the last. From this chart, one can look for the most economical size. It shows the price that must be paid for keeping the chip small, although naturally enough large areas also eventually bring on price penalties.

A note of caution about Fig. 4. Please keep in mind that this is a generalized structure, with all the little lumps and bumps of an actual price structure removed. The chart was based on nonautomated assembly procedures, so any firm that assembles certain sizes or values with a highly automated procedure may offer these parts at prices below that reflected on the chart. Very common values and very common sizes tend to be less expensive than odd values and sizes.

Reliability

Reliability is often an important consideration in choosing a chip. The multilayer ceramic chip is a good candidate in general for high-reliability applications, chiefly because of its monolithic construction. In a multilayer chip there is one universal rule to follow in designing for reliability; that is, keep the thickness high. Failure rates are, in general, proportional to the cube of the electronic field. If one has room to use a larger area, he can get higher reliability because this will allow thicker ceramic to be used. Reference to Fig. 4 will indicate that this change is usually not very expensive; in fact, it may sometimes result in lower cost. If there is no space to spare, then the number of layers can be increased. This raises the cost, of course; and, all other things being equal, adding layers tends to reduce reliability to a certain degree, but not nearly by the amount that is gained by using thicker ceramic.

The single-layer chip is, in itself, an inherently more reliable design than the multilayer chip—provided the electrodes are pulled back from the edge. But connection difficulties usually put it back, at best, on a par with the multilayer chip's general level of reliability.

Electrical side effects

In an earlier section the interrelationships between *K* and temperature coefficient were discussed. It should be noted that with ferroelectric chips there is a host of other characteristics that are often also part of the design tradeoff. It is important for the user to know how the various types of ceramic behave, and to avoid the types that will affect the circuit adversely.

Temperature effects on K, Q, conductivity, and life. Figure 5 shows the variation of K with temperature of the K1200 and the K6000 formulation. Note that at the higher temperatures, the stable K1200 material has as high a K as does the K6000 material.

Figure 6 shows the variation of dissipation factor with temperature for K1200 and K6000 materials. (Neither Q nor K is a function of temperature with the NPO formulation.) The jump in dissipation factor of K6000 at lower temperatures is associated with the fact that these temperatures are below the Curie point. The K1200's Curie point is still at +125°C, so the dissipation factor stays relatively high across the entire temperature range.

Conductivity is, generally speaking, an exponential function of the type

$$\sigma = \sigma_0 \exp(-w/kT)$$

where σ_0 and w are constants that vary from one formulation to the next and also with construction design and methods. The w constant is generally a function of the formulation, whereas the σ_0 constant is controlled both by the formulation and the construction. Generally speaking, σ_0 is lower and w is higher in a material of lower dielectric constant. (Thus NPO has a high room-temperature insulation resistance and is less affected by temperature than is K6000.) The conductivity of most high-K ceramics increases by about an order of magnitude as temperature increases from 25°C to 125–150°C. The user has an opportunity to alter this characteristic for the worse by the manner in which he attaches the chip and by subsequent treatment thereof.

Life expectancy in ceramic capacitors is affected by chemical changes, which are related to temperature in a manner similar to the conductivity expression discussed in the preceding paragraph. The constants of the expression are such that the degradation rate doubles every 10°C. The variation between ceramic formulas is not great. This expression, along with the voltage relationship, is one of the cornerstones of accelerated life testing.

Frequency effects. Capacitance and dissipation factor are frequency-dependent in the high-K formulations, but not to any significant extent in the NPO types. Figures 7 and 8 show the variations typically seen in Q and K in high-K formulations. The relationship is not connected with dielectric constant, but rather with the nature of the material used. Variations as great as those shown in Figs. 6 and 7 are evident in various commercial formulations of similar dielectric constant.

Many of the compounds used in formulating temperature-stable ceramics have an adverse effect on frequency stability, so the performance tradeoff crops up again. Some degradation with frequency is traded for temperature stability, high dielectric strength, and good life characteristics. Thus many of today's temperature-stable high-K formulations are not as frequency-stable as one might desire.

Voltage effects. There are three effects that voltage has on high-K dielectric ceramics—K, dissipation factor, and life. Other characteristics are not very sensitive to voltage.

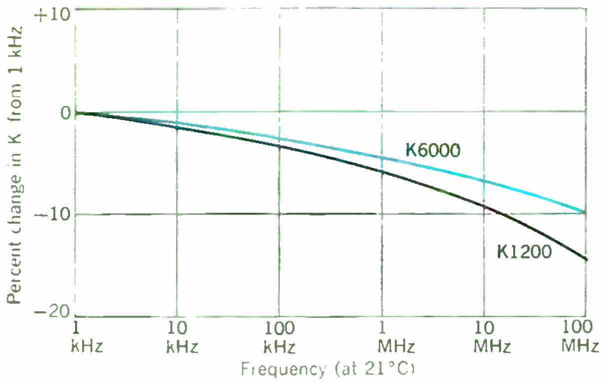


FIGURE 7. Typical variations of dielectric constant with frequency for commercial high-K ceramics.

FIGURE 8. Typical variations of dissipation factor with frequency for commercial high-K ceramics.

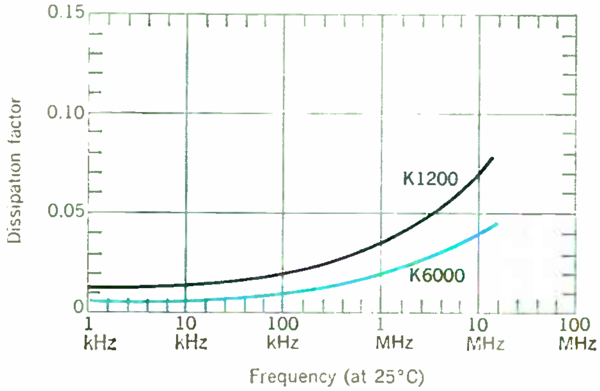


Figure 9 shows the effect of a dc field on K for the two high-K ceramic formulations (NPO is not voltage-sensitive). The percent change is shown as a function of voltage per unit thickness. Therefore, in order to know what capacitance one will see at a particular bias, the thickness of the dielectric must be known. Since thicknesses of 25 to 75 μm (1 to 3 mils) are common in multilayer ceramic capacitors, the effect of a 50-volt bias can be considerable, as can the variations resulting from differences in ceramic thicknesses. In a broad sense, the dc voltage coefficient of K is proportional to dielectric constant, but for a given dielectric constant the coefficient is subject to some manipulation. Such factors as crystal size and the nature of the chemical used can exert considerable influence on the dc voltage coefficient.

High dc fields have a clamping effect that reduces not only K but also dissipation factor. Higher-K ceramics are more sensitive than lower K. About 1.2 V/μm (30 V/mil) will reduce the dissipation factor by half in K6000 but by only 20 percent in K1200.

The life of the part is influenced by voltage as well as by temperature. The expected life varies inversely as the voltage cubed, so a part working in half the dc field of another identical part will last approximately eight times as long before its conductivity reaches a point of being too low to meet specifications.

AC effects. Both K and Q can be strongly affected by an ac field. Generally speaking, capacitance goes up with an increasing ac field (for values below about 1 V/μm or 25 V/mil).

A 1-kHz signal is commonly used for measuring

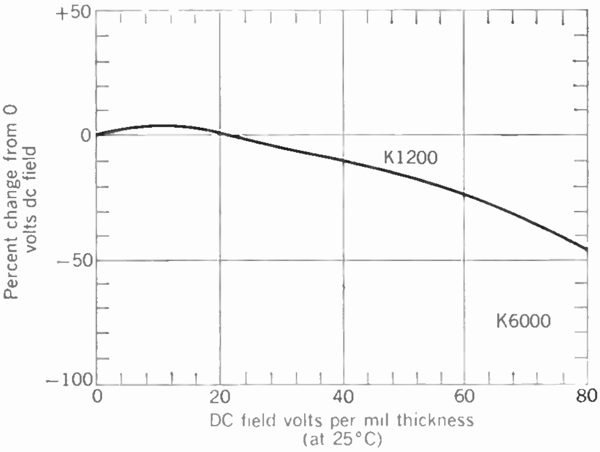


FIGURE 9. Typical dc bias effects on capacitance for commercial high-K ceramics.

capacitance of high-*K* ceramics and the strength of this signal can have considerable effect on the readings one sees, especially in the K6000 formulation. If one varies a 1-kHz signal from 0.5 to 5 volts, a K6000 capacitor with layers 50 μm (2 mils) thick can easily show a 10 percent variation. This has been the cause of many misunderstandings in correlation readings from one firm to another and it is important that one understands what the measuring voltage is to be. K1200 types and lower are very much less affected by ac measuring voltage. NPO is not affected at all.

Time effects. Ceramic capacitors lose some of their capacitance with time. This is a logarithmic relationship and amounts to 1 to 5 percent per time decade. High-*K* types (K6000) range from 3 to 6 percent per decade and stable-*K* from 1 to 3 percent per decade. NPO does not age. Aging starts from the time of transition through the Curie temperature (in other words, the time of last heating—usually the manufacturer's silver firing and/or the user's soldering date). It is associated with a relaxation phenomenon, where stresses within the crystal structure are slowly relieved. As a general rule, high-*K* materials age more than low-*K* materials, but again there is some variation from formulation to formulation. As long as aging is below 5 percent per decade, the changes after a day or so are not significant for most applications.

By this time it should be apparent to all that ferroelectric dielectrics are quite complex, both in their makeup and in their performance. This is, unfortunately, the price one pays for the high *K*. It is not possible to ask for a high *K* and wipe out all the time-, voltage-, and frequency-dependence. The very mechanism that provides high *K* is also responsible for these other characteristics.

The future

In efforts to increase penetration and to keep his product competitive with other types of capacitors, the chip manufacturer is faced with several not always mutually compatible goals. He will try to offer more capacitance in smaller packages, remove or reduce some of the most objectional characteristics discussed in the previous section, and at the same time lower the cost. The degree of success in reaching these goals is subject to speculation, but there is no doubt that in the coming years the already attractive chip will be even more widely accepted.

Many of the producers of chips have been investing heavily in automation, thereby causing costs to come down rapidly. Cost reduction in the past several years has cut prices in half or more, and there is no reason not to expect a continuation of this trend until the multilayer chip is in the same price range as today's disk capacitor—just a few pennies.

Materials and process research and development programs are also progressing rapidly. Work being done now centers around increasing *K*, especially in the stable materials. However, increasing *K* is not the only solution; decreasing thickness is another approach. Thinner sections put the emphasis on understanding and controlling voltage sensitivity and dielectric strength. Working with a fraction of a mil thickness forces the dielectric into high fields for any but nominal dc voltages. It is not unreasonable to expect improvements in capacitance approaching an order of magnitude within the next decade or less.

Conclusion

The chip makers are working hard to increase capacitance and to reduce price. Considerable progress can be expected within the next five to ten years. An order-of-magnitude increase in capacitance range and decreasing prices will probably be forthcoming. This progress speaks well for the continued health of the multilayer ceramic capacitor in general and its application in integrated circuits in particular. Other thick-film capacitors, such as the screened-on type, will find in the multilayer capacitor some aggressive competition.

Essentially full text of a paper presented at the Hybrid Microelectronics Symposium, Chicago, Ill., Oct. 28–30, 1968.

Donald W. Hamer (SM) received the B.S. degree in ceramic engineering from the University of Illinois in 1945, the M.B.A. degree from the University of Chicago in 1958, and the B.S. degree in electrical engineering from Pennsylvania State University in 1968. In 1956, after spending ten years outside the electronics industry (in pottery manufacturing, Navy duty in Korea, and refractory manufacturing), he joined the Kemetal Division of Radio Industries, now a part of TRW's electronic components group. At Kemetal he did the original formulation and process development and set up the initial manufacturing capabilities for titanate ceramic dielectrics, and subsequently served as the division's first chief engineer. In 1961 he joined the Solar Manufacturing Corporation in Los Angeles. Mr. Hamer joined Erie Technological Products' Technical Materials Division as chief engineer in 1963, in which capacity he supervised the development of multilayer ceramic capacitors, alumina substrates, and fired-on precious metal conductors. In 1967 Erie established its Materials Research Laboratory and named him director. In 1968 he was appointed corporate director of research and development.

Telecommunications and electronics in Japan

In a small country such as Japan, where the most important natural resource is the talents of its people, "knowledge industry" technologies, especially telecommunications and electric data processing, can play a major role in the nation's economy

Koji Kobayashi *Nippon Electric Company, Ltd.*

Prior to World War II, Japanese technology was concentrated mainly on military production. However, since the end of the war, the country's industries, particularly the electronics industry, have grown rapidly as the result of a concerted effort to achieve a viable postwar economy. And an important contribution to this growth has been the adoption of modern management and control technologies developed mainly in the United States.

Modern science and technology in Japan began in the Meiji Restoration just 100 years ago. Up to that time, the level of Japan's science and technology, in the Tokugawa Era, was far lower than that of various Western countries. However, after the Meiji Restoration, Japan was able to achieve modernization by rapidly digesting Western material civilization. This ability is believed to have been fostered by the Japanese traditional culture for several hundred years. In particular, the policy of seclusion, which had been followed throughout the regime of the Tokugawa Shogunate for more than 300 years, resulted in the slow pace of the internationalization of Japan. On the other hand, Japan was successful in organizing the fields of politics, economy, education, etc., by its own efforts, without aid from other countries.

Before World War II, Japanese technology seemed to be inclined to military production. Some typical examples are the army's Zero fighter, the navy warships *Yamato* and *Musashi*, and various kinds of optical weapons, torpedos, etc. However, the level of those technologies common to the general industry remained rather low. Although there were some pioneering inventions and developments in the field of telecommunications and electronics during this period in Japan, these were rather sporadic and were not fully industrialized. Representative instances of such work include: radiotelegraph, M. Matsushiro (1897); radiotelephone, Torigata, Yokoyama, and Kitamura (1914); power-line carrier telephone, U. Torigata (1919); the Yagi–Uda antenna, Yagi and Uda (1926); the magnetron, K. Okabe (1927); the ferrite magnet, Kato and Takei (1930); the R-cut crystal oscillator, I. Koga (1932); the magnetic tape recorder, Nagai and Igarashi (1938); strong permanent magnets, such as the KS magnet by Honda and Takagi (1913), a new KS magnet by Honda, Masumoto, and Shirakawa (1936), the MK magnet by T. Mishima (1931), and the MT magnet by Mishima and Makino (1948); a picture transmission system, Niwa and M. Kobayashi (1928); and research on propagation of radio short waves, Namba and Maeda (beginning in 1933).

One of the most successful examples of such developments was the nonloaded cable-carrier system proposed in 1932 by Drs. S. Matsumae and N. Shinohara of the Ministry of Communications. The development originated in the movement to promote domestic technology that began about that time. In accomplishing this system, the Ministry received cooperation from both cable and equipment manufacturers. The cable manufacturers developed a nonloaded cable of minimum crosstalk and the equipment makers developed high-gain repeaters and multiplex terminals. The writer himself was engaged directly in the latter work as a young engineer. Thus, in the adoption of this system, a 3000-km 24-channel nonloaded cable line, probably one of the longest in the world at the time, was completed between Japan and Manchuria in 1939. Another pre-World War II invention was the very-high-frequency multiplex telephone system, which was developed jointly by Dr. Shigeru Yonezawa of the Ministry of Communications and Dr. N. Tanaka of the Nippon Electric Company; it was put into commercial use in 1939. This route spans 60 km, linking Japan's Main Island with Hokkaido

Island across the Straits of Tsugaru. With a six-channel telephone, this is believed to be the first multiplex radiotelephone system in commercial use in the world.

With the end of the war, every technology in Japanese industries had to be started afresh from the ashes of war damage. The most significant steps that Japanese industries, including the telecommunication industry, took for Japan's postwar rehabilitation were the modern management and control technologies that originally had been developed elsewhere, mainly in the United States.

In particular, quality control became important to Japanese industry in order to wipe out the bad connotation of the label "Made in Japan" that had spread throughout the world before the war. Further, the industries that lost such a big market as the military concentrated on creating a commercial market, and were successful. In the course of such efforts, they introduced new technologies from other countries, especially from the United States; at the same time, they tried to increase their own capability in research and development.

Development of telecommunications

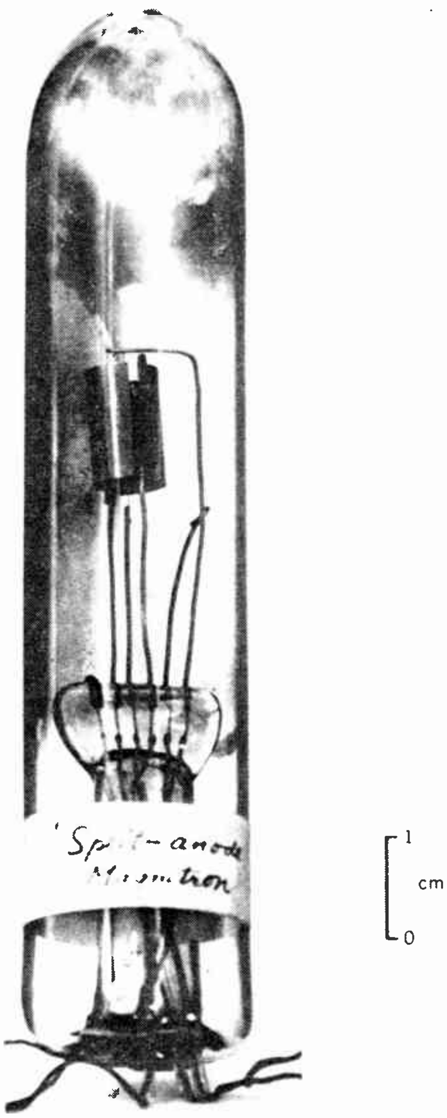
It is noteworthy that as a consequence of the Radio Law and the Broadcast Law, enacted on June 1, 1950, those radio waves thus far monopolized by the Japanese government (military) were opened to public use. Two years later, in 1952, the Nippon Telegraph and Telephone Public Corporation Law was established. These laws initiated the following:

1. The establishment of the Nippon Telegraph and Telephone Public Corporation (NTT) and the release of telecommunications from government operation.
2. The release of broadcasting from the monopoly of the Japan Broadcasting Corporation (NHK), following the start of commercial broadcasting.
3. The start of television broadcasting.
4. An increase in corporate communication by the use of radio communications in public and private enterprises, such as railway and electric power companies.

June 1 is celebrated by the Japanese nation each year as "Radio Wave Day."

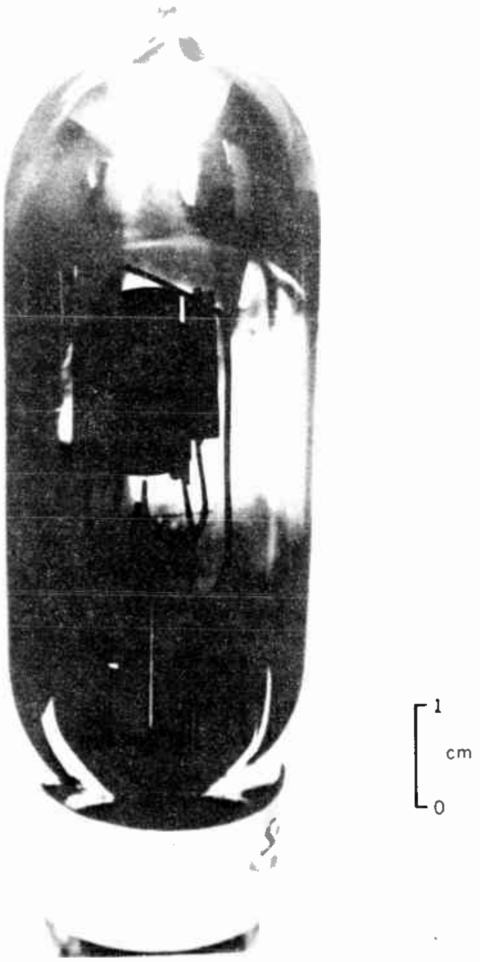
After the war, demands for national defense equipment decreased substantially. Instead, public enterprises requested from industry commercial facilities of the highest technology, performance, and quality. (Public enterprises include NTT, NHK, Japan National Railways, electric power companies, etc.) Industry cooperated, and thus, generally speaking, spent twice as much money as the government on research and development. In 1966, Japan's research expenditures amounted to about 490 billion yen (1360 million dollars) in total, of which some 60 percent or 300 billion yen (830 million dollars) was borne by private industries.

Dr. T. Kajii, the first president of NTT, shaped the basic policy that "all technological innovations are to be adopted to rapidly develop modern telecommunication network in Japan," and organized a development laboratory called the "Electrical Communication Laboratory" (ECL), the largest of its kind in Japan. In this way, a close cooperative tie-up with the telecommunications equipment industry was established. Around 1953, NTT started its four consecutive Five-Year Expansion Programs, aimed at realizing (1) a nationwide automatic telephone exchange network, (2) the clearing of waiting subscribers and attainment of 20 million subscribers, and (3) an entry into a new service age in 1972.



THE FIRST two-split-anode magnetron (1927).

This year makes the end of the third Five-Year Program and the start of the fourth one and, up to the present, many technological innovations have been adopted. The resultant curtailment or savings in facilities spending is estimated by NTT at more than 400 billion yen, or 1100 million dollars. As a beginning, telephone customers' equipment was restored. The no. 4 type of telephone set was put into commercial use in 1950, followed by the no. 600 type in 1961. In order to expand toll telephone lines, carrier transmission equipment was first miniaturized and then transistorized, resulting in an increase in capacity per bay from 48 channels to 300 channels; improvements are still going on. Standard coaxial cable systems and small-diameter coaxial cable systems were put into commercial use in 1956 and 1962 respectively. As a result, a 2700-channel system at 12 MHz was made standard. In addition, a 24-channel pulse-code-modulation



THE FIRST four-split-anode magnetron.

system was put into commercial use in 1965.

To realize nationwide television broadcasting and supermultiplex long-distance telephone, a microwave broadband relay system was developed. As the first part of this system, a Tokyo-Osaka route was opened in 1954. The NTT decided to use traveling-wave tubes, which contributed greatly to the rapid development of the microwave system. In addition, the over-the-horizon radio relay system was made possible by the invention of a high-sensitivity receiving system with phase-lock threshold extension by Drs. Morita and Ito of the Nippon Electric Company in 1957. Utilizing this new system, a television relay was realized between Okinawa and Japan. It was also adopted by both the U.S. Army and Air Force stationed in Japan. As far as improvements in equipment are concerned, the adoption of solid-state devices in place of electron tubes accelerated the improvement of performance, miniaturization, and cost reduction, and contributed toward export of microwave systems.

In order to cope with the increase in the number of subscribers and toll telephone lines, a new type of crossbar switching system was put into operation in 1958.

Tomorrow's technology

There are three major items being strenuously developed as technology for tomorrow:

1. A data communication system. A data system is being developed by NTT to meet the requirements of economic and industrial circles toward the use of on-line and real-time electronic data-processing systems.

2. An electronic switching system. In an attempt to be prepared for new services expected to come into bloom after 1972, a switching system is being developed under the leadership of NTT's Electrical Communication Laboratory, with the cooperation of manufacturing companies.

3. A satellite communication system. Satellites are expected to be useful not only for international communication but also for domestic communication in the future. Meanwhile, a Satellite Telecommunication by Automatic Routing (STAR) system has been developed through the collaboration of the Hughes Aircraft Company in the United States and the Nippon Electric Company. In 1965, public simulation experiments, utilizing a transponder similar to that used in the Early Bird, were carried out in Tokyo.

The hub of the broadcast enterprise in Japan is the Japan Broadcasting Corporation (NHK), a public corporation. It has enjoyed rapid progress since 1951, when commercial broadcasts started. As of 1967, radio stations in Japan totaled about 560, including some 150 commercial broadcast stations. In the same year, television stations numbered about 1500, including approximately 460 commercial television stations and about 900 nonattended satellite stations. This was some 15 years after the start of television broadcasting.

In order to enjoy radio or television programs in Japan, it is necessary for each household to make a contract with NHK and to pay certain receiving fees, as required by law. At the end of 1967, the number of contracted households and the rate of popularization were 20 million and 83 percent, respectively, for television alone; if radio were included, these figures would be 22.5 million and 93 percent. This number ranks second in the world, following the United States.

The Japanese nation, rich in delicacy, always requests radio and television broadcasts of high quality. In order to meet the requirements, NHK established a technical development laboratory and developed many useful broadcast technologies with close cooperation from manufacturing companies. Further, NHK organized a basic research laboratory in 1965 to carry out research work for the audio and video sciences. The successful television relay to the world of the Olympics in 1964 was attributable to the full collaboration of radio research laboratories of the Ministry of Posts and Telecommunications, Kokusai Denshin Denwa Kaisha (KDD), NHK, and the Nippon Electric Company.

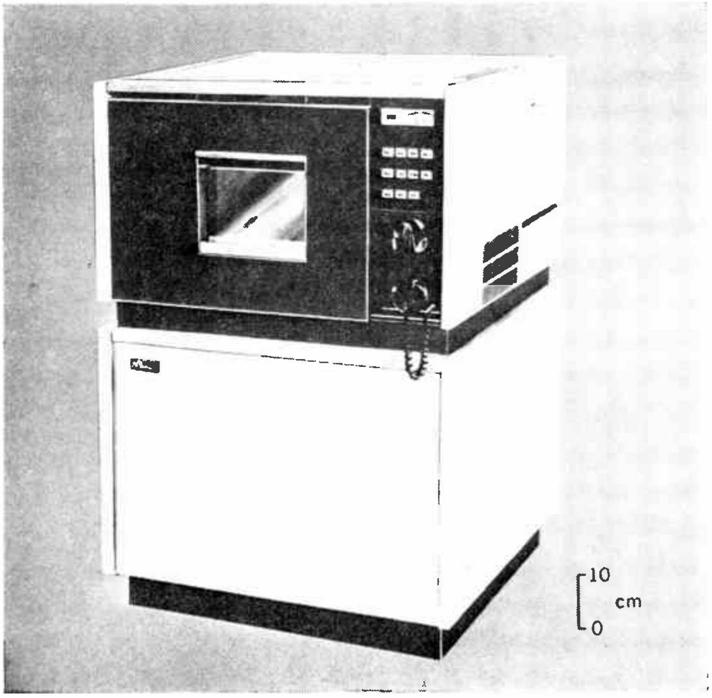
Such an increase in radio and television subscribers brought about a rapid expansion of the receiver manufacturing industry. In particular, the semiconductor industry was built up, utilizing technology from the United States, and created a new market with the so-called "handbag radio." This has accelerated the growth of Japan's entire electronics industry. The idea of creating such a new market, coupled with the efforts made to develop the new technologies required for its realization, is considered one of the important factors in the rapid growth of Japanese industry.

Railways, and computers

Railway traffic is also contributing importantly to the development of culture, economy, and industry in Japan. Physically, Japan is a narrow country supporting some 100 million people. Its population is concentrated mainly on the Pacific coast between Tokyo, Nagoya, and Osaka. According to Professor Tange, these city belts will emerge as a Tokaido megalopolis in 20 years. In Japan, railway service is considered most important as a means of mass transportation. As a result, Japan National Railways, as a public corporation, now operates 75 percent of the total operating mileage, thus playing a leading part in the transportation picture. With its excellent Railway Technical Research Laboratory, this company completed the New Tokaido Line in 1964, some 340 miles (540 km) from Tokyo to Osaka via Nagoya and Kyoto, and brightened the worldwide future of railway enterprise, which has been considered to be a declining industry. As a result, it has accelerated the development of automatic train and traffic control systems and many other electronic automation technologies. At present, construction is under way to extend the New Tokaido Line to western Japan.

The computer industry in Japan started quite a bit later than that in the United States. However, some scanty R & D was being conducted regarding computers of both relay and electron tube types. In the course of rapid economic growth, Japanese economic circles and industries felt it necessary to use electronic data-processing systems for both management and scientific applications, and desired the domestic development and production of computers for this purpose. The Japanese government established the Electronic Promotion Law in 1958 and endeavored to build up the computer industry domestically. In accordance with this, the industries organized the Japan Electronic Industry Development Association to help the local production of computers, automation equipment, and related components. An electronic computer, incorporating parametrons invented by Dr. E. Goto in 1954, was put into commercial use in 1957 as the first of such commercial electronic computers. In 1958, the first transistorized computer was put on the commercial market. It was modeled after the transistorized computer developed by the Electrotechnical Laboratory of the Ministry of International Trade and Industry between 1956 and 1957. Thus, the computer industry entered into production of a second-generation computer from the very beginning. According to statistics in 1967, the ratio of Japanese-made computers to imports was 47 to 53 percent in value and 72 to 28 percent in quantity.

FULLY automatic receiving equipment for use in telegraphic picture transmission.



The 'knowledge industry'

In Japan, the "knowledge industry" (a term used for the first time perhaps by Prof. F. Machlup of Princeton University) accounted for some 16 percent of the country's gross national product as of 1963. Telecommunications and electronics are the important technologies supporting this industry. Japan lacks in natural resources, but abounds in diligent and well-educated brainpower. Thus, the "knowledge industry" is believed to be one of the most suitable and promising industrial fields for Japan.

Koji Kobayashi (F), who has been president of the Nippon Electric Company, Ltd., Tokyo, since 1964, joined the company in 1929 following his graduation from Tokyo Imperial University with the bachelor's degree in electrical engineering. He received the Ph.D. degree from the same institution in 1939 and the honorary degree of doctor of laws from Monmouth College (Illinois) in 1968. Since 1963 he has also been president of the Nippon Aviotronics Company, which was organized as a joint venture of the Hughes Aircraft Company in the United States and the Nippon Electric Company; he is serving, in addition, as president of the New Nippon Electric Company, Ltd., and as chairman of the board of the Nippon Electric Industry Company, Ltd. The Chairman of the IEEE Tokyo Section, his other activities in the industry include the presidency of the Japan Electronic Industry Development Association and membership on the Electronic Industry Council, Ministry of International Trade and Industry. Dr. Kobayashi has received a number of awards for his work; he was presented the Purple Ribbon Medal by the Emperor of Japan in 1957 for his outstanding inventions and the Blue Ribbon Medal in 1964 for his contributions to the communication and electronics industry in Japan. In 1965 he received the Mainichi Press Prize for the development of industrial technology. The American Academy of Achievement selected him as recipient of its Golden Plate Award in 1966.

Electrical features of the Grand Coulee third power plant

With such unprecedented features as 600-MW generating units and 525-kV insulated cable, the world's largest single power plant will contribute 7200 MW to a power complex that can boast a total capacity of nearly 10 000 MW

John V. Baptist, Roy Y. Nitta

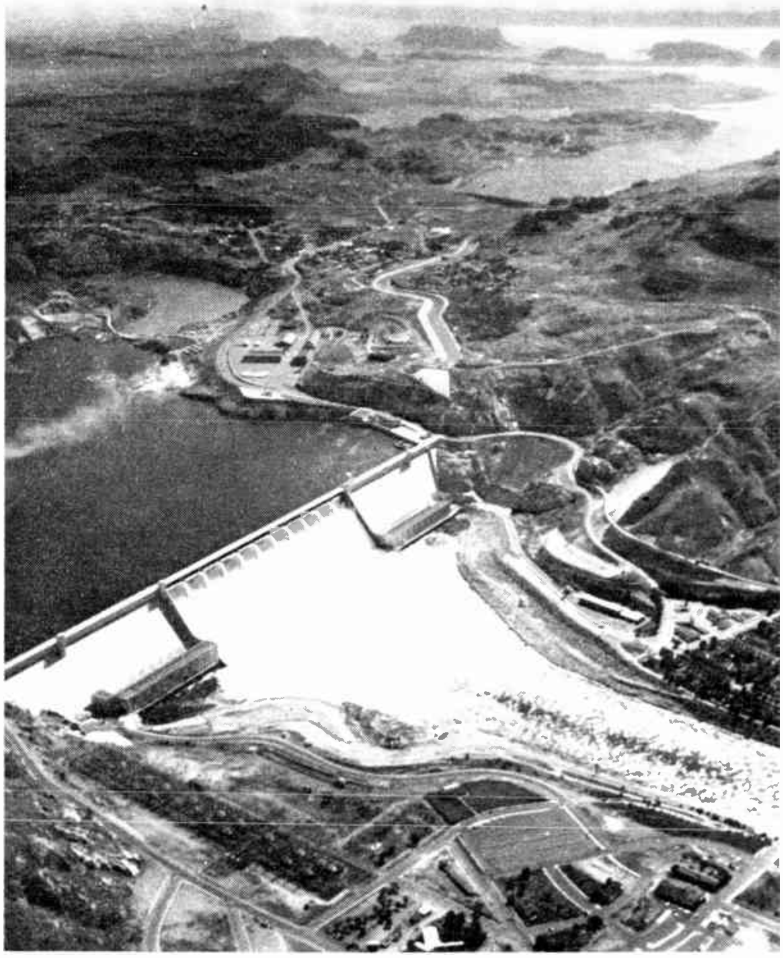
U.S. Department of the Interior

The Grand Coulee third power plant is a major feature of the Columbia Basin Project, sponsored by the U.S. Department of the Interior's Bureau of Reclamation. The planning and design of the facilities in this challenging undertaking are currently under way. It is anticipated that a number of "firsts" and records will be achieved in its construction, and that it will ultimately be the largest power plant in the world. This article describes the status of the electrical designs for the proposed 7200-MW Grand Coulee third power plant.

The Grand Coulee Dam is of the concrete gravity type, located on the Columbia River about 145 km west of Spokane, Wash. Completed in 1942, it is 1272 meters long, 168 meters high, and contains 10.585 million cubic yards (8.1 million m^3) of concrete. Franklin D. Roosevelt Lake, impounded by the dam, has a storage capacity of 9 562 000 acre-feet, and extends approximately 240 km to the Canadian border. The reservoir provides flood control and regulates the flow of the Columbia River to enable the Grand Coulee Dam and other downstream dams to produce more firm power than would otherwise be possible. It also provides over 25 percent of the developed power head on the Columbia in the United States, and has almost 30 percent of the total generating capability on the present United States Columbia River power system.

The present Grand Coulee generating facilities comprise two power plants—one on the right bank and one on the left bank of the river, against the toe of the dam (Fig. 1). Each plant contains nine 108-MW generators. In addition, there are three 10-MW station-service generators installed in the left power plant. The total nameplate rating of the two plants is now 1974 MW. The station-service units supply station power to the two

FIGURE 1. Aerial view of Grand Coulee Dam and present left and right power plants.



existing power plants and are also connected to the 115-kV system.

Each of the 108-MW-rated units in the two plants has been found to have a potential capability of approximately 130 000 kVA, and has been operated for sustained periods of load approaching this amount. Currently, there is a program to rewind the 108-MW units. The new armature windings are insulated with class B insulation, impregnated with an epoxy or polyester resin bonding agent. Three of the units in the left power plant have already been rewound, and each of these units now has a capability of up to 143 750 kVA at overload. Two additional armature windings are currently being manufactured that also have a "within 60°C machine temperature rise limits" capacity of 125 000 kVA and overload capability of up to 143 750 kVA. It is anticipated that the remainder of the units will be rewound at the rate of one a year. The ratings for these 13 windings, which have not yet been purchased, have not been firmly established at this time, since they may be influenced by the change in hydraulic conditions brought about by the addition of the third power plant and other downstream conditions. If the rewinding program were to continue on the same basis as those already installed, the "within 60°C machine temperature rise limits" capacity of the existing plants will approach 2280 MW.

The existing major switching facilities at the dam consist of a 115-kV switchyard on the left abutment of the dam and two 230-kV switchyards, one on each side of the river. The three yards make up a single operating installation, connected together with tie circuits. The 115-kV facilities are tied to the 230-kV circuits through the 108-MVA three-winding transformer bank for one of the generators on the left side of the river. Another generator is connected directly to the 115-kV yard. Power from these yards is delivered to transmission lines of the Bonneville Power Administration, the marketing agency for federally produced power in the Pacific North-

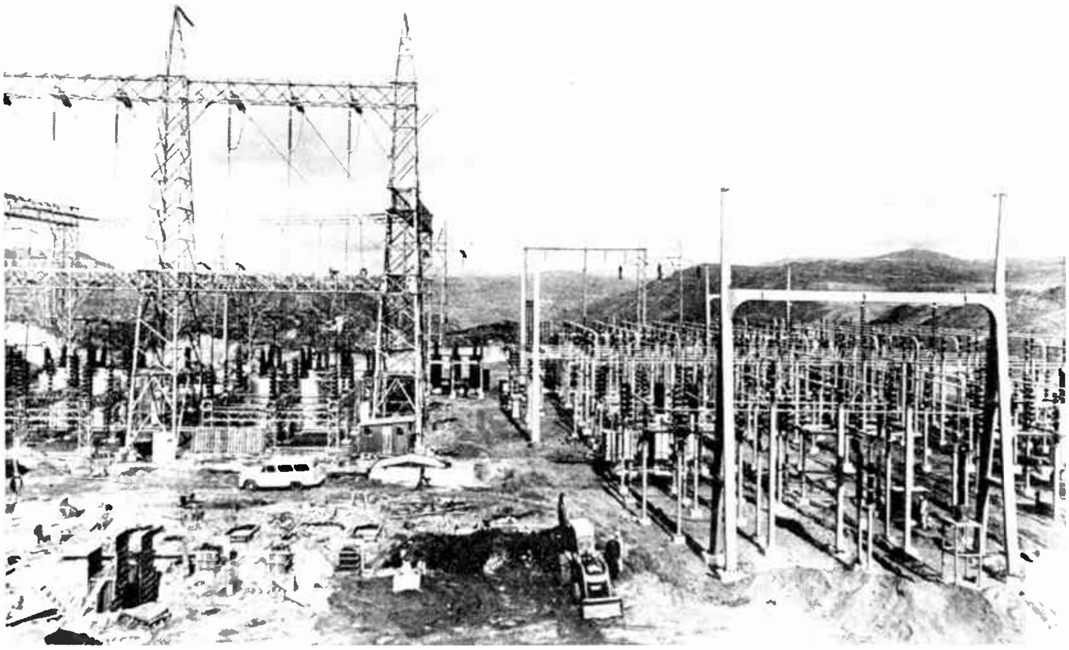
Modification of existing facilities

The site chosen for the third power plant will cause the removal of the existing 230-kV right switchyard. The first major contract, awarded in July 1967, is for modification of the existing 230-kV switchyards and is scheduled for completion in January 1970. The work includes the removal from service, by May 1969, of the right switchyard near the east abutment of the dam; combining the right and left switchyards into a consolidated 230-kV yard on the left abutment of the dam; and installing the bussing necessary to reconnect the existing units from their present unit system to essentially six groups of three units each, the exception being that units L7 and L8 will be paired and unit L9 will supply power to the 115-kV yard.

The existing overhead power-plant circuits will be removed and replaced with high-pressure oil, pipe-type cable circuits extending from the power plants to the 115- and 230-kV switchyards. The new cable circuits for the right power plant run from the transformer deck through an existing gallery in the dam into a newly excavated tunnel on the left bank to the consolidated switchyard. The circuits from the left power plant are also routed through the new tunnel.

The consolidated 230-kV switchyard will be of low-profile design (Figs. 2 and 3) using ground-mounted pipe busses, and will be built in stages to minimize outages. The consolidated yard will occupy an area of about 119

FIGURE 2. Existing left-switchyard structures (left) and low-profile structures of consolidated switchyard (right) under construction.



by 400 meters. It will be of the main and transfer bus type and includes 13 line bays, seven unit bays, three transfer bays, and two bus-sectionalizing bays. Space will be provided for six additional bays to accommodate future lines and possible braking resistors.

General plans for the third power plant

The Grand Coulee third power plant (Fig. 4) is to be constructed on the east bank of the Columbia River, immediately downstream from the dam. The power plant will be an indoor structure, and the building housing the initial (authorized) six-unit installation will be 326 meters long, 65 meters wide, and will rise 79 meters above its foundation. It will be of reinforced-concrete construction and will have a 38-meter-wide concrete superstructure, along with six unit bays, each 36 meters long, to house six 600-MW generating units, a 63-meter-long generator-erection bay at the north end, and a 45-meter-long turbine-erection bay at the opposite end of the building (near the existing right power plant).

Subject to future authorization by the Congress, the plant will be extended to house six additional 600-MW units for a total installation of twelve 600-MW units. The forebay for the power plant has been designed and will be constructed with sufficient width to permit the subsequent installation of the six additional units. The third plant will thus have an installed nameplate capacity of 7200 MW. Adding this total to the 1974-MW nameplate capacity of the existing plants, and another minimum of 200 MW of generation that will be installed as pump–generator facilities in the existing Grand Coulee pumping plant, the Grand Coulee power complex will have a total combined nameplate rating of 9374 MW—the largest single power concentration in the world. The possible ultimate total "within 60°C temperature rise" capability of 2280 MW of the existing plants was not

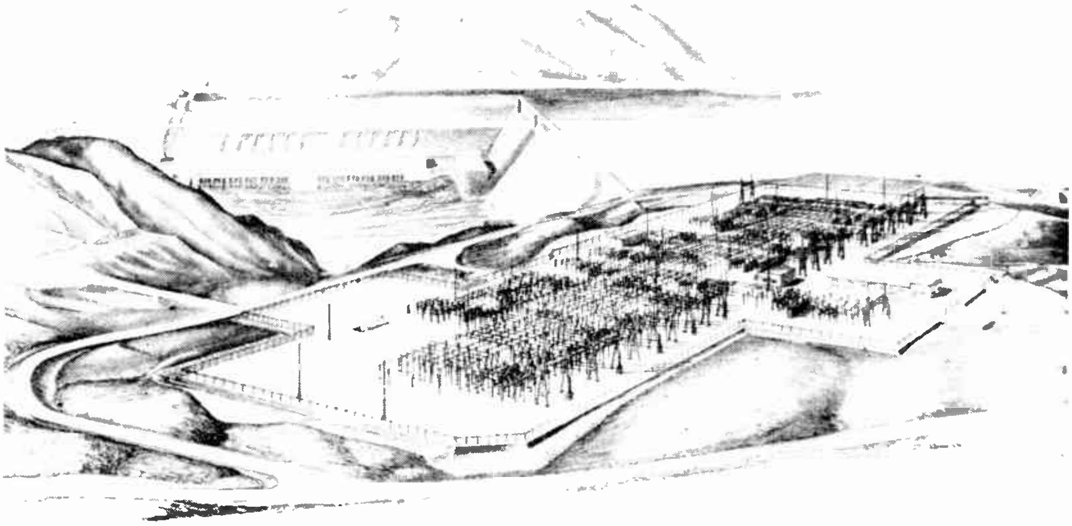
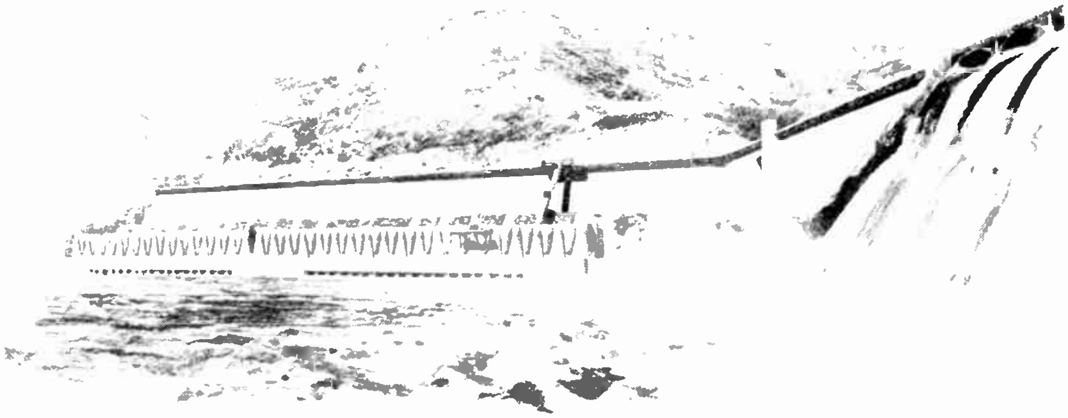


FIGURE 3. Artist's conception of consolidated switchyard.

FIGURE 4. Artist's conception of third power plant showing ultimate facility for 12 units.



used in this combined total because, as previously mentioned, higher tailwater conditions will preclude simultaneous maximum capacity operation. The additional pump–generator units to be installed in the existing pumping plant will be used to pump irrigation water during the late spring and summer months. During the nonirrigation season, water for power production will be pumped during low-power-demand periods and the flow reversed to generate power for peak demand periods. The generation capability of these units has not been determined and will be dependent on the canal capacity and lake elevations.

During the initial construction of the power plant, the substructure, intermediate structure, and superstructure will be completed for the six unit bays and the turbine- and generator-erection bays. Second-stage concrete will be placed for the first three units, and the turbines and generators will be installed for these three bays. Compacted backfill will be placed upstream to provide access to the right power plant, and the tailrace channel will be completed for the first six units of the third power plant. Installation of additional units will include placement of second-stage concrete and installation of the generating units.

Waterways and turbines. The forebay dam for the third power plant will be a concrete gravity structure, constructed as an approximate right-angle extension of the existing dam along its right abutment. The forebay will be an open cut to Franklin D. Roosevelt Lake, and the water from the forebay will be carried to the power plant turbines through 12-meter-diameter steel penstocks, which will be embedded in the forebay dam. A view of the forebay can be seen in Fig. 5.

The turbines and generators for the third power plant will be the largest hydroelectric units in the world. Each turbine will have a nameplate rating of 820 000 hp, 86.9-meter head, 72 r/min. The approximate weights and sizes of the turbines are given in Table I. To date, no procurement action has been initiated to secure a governing system for the turbines.

Power plant. The electrical arrangement will be a unit system consisting of one generator connected to a bank of step-up transformers through a breaker to the 500-kV transmission busses. The switchyard breaker arrangement is such that a fault will not interrupt more than one unit circuit to drop generation from the system. The overall switching arrangement has not been finalized, but present plans are for a breaker-and-a-half switching scheme.

Generators. Solicited offers for procurement of the first three generators were issued on January 19, 1968. Bidding on these first three units was restricted to domestic manufacturers. An offer was considered domestic if more than 50 percent of the cost was of United States origin. This did not exclude the possibility of manufacturers procuring components from abroad. Bidding rules for the remaining generating units have not yet been determined.

Procurement of the first three units was by the two-

I. Turbine statistics of the Grand Coulee third power plant

Spiral case	
Maximum inlet diameter	34′ 8″ (10.6 m)
Weight	530 tons (480 tonnes)
Runner	
Outside diameter	32′ 7″ (10 m)
Height	15′ 6″ (4.7 m)
Weight	550 tons (500 tonnes)
Shaft	
Diameter	8′ (2.4 m)
Weight	300 tons (272 tonnes)
Wicket gates	
Height	7′ 7″ (2.3 m)
Length	3′ 9″ (1.1 m)
Weight	10 tons (9 tonnes)
Head cover	
Outside diameter	39′ 2″ (12 m)

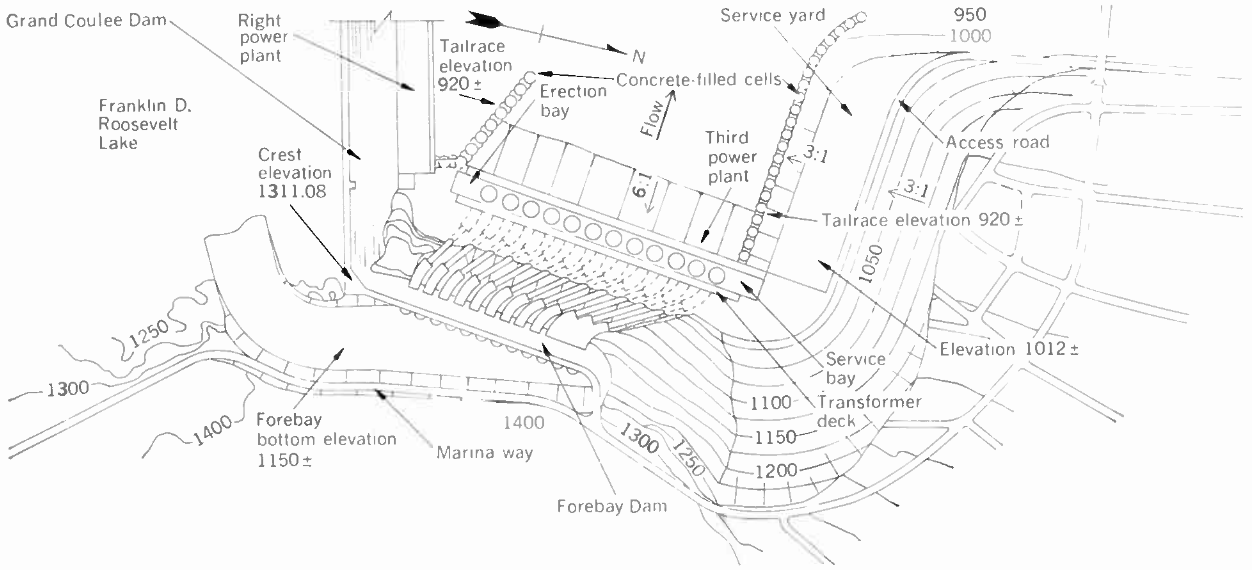


FIGURE 5. Proposed plan of Grand Coulee third power plant and forebay.

step formal advertising method, as prescribed by federal procurement regulations. Prospective offerers were given 78 calendar days to submit complete technical proposals without price. The offerers were encouraged to propose their own designs as alternatives to the requirements of the specifications, except for some characteristics, such as number of bearings, inertia, certain electrical features, and speed, which were fixed. Two offers were received and each offer was examined; where alternative designs were involved, they were evaluated for technical acceptability.

After the offers were evaluated and found to be acceptable from the technical standpoint, each offerer was requested to submit a price quotation for furnishing, installing, and testing the three generators. These price offers were taken on June 18, 1968.

For specification purposes, a conventionally water-air-cooled generator was assumed in order to establish a reference base for performance and standards. This assumption was made on the premise that the specified minimum WR^2 requirement would dictate a machine that was too large physically to gain benefit from direct water cooling. However, the two-step advertising procedure mentioned earlier permitted full consideration of the water-cooled type as well as any other innovative approaches an offerer might wish to propose.

Data that are furnished by the offerers and required on the "bidding schedule" of our solicitation are considered public information. Prior to awarding a contract, any other technical data or description of techniques considered by the offerers to be proprietary or confidential were so identified by the offerers and treated accordingly by the government. These conditions have been lifted by the successful offerer in order that the related design work be continued. All data received from the unsuccessful offerer have been returned.

II. Generator statistics of the third power plant

Rotor	
Diameter	60′ (18.3 m)
Height (including fans)	16′ 11″ (5.2 m)
Weight (without shaft)	1832.5 tons (1660 tonnes)
Stator	
Diameter	67′ (20.4 m)
Height	17′ (5.2 m)
Weight	825 tons (750 tonnes)
Air housing	
Diameter	83′ 10″ (25.5 m)
Total weight	3383.5 tons (3075 tonnes)
Shaft	
Diameter	8′ 4″ (2.5 m)
Wall thickness	8″ (20 cm)

III. Generator performance characteristics

kVA rating	615 385
Power factor	97.5%
Rated voltage	15 kV
Speed	72 r/min
Efficiency	98.37%
WR^2	2.542×10^9 lb·ft^2
X'_d	30%
Short-circuit ratio	Not less than 1.21
X''_q to X''_d ratio	Not more than 1.35
Excitation system response time	Not more than 0.05 second

There is a value engineering clause in the specifications that permits the contractor to propose alternative designs that would reduce the cost to the government. If such designs were approved by the Bureau, the contractor and the government would share the savings equally.

The generators will be very large physically; Table II lists dimensions and weights obtained from bid data and should reflect very nearly the final size of the units.

Since specifications were written assuming extensions of conventional designs, the requirements for insulation, cooling, bearings, and structural details are not greatly different, except for scale, from our previous specifications.

The generator windings will be insulated with class B insulation, and the coils will be impregnated with an epoxy or polyester resin bonding agent. Each machine will have its thrust bearing below the rotor and will have one guide bearing above and one below the rotor. The thrust bearing will be constructed so that the shoes can be handled and removed through the arms of the rotor spider and out of the top of the air housing. The bearing will be equipped with a high-pressure lubricating system to force oil between the bearing surfaces and the runner plate on start-up and shutdown. There are no special or unusual instrumentation or monitoring features being considered for the generators at this time. The stator frame will be recessed in the generator floor, and the top of the generator will be flush with the generator floor. The major generator characteristics are given in Table III.

Each generator will be equipped with an exciter and excitation control equipment consisting of static components. The system will be capable of reversing its output voltage in response to a control signal. Since the power for the excitation system will be supplied from the generator terminals, the system is not dependent upon an outside source of power. The static exciter will be of modular construction so that components can be readily removed for inspection or replacement. The circuitry for the exciter will incorporate speed-signal supplemental controls for damping system swings. It will also provide for prompt removal of voltage reference upon detection of a fault in the generator, transformer, or transformer circuits, and bring the generator field current to zero as rapidly as possible.

Generator busses. It is planned to utilize isolated-phase bus for the connections between the generator terminals and the transformers. No procurement action for these devices has been initiated at this time. However, preliminary plans include forced-air-cooled bus from the generator lead extension housing through the powerhouse wall to the transformer bank. The current-carrying capacity of the bus has not been selected at this time. Several factors are being considered in determining the rating of the bus.

In addition to carrying the rated voltage full-load current of the generators, the bus must be able to carry the 115 percent overload capability of the generator. Bus ratings are also influenced by conductor temperature rise limits. Under the conditions of 115 percent load and a 50°C-temperature-rise limit, the bus rating could be as high as 31 100 amperes.

Preliminary investigations show that the bus conductors may be 61 cm in diameter and the housing diameter about 107 cm in diameter. It is anticipated that the housing will be all-welded aluminum.

Transformers. The design of the powerhouse locates the unit step-up transformers on a reinforced-concrete structural deck upstream from the generator location. The transformers will either be forced-oil-water-cooled or forced-oil-air-cooled units, the selection being influenced by the total weight and space requirements.

The transformation from generator voltage to transmission voltage will be by three single-phase 236-MVA units. Each unit is estimated to be about 620 cm wide by 750 cm deep by 1610 cm high. Its total weight will be approximately 342 tonnes (378 tons).

The electrical characteristics of the transformers have not yet been established. A study is currently under way to ascertain these requirements.

Transformer circuits. Investigating means of connecting the transformers to the transmission busses has revealed a number of difficult obstacles in the case of overhead circuits. Therefore, the connections from the step-up transformer banks to the 500-kV transmission busses will consist of underground 525-kV cable circuits from the transformer bushings to a cable-spreading yard and overhead circuits from the spreading yard to the 500-kV switchyard.

The output of each generator will be carried from the power plant to the spreading yard in one three-phase cable circuit. Only one cable per phase will be used. Solicitation for the cable was issued June 11, 1968, and requires the successful offerer to furnish and install the cable.

Since this will be the first 525-kV cable installation in the world, Bureau engineers believe that a "furnish and install" contract is the best method to use to obtain the most satisfactory installation. Rigid offerer's qualifications are included in the specifications. Offers from only those manufacturers who have successfully manufactured high-voltage cable insulated for voltages higher than 275 kV will be accepted. To qualify as an offerer, a cable installer must have installed at least ten cable systems insulated for 115 kV or higher. Offers will be taken for either high-pressure or medium-pressure oil, pipe-type or self-contained-type cable systems. Several manufacturers, both from the U.S. and other countries, have indicated that they have an interest in submitting offers on this installation.

Current plans call for tunnels to be provided from the power plant to the dam to accommodate cables for the first six units. The cables will run from the potheads on the transformer deck down through the deck into the tunnels. The circuits will be run through existing galleries in the dam and through tunnels in the left bank to a cable-spreading area. There will be a difference of about 195 meters in elevation between the spreader-yard end and the powerhouse end. The estimated circuit length will be approximately 2140 meters.

If a pipe-type system is obtained, each single-phase cable will be paper-insulated, oil-impregnated, and about 10 cm in diameter. If a self-contained system is obtained, each cable will also be paper-insulated and about 12.7 cm in diameter. It is anticipated that corrugated aluminum sheathing will be used to contain the oil pressures in the self-contained cable. In either case, stainless-steel armoring will probably be used for mechanical strength.

Potheads for 525-kV insulation have been made for testing purposes. They are about 5.5 meters in height and 53 cm in diameter. One manufacturer has built high-strength porcelains for these potheads in one piece. Others have used fiberglass-reinforced tubes inside the porcelain housings to withstand the high pressures that are created.

Switchyard

Sites on the left side of the river are being investigated for construction of the switchyard to deliver the output of the third power plant to the Bonneville Power Administration transmission system. Switching schemes involving various voltages and transmission lines are being considered.

In addition to developing switchyard facilities for the 500 kV of the first six units, additional facilities for 700 kV and 1000 kV are being considered for the ultimate installation. This switchyard will be connected to the 230-kV consolidated switchyard by means of an autotransformer bank rated 1200 MVA. Provisions will be made for a future additional interconnection of these switchyards through another transformer bank of the same VA rating.

Handling facilities

The turbines and generators, because of their huge dimensions and weights, impose complex problems in the transportation of components to the site of construction, and in the handling of their assembled units within the powerhouse. In addition, the extremely tight schedule for construction of the plant will make it necessary for the turbine and generator contractors to work simultaneously, for at least part of the time, in a single generating unit pit.

In-place construction techniques for rotor erection were considered, but because of the crowded working conditions in the pit, and in order to reduce erection time, it was decided to require the assembly of the generator rotors in the generator-erection bay. This method introduced a need for a lifting device that could handle an assembled rotor (less shaft) that will weigh about 1660 tonnes. A hydraulic-hoist-type gantry, which will straddle the generators and operate over the entire length of the plant, will be provided for moving and setting the rotors. The gantry will also be capable of moving laterally a distance of 5 cm, which is sufficient for setting the rotor into the stator. The lifting capacity of the gantry will be established at 1800 tonnes.

In addition to the gantry, two 250-tonne overhead traveling cranes and one 45-tonne overhead traveling crane will be installed in the powerhouse for lifting purposes.

Control system

The control system for the third power plant can be separated into three categories. First, the primary control of the plant will operate from the dispatcher's room, which is located within the left power plant. This will be a centralized control system similar to the system now employed for the control of the left and right power plants. All control functions will be on a selective basis.

The second control of the third plant will be placed directly within the plant in a centrally located control console, a duplicate of the one in the left power plant.

Control functions from this point will also be on a selective basis. The third and final point of control will be from the control boards located in each unit bay. The control equipment located in the unit bays will control the individual units.

Closed-circuit television will be used for monitoring each unit control board; hence, there will be a television camera installed in each bay at each control board. One receiver will be installed in the central control area of the third power plant and one in the dispatcher's room in the left power plant. With each receiver, it will be necessary to select the unit control board that is to be viewed from these remote locations.

Control of the 500-kV switchyard may be performed locally at the switchyard. In addition to this local control, supervisory control equipment will also be installed to permit control of the switchyard from the third power plant central control area and the dispatcher's room in the left power plant.

Station-service system

The station-service power for the third power plant (six-unit plant) will be obtained by tapping the bus of two of the existing units in the right power plant. The generator voltage will be stepped down to 6900 volts and connected to a double-bus, double-breaker switchgear similar to those that have been installed in the left and right power plants.

In addition to the supply from the generator busses, one feeder will be installed to connect this switchgear to the station-service 6900-volt switchgear into the right power plant. All station power for the third power plant (six-unit plant) will be taken from this 6900-volt switchgear. In addition, plans call for installation of step-down substations at various locations in the plant to provide the necessary power for the station auxiliaries.

Schedule of construction

As of October 1, 1968, contracts totaling more than $55 million had been awarded for the initial phases of the third power plant site construction. As previously mentioned, work under the contract for modifications of the existing facilities were begun in July 1967, with completion scheduled in January 1970. A second contract was begun in December 1967, with scheduled completion by October 1969. The contract includes initial excavation of rock and common material required for construction of the forebay channel and forebay dam; construction of a road over the cofferdam; excavation and removal of 80 meters of concrete from the right end of the Grand Coulee Dam; excavation for the cable-spreader yard on the right abutment; and earthwork for about 3.2 km of haul and access roads. A third contract was awarded in July 1968, with scheduled completion in January 1970, covering the initial excavation of the power plant.

The prime contract for the construction of the forebay, forebay dam, the third power plant, tailrace, and a permanent visitors' center is currently scheduled to be awarded in September 1969. Power plant construction should be advanced to the stage that work on completion of the plant can begin by early summer of 1971.

The contract for the 1800-tonne gantry is expected to be awarded in the summer of 1970; erection of the gantry is expected to be completed by the summer of 1972.

The first 600-MW generating unit is scheduled to be placed in service by September 1973, and the other two units are to follow at six-month intervals.

Summary

The Grand Coulee third power plant will include many unprecedented features, such as 600-MW hydroelectric generating units, 525-kV insulated cable, and a 1800-tonne-capacity lifting gantry. Although construction work has already begun on this project, many details are still under intensive investigation and study.

Based on a paper presented at the IEEE Summer Power Meeting, Chicago, Ill., June 23–28, 1968.

John V. Baptist (SM) received the bachelor of science degree in electrical engineering from Kansas State University in 1935. After being employed by electric utility companies in eastern Kansas from 1935 to 1941, he became a junior engineer for the U.S. Bureau of Reclamation in 1941. Shortly thereafter, he entered the military service as an officer in charge of installation and maintenance of radar and communications equipment. Following his return to the Bureau of Reclamation in 1946, he was employed by the Electrical Branch, progressing from a specialist in the design of control, instrumentation, and relaying systems for hydroelectric power and pumping plants to his present position as chief of the Electrical Branch, Division of Design, in the Bureau's Office of Chief Engineer, Denver, Colo. The Electrical Branch performs all major electrical designs for the power plants, pumping plants, switchyards, substations, and transmission lines that are constructed by the Bureau. Mr. Baptist is a member of Phi Kappa Phi and Sigma Tau, and is on the U.S. Committee on Large Dams. He is a registered professional engineer in the State of Colorado.

Roy Y. Nitta is presently head of the U.S. Bureau of Reclamation's Power and Pumping Plant Design Section, Electrical Branch, Division of Design, at the Office of the Chief Engineer in Denver, Colo. Mr. Nitta received the bachelor of science degree in electrical engineering from the University of Denver in 1942. Following graduation, he served in the United States Army Signal Corps until 1945. Joining the Bureau of Reclamation in 1946 as a member of the Electrical Branch, he participated in the various aspects of designing power plants, pumping plants, switchyards, and substations. In particular, his experience has included control, instrumentation, relaying, and load dispatching systems; complete layouts and arrangements of plants, yards, and stations; equipment selection and application; and EHV cable systems. Mr. Nitta is a registered professional engineer in the State of Colorado.

Scanning the issues

Advance tables of contents
Translated journals
Special publications

Decision Theory. "Decisions! Decisions! Decisions!" That is the complaint one hears as often as one hears discussions of the weather and the state of the health—of individuals and of whole societies. But decisions are an essential ingredient of man's internal and intellectual climate. Hardly a moment goes by that one is not faced with the necessity of making decisions—small ones and large ones. Should one continue reading this paragraph? Should one change one's job or style or life? And so on. Usually one makes such decisions on the basis of a ready common sense watered with whatever information one can muster.

Nowadays, however, some aspects of common-sense decision making are being formalized through the use of mathematical and other tools under the name of "decision theory." The state of the art of this field, which one writer views as having progressed in the last few years from a theorist's toy to an important ally of the decision maker, is now treated in a special issue of the IEEE TRANSACTIONS ON SYSTEMS SCIENCE AND CYBERNETICS. The papers in the issue range, as guest editor Ronald Howard notes, from blue-sky theory to current practice and from the narrow corners of the field to its full breadth. Some significant applications, according to Howard, already include the desirability of kidney transplants, electric power system planning, and the development of policies for space exploration. No one can say, Howard says, when the limits of the decision analysis revolution will be reached, and he suggests that whether or not limits even exist depends more on man's psychology than on his intellect.

For those who have no grounding in formal decision theory, the issue provides an enlightening tutorial introduction to this area of systems science by D. Warner North of the Stanford Research Institute. North leavens his explication with a homely example that makes the piece entertaining as well as instructive. His viewpoint is taken from engineering rather than from science. He explains: We are not observing the way people make decisions; rather, we are participants in the decision-making process. Our concern is in actually making a decision, i.e., making a choice between alternative ways of allocating resources. We must assume that at least two distinct alternatives exist (or else there is no element of choice and, consequently, no problem). Alternatives are distinct only if they result in different (uncertain) rewards or penalties for the decision maker; once the decision has been made and the uncertainty resolved, the resource allocation can be changed only by incurring some penalty. To deal with such problems on a rational basis, North says, it is necessary to develop a theoretical structure for decision making that includes uncertainty.

In explicating the foundations of such a rational decision theory, North summarizes the necessary assumptions on an axiomatic basis. He identifies two dimensions to the way in which a decision problem is represented mathematically—value, by means of utility theory, and information, by means of probability theory. In this representation, he says, the large and complex problems of systems analysis become conceptually equivalent to simple problems in our daily life that are solved by common sense. In his description, North subordinates mathematical rigor in favor of a clear and readily readable exposition of fundamental concepts. However, for the purists he provides references to a sampling of what he calls the elegant and rigorous treatments of decision theory that are available.

Although decision theory is a way of formalizing common sense, North warns that it gives no magical formulas for correct decisions. In fact, he says, it forces the decision maker to rely more strongly than ever on his own preferences and judgments. But it does give the user a logical framework in which to work, a framework that is adaptable in principle to all decision problems, from the simplest to the most complex. As modern society continues to grow in size and complexity, North concludes, such a framework for decision making will become more and more necessary.

Among the many other interesting papers in this special issue, there are "The Foundations of Decision Analysis," by editor R. A. Howard, which, he says, provides a bridge between the formalism of decision theory and the practical requirements of analyzing decisions; in "Decision Analysis in a Corporation," by R. B. Wilson, the introduction is extended with special reference to the problems of group decision making. One whole group of papers deals with the encoding of subjective information so necessary in modeling the decision-making process. In this group, a number of papers deal with the subjective treatment of probability, and a number of the others take up the assessment of value and risk preference. Taken together, Howard writes, these papers on subjective topics provide a clear and comprehensive view of one of the unique aspects of decision analysis.

Still another group of papers deals with the frontiers of structural research, concerning the modeling question of decision analysis—namely, the problem of how to capture in a formal structure the many logical relationships involved in a decision.

The application of decision analysis is represented by two papers: "Decision Analysis for Product Development" by W. H. Cook is a detailed study of how decision analysis is used in a major corporate decision on which of several product development alternatives should be pursued; and "An Application of Decision Theory to a Medical Diagnosis–Treatment Problem" by A. S. Ginsburg and F. L. Offensend works on the most ancient of questions faced by the physician, "What course of action, in the form of diagnostic tests and/or treatments, should be taken?"

The breadth of potential applicability of decision theory specially recommends this first special issue on the subject to your attention. All you need to do is decide to get it and start reading. (*IEEE Trans. on Systems Science and Cybernetics*, September 1968.)

On Digital Filtering. In recognition of the stimulus that the development of large-scale integration (LSI) is likely to have on the spread of digital-filtering techniques, the IEEE TRANSACTIONS ON AUDIO AND ELECTROACOUSTICS has published a special issue on digital filters. To set the stage for the issue, the G-AE Concepts Subcommittee, under the chairmanship of C. M. Rader of the M.I.T. Lincoln Laboratory, has prepared a background paper that should be of wide interest.

As the committee notes, digital filtering is the process of spectrum shaping of signal waveforms using digital components as the basic elements for implementation. The technique of digital filtering is used extensively in computer simulation of analog filters, and, in fact, the origin of the technique can be traced to the earliest efforts to simulate analog signal processing schemes on general-purpose digital computers. For example, as editor Reg Kaenel writes, the development of many speech-processing systems, such as vocoders, has made great use of this type of simulation, thus contributing to the understanding of the technique.

In its paper, the committee observes that an increasingly large number of examples can be identified in which digital filtering appears to be more practical than analog processing for performing such operations as interpolation, extrapolation, smoothing, and spectral decomposition. This is especially true when the data to be operated upon are generated in digital form, as, for example, by a digital transducer. The unique advantages offered by digital techniques include the following: potentially small-sized integrated-circuit implementation; very predictable stable performance of arbitrarily high precision; absence of impedance-matching problems; no restrictions on the location of critical filter frequencies; greater flexibility because of the ease with which the filter response can be changed by varying the proper coefficients; and the intrinsic possibility of time-sharing major implementation segments. These advantages, the committee concludes, together with LSI, promise to make the digital-filtering technique eminently suitable for the exacting requirements of modern communications-oriented computing facilities. In fact, the rapid development of LSI has greatly increased the possibility of digital-filtering techniques. These trends, the committee predicts, promise to end the virtual monopoly of analog components for realizing real-time filters.

In its paper, the committee introduces the z-transform of a discrete-time series, considers the use of this transform in linear system analysis, and then discusses the relationship between discrete and continuous signals. Since the dozen papers that comprise the special issue cover digital filter implementations in one form or another, the background paper considers only an overview of such implementations.

In its conclusion, the committee notes that one possible advantage of digital filtering over analog filtering is the high precision attainable. The sources of error due to finite-length arithmetic, it is said, are somewhat understood and can be minimized in many cases by analytical considerations. As the size and cost of discrete components continue to decrease, the techniques of digital filtering will undoubtedly grow in importance, with the consequence that the need to know about these methods will grow apace. ("On Digital Filtering," *IEEE Trans. on Audio and Electroacoustics*, September 1968.)

Electronic Scanning Comes of Age. It has been said, writes Robert C. Hansen in a special issue of the PROCEEDINGS OF THE IEEE, that nothing can resist an idea whose time has come. And, he adds, it seems that this is the time of the phased array. So the phased array, which includes multiple-(fixed-) beam arrays as well as electronically scanned arrays, is made the subject of this special issue.

Phased-array theory, writes guest editor Hansen, goes back at least to Schelkunoff's work at Bell Labs during the '30s. During World War II there was a small flurry of experimental electronic scanning work, but very little implementation into operational systems. But then, in the '50s, a renaissance in phased arrays started under the sponsorship of ASD and the Air Force Cambridge Research Laboratories. The technology advanced slowly, however, especially in the component area, until large-system requirements drew the kind of support that was needed in R & D. The large-system requirements were, Hansen duly notes, primarily antiballistic missile defense, and secondarily satellite and aircraft tracking. By the time these efforts got going, the frustrating and disappointing period of working on an idea whose time had not yet come, as Hansen suggests, was past.

The basic reason, Hansen's description goes on, that phased arrays have advanced so slowly is simply that an array is a hard way to form, and move, a beam. There are relatively few applications where a fixed-beam reflector or array cannot be mechanically rotated, or where the rapid agility of an electronically steered beam is needed. Most phased-array applications require narrow beams and, hence, many elements.

The primary technical problem, then, is how to feed 10^3, 10^4, or even 10^5 elements. Space feeds are bulky and corporate feeds are complex and lossy.

The second major technical problem is the phase shift or time delay. The difficult tradeoffs involved here that Hansen identifies are the power handling, loss, number of units needed, and the number of bits in the case of the digital system.

A third major problem is that of beam forming and control, namely, finding the optimum configuration of feed network, phase shifters, analog or digital control units, and computer, given a set of value and cost criteria.

Last of the major problems identified by Hansen is that of mutual coupling and how to compensate for it. Inadequate compensation may result in shortened tube life; the least deleterious result will be poor system performance due to power loss and high noise temperature.

Add to these major problems a whole host of minor problems and it is easy to see, concludes Hansen, why every radar does not have a phased array.

And then, in a nice and amusing human note that comes somehow *sub voce* in Hansen's lucid account, he adds the following: Fortunately for those of us interested in electronic scanning, certain important tasks can *only* be accomplished with it!

The drive of the engineer to conquer new territories is clearly irrepressible.

The issue itself contains a wealth of papers, including some invited papers,

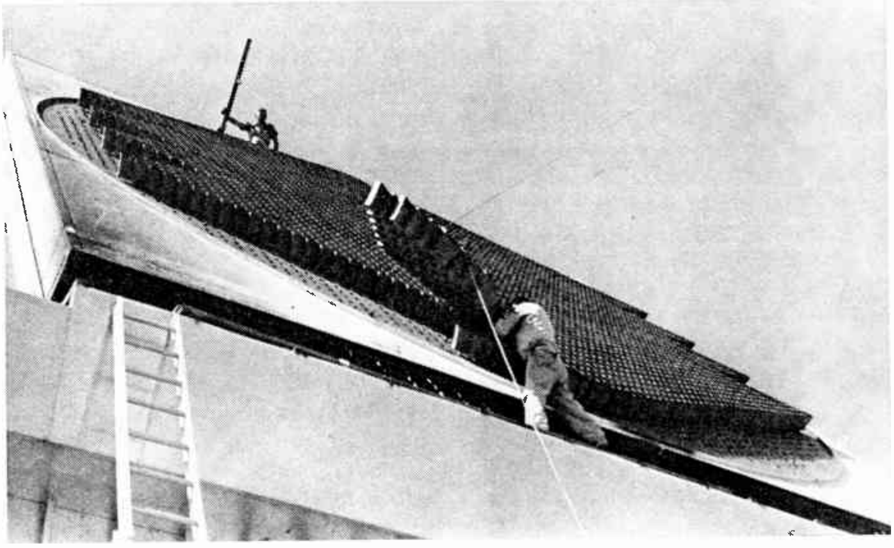


FIGURE 1. The installation of Hapdar, an operational phased-array radar.

that may have a reach beyond the electronic-scanning audience. We might single out a few from the five major categories into which the issue is divided: array theory; mutual coupling theory; components; array elements; array implementation.

The first, of course, should be the survey paper by Kahrilas, "Design of Electronic Scanning Radar Systems (ESRS)," which, according to him, is a "cursory" development of an ESR systems design procedure with an emphasis on the theoretical and practical design considerations unique to ESRs. There is, however, he says, no unique approach to the design of such systems.

There is one extract that is worth pulling out of Kahrilas, which should properly have been mentioned earlier. With electronic scanning, he writes, it is possible to obtain practically instantaneous slewing of an antenna beam to any position in a designated sector. The important thing is that electronic scanning is a method of positioning an electromagnetic beam in space by electronic means with the antenna aperture remaining fixed in space and no mechanics involved in the scanning process.

(Today's overburdened reader might well wish that there were some equivalent means of having the literature of his field scanned electronically for him.)

Also under the array theory category is an important theoretical contribution by Tseng and Cheng that produces Chebyshev-type optimization over a range of scan angles; a set of Chebyshev approximations by C. J. Drane; a treatment of the important problem of beam steering by Hatcher and Strauss; and others.

Editor Hansen observes that one of the current interesting problems in arrays is the appearance of "blind spots" or "element pattern nulls" at angles near the appearance of a grating lobe. This phenomenon surprised workers in the field and led them into interesting theoretical work, some of which is reported in this issue under the category of mutual coupling theory.

Of perhaps wider interest are the papers collected under the category of components. Two of these are invited papers, of which one, by Louis Stark and his colleagues at Hughes Aircraft, "Microwave Components for Wide-Band Phased Arrays," describes the design and performance of microwave components, and the other, "Review of Semiconductor Microwave Phase Shifters," by J. F. White, looks at the progress and potential of these discrete devices. These two papers, along with a paper by Hord *et al.*, cover the art in components that have allowed large phased arrays to be realized.

Four papers in the issue discuss array elements. In "A Systematic Approach to the Design of a Radiator Element for a Phased-Array Antenna," H. A. Wheeler comments on the problems peculiar to the design of a radiator element, describes some design techniques, and outlines a design procedure based on these techniques.

The last category of papers covers actual array systems. Hansen comments that it will be no surprise to the knowledgeable reader that information on some of the most interesting systems is classified—for example, the Nike-Zeus radars, the Sentinel radars, the carrier *Enterprise* array radar, and others. However, Peter J. Kahrilas provides a story on Hapdar, an operational phase-array radar, which was built to achieve maximum performance per dollar of cost. Figure 1 shows one aspect of the Hapdar installation. The article on "Phased Arrays for Radars" by T. C. Cheston, which appeared in the November 1968 issue of IEEE SPECTRUM, also contains a brief discussion of Hapdar.

All in all, then, it is fortunate for those engineers who wish to know about this idea that has come of age that a thorough and readily available literature awaits their perusal. (*Proceedings of the IEEE*, November 1968.)

Simple DC Transformer. From a South Australian engineer, there comes a description of a relatively simple and economical circuit that functions as a bilateral variable-ratio direct-voltage transformer. The circuit, which consists essentially of two thyristors, two diodes, and an inductance, plus trigger control circuits, could be used for such applications as regenerative braking and control of a dc shunt motor. When used for the control of dc machines, dynamic braking is accomplished automatically by feedback of power to the supply.

The author, Brian H. Smith, describes the operation of the circuit, which, he says, is a modification of the familiar pulse converter, and he discusses some of the factors that influence the choice of parameter values to achieve an optimum design for a particular application. He discusses the mechanism of commutation in the device, along with methods of selecting optimum values for the circuit elements involved, in a companion paper, "Commutation in Bilateral Variable-Ratio DC Pulse Converter."

As the author points out, and as practicing engineers know, there are many occasions where there is needed a simple, efficient bilateral variable-ratio device capable of converting from one direct voltage to another. This successfully tested design of a kind of dc transformer may belong in your file of circuit recipes. (Brian H. Smith, "A Simple Bilateral Variable-Ratio DC Pulse Converter," *IEEE Trans. on Industrial Electronics and Control Instrumentation*, November 1968.)

Nilo Lindgren