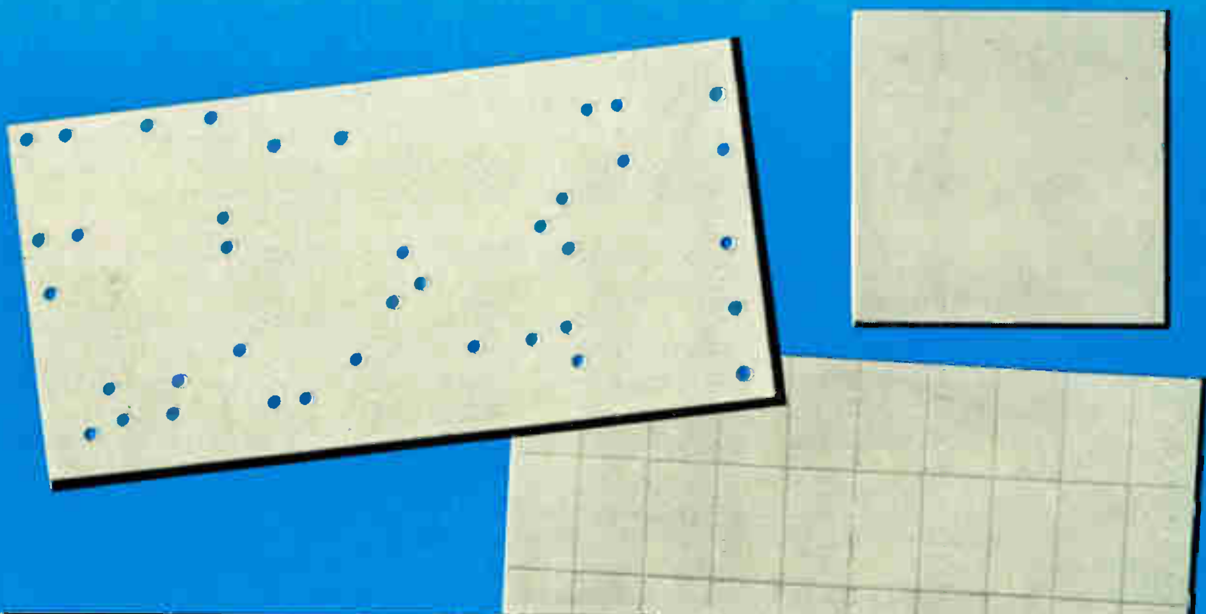
VOLUME 7 □ NUMBER 1 IEEE spectrum JANUARY 1970

IEEE spectrum

features

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

the 4-day substrate



4-DAY SUBSTRATES

Ceramic (10 or 25 mils thick)	Part Type	Max. No. Pieces	Max. Size (sq. in.)
ADS-96F (96% Al_2O_3, as-fired)	Sandard-Size Blanks*	2500	4
		1500	8
		500	16
	Custom-Cut Blanks**	1000	4
		500	8
		250	16
ADS-995 (99.5% Al_2O_3, as-fired)	Standard-Size Blanks*	1250	4
		750	8
		250	16
	Custom-Cut Blanks**	500	4
		250	8
		100	16

*Available in over 110 shapes and sizes (listed in DS-1402).
**Cut to your specifications.

7-DAY SUBSTRATES

Coors ships the following substrates within 7 calendar days of receipt of order.

Standard-Size Blanks — 15, 20, 25, or 30 mils thick; same quantities and wide choice of sizes as 4-day substrates.

Custom-Cut Blanks — 15, 20, 25, or 30 mils thick; same quantities as 4-day substrates.

Parts With Holes — maximum 250 pieces per order; maximum 50 holes per piece; minimum 15-mil hole diameters; 10, 25, or 35 mils thick.

Strate-Breaks® — maximum 500 pieces per order; maximum 100 segments per piece; 1 to 4 sq. in. total part size; 10 or 25 mils thick. (Sizes listed in DS-1402.)

Why wait 8 to 16 weeks (or more) for ceramic substrates?

Coors ships the parts shown at the left in 4 calendar days (or less).

And the cost is lower than you might think.

For details on part sizes, shapes, tolerances, and materials, send for data sheet 1402.

For prices, mail or phone specifications to:

Coors Porcelain Company
Tape Products Dept.
604 Ninth St.
Golden, Colorado 80401
(303) 279-6565, Ext. 404

Coors / CERAMICS

Circle No. 3 on Reader Service Card.

the cover

High-performance silicon devices today are being used widely in the synthesis of linear active networks, primarily because of their increasingly low cost. This month's cover schematically depicts a typical nonideal "op amp," in the inverting mode. For a discussion of these devices see the article beginning on page 42

departments

IEEE SPECTRUM is published monthly by The Institute of Electrical and Electronics Engineers, Inc. Headquarters address: 345 East 47 Street, New York, N.Y. 10017. Cable address: ITRIPLEE. Telephone: 212—752-6800. Published at 20th and Northampton Sts., Easton, Pa. 18042. **Change of address** can be made effective with any particular issue if received by IEEE Headquarters by the 10th of the preceding month. **Annual subscription:** IEEE members, first subscription $3.00 included in dues. Single copies $1.50. Nonmember subscriptions and additional member subscriptions available in either microfiche or printed form. Prices obtainable on request. **Editorial correspondence** should be addressed to IEEE SPECTRUM at IEEE Headquarters. **Advertising correspondence** should be addressed to IEEE Advertising Department, at IEEE Headquarters. Telephone: 212—752-6800.

OTHER IEEE PUBLICATIONS: IEEE also publishes the PROCEEDINGS OF THE IEEE, the IEEE STUDENT JOURNAL, and more than 30 Transactions for IEEE Groups with specialized interests within the electrical and electronics field. Manuscripts for any IEEE publication should be sent to the editor of that publication whose name and address are shown on page 17 of this issue. When in doubt, send the manuscript to E. K. Gannett, Editorial Services, at IEEE Headquarters, for forwarding to the correct party.

In June and July 1968, we discussed on this page, the design considerations of infrared heterodyne receivers and described a packaged receiver developed for NASA/GSFC laser communications experiments. This month, we discuss another outgrowth of this work. Contributors include W. Chiou, T. Flattau, B. Peyton, and R. Lange.

HIGH-SENSITIVITY INFRARED HETERODYNE RECEIVERS WITH GIGAHERTZ IF BANDWIDTH

Part III: Packaged 10.6-μm Heterodyne 1.2-GHz Bandwidth Receiver For CO_2 Laser Radar*

The receiver described in July 1968 demonstrated that sensitivity near the quantum-noise limit could be obtained in combination with large instantaneous bandwidth extending to the microwave region.** The receiver described here is based on the previously utilized technology.

In addition, the new receiver has the following characteristics:

- The frequency response was extended to a 1.2-GHz instantaneous bandwidth
- The sensitivity at the higher IF frequencies was significantly improved
- IF and signal-processing electronics were added

Packaged Receiver

The packaged receiver (see Figure 1) consists of four units:

- Mixer-amplifier package
- Mixer monitor and control unit
- IF and signal-processing unit
- Ion-vacuum pump control unit.

As in the NASA/GSFC receiver, the mixer-amplifier package uses a compensated Ge:Cu photoconductive mixer mounted in a cryogenic dewar. The infrared signal and local-oscillator radiation are incident on the mixer element through an Irtran 4 antireflection-coated window on the flat bottom face of the dewar. The output of the Ge:Cu element is connected to either the low-noise wideband IF preamplifier, with nominally 30 db of gain for heterodyne reception, or an audio-frequency amplifier for envelope detection and boresighting measurements (local oscillator tuned off).

A 75-foot cable connects the cryogenic mixer-amplifier package, which is to be mounted in the optical system, to the rack-mounted equipment. The mixer monitor and control unit supplies bias and amplifier voltages, and monitors mixer bias, LO power, and cryogenic liquid level.

The noise equivalent power (P_{min}) of the receiver, as measured in a homodyne setup with the signal beam intensity modulated at 1.4 kHz, was 8.1×10^{-20} W/Hz. This value is close to the theoretical value for the quantum-noise sensitivity limit.

$$2 h\nu B/\eta = 6.7 \times 10^{-20} \text{ W/Hz},$$

where

$h\nu$ = photon energy = 1.87×10^{-20} watt at 10.6 μm

h = Planck's constant

ν = frequency

η = quantum efficiency ≈ 0.56

The factor of 2 arises due to generation-recombination (g-r) noise.

Detailed measurements of the g-r noise spectrum up to 1.2 GHz yielded, when referenced to P_{min}, values for the noise equivalent power of less than 2.8×10^{-19} W/Hz over the entire 10 MHz to 1.2 GHz range.

* This work was supported by Rome Air Development Center, Griffiss AFB, N. Y.

IF and Signal-Processing Electronics

To achieve simple and effective acquisition and tracking capability, the following characteristics were provided in the IF and signal processing electronics:

- Rapid signal-acquisition capability to define the narrow spectral region occupied by the signal—needed due to large frequency uncertainty of the incoming signal
- Automatic frequency control (AFC)—for tracking the varying doppler-shifted signal
- Automatic gain control (AGC)—to keep receiver output signal level within a narrow range
- Range gate—to improve signal-to-noise energy ratio, and reduce spurious receiver response

The IF and signal-processing electronics (Figure 2) employ two frequency conversions. The first converts the incoming IF signal (10 MHz to 1.2 GHz) in a wideband mixer to a second narrower band IF amplifier. The frequency of the incoming IF signal is determined by monitoring the variable-frequency local oscillator (VCO) frequency. The second IF is filtered to remove undesired signals, such as those generated by the VCO. It is then down-converted to a 60-MHz third IF by mixing with the output of a crystal-controlled oscillator-multiplier. A 1-MHz bandwidth filter centered at 60 MHz restricts the bandwidth of this signal and determines the receiver noise bandwidth. The 60-MHz third IF is then amplified and detected. The processing of this signal includes pulsed AGC, AFC, and range gates.

The signal-processing electronics were designed to operate most effectively for the detection of radar signals with a pulse width of 10 microseconds and a pulse repetition rate of either 300 or 1000 pps. The AGC and AFC use a weighted averaging sampling technique. Operation of the unit was simplified by including indicators to monitor AGC, AFC, and signal acquisition. VCO control voltage, video output, 60-MHz output, and discriminator output can also be monitored.

Important characteristics are tabulated below:

Pulse Repetition Frequency (pps)	300	1000
Frequency search rate (MHz/s)	20	60
AGC response time (ms)	60	30
AFC response time (ms)	40	30
Frequency uncertainty with constant doppler (kHz)	30	50
Additional frequency error for 20 MHz/s tracking (kHz)	18	6
AGC dynamic range (dB)	30	30
Discriminator output	5.5 V/MHz into 2 kΩ	

This receiver will be used in a ground-based high-power CO_2 laser radar system at RADC.

We hope to report in the future on other infrared heterodyne receivers and electro-optical devices presently under development.

REFERENCE:

** F. Arams, R. Lange, B. Peyton, and E. Sard, "Packaged infrared 10.6 micron heterodyne 1-GHz bandwidth receiver," Digest, 1968 International Electron Device Meeting, Washington, D.C., October 23–25, 1968.

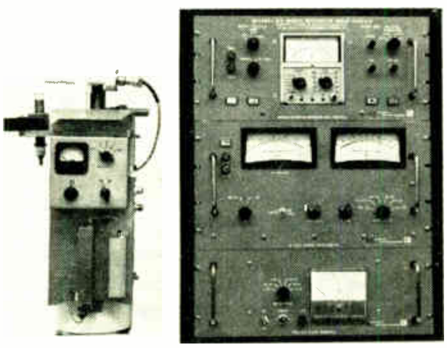


FIG. 1. INFRARED LASER RADAR HETERODYNE RECEIVER

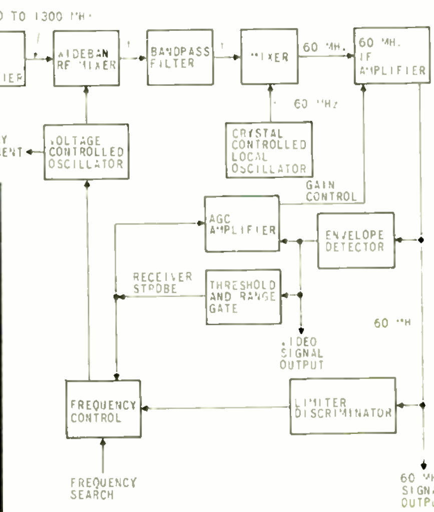


FIG. 2. BLOCK DIAGRAM

Forum

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

Group insurance

I would like to bring to the attention of the Directors of the IEEE and of the membership the comparison between the life insurance rates offered by the group plans of the Association of Professional Engineers of Ontario and the IEEE. In these calculations, a premium dividend of 35 percent was assumed for the IEEE plan, which was the highest ever given and is not guaranteed. However, even in this case, the IEEE plan's rates are very much higher, which prompts me to ask why this particular plan was accepted. I would be grateful if you would print this letter, along with a reply of the company involved, as many no doubt feel that a more open discussion of the IEEE Group Plans may be in order. The figures follow.

AGE	APEO	IEEE	% Higher IEEE
Under 30*	$ 1.20	$ 1.83	53
30–39	1.68	2.33	39
40–49	3.28	4.95	51
50–59	6.94	11.86	71
60–70	16.00	17.67	10

Rates quoted are per thousand, taking into account existing premiums. Ten-year averages are taken to equate both plans.

* 29 or less for IEEE plan.

The conclusion is thus that the IEEE members are paying about 50 percent more for their insurance than the APEO members.

Paul J. Bénéteau
Milan, Italy

The Group Life Plan available to professional engineers in Canada is indeed less expensive than the IEEE plan; exact comparisons are very complex, since the plans differ in many ways. Because of more favorable mortality rates in Canada, life insurance rates in the U.S. are normally higher than in Canada. The IEEE plan received favorable special mention by Consumers' Union in *Consumer Reports*, March 1967.

The Editor

Portable pensions

I would like to see the IEEE become more involved in an area that I think is a sore spot in our industry.

Because of the nature of the contracts in many companies for which an aerospace engineer works, the engineer may find himself changing companies more often than other individuals. It is true that many of us have accepted this as part of the game; however, I feel that there is something that could be done to make the situation more bearable. That something is a central organization for the control of retirement benefits and medical and life insurance. These benefits are presently offered by companies, but when an individual changes companies, either voluntarily or because of loss of contracts, he loses these benefits in most cases. To date in my career, I have encountered only a few people who have met all the requirements of a company plan such that they have retained these benefits.

If this situation were changed so these benefits were not lost in a move, I believe it would make the engineering profession, as presently practiced, much more desirable.

Dennis W. Beech
Newport Beach, Calif.

Engineers and politics

In this technological age it is vitally important that the technical community voice its opinions. This is especially important since most projects proposed to our legislators have testimony only from self-interested agencies. And can you imagine NIH proposing transfer of a billion dollars of its cancer research funds to another agency for oceanography research?

I propose that each year when the dues bill is sent to a member that the IEEE enclose a ballot listing ten major *technical* projects. The member would vote in five categories: (1) cut 50 percent, (2) cut 10 percent, (3) no change, (4) increase 10 percent, and (5) increase 50 percent. The results would be published each month in *IEEE Spectrum*.

The annual cost for processing (say $20 000) could probably be picked up by a foundation.

The only objection I can think of would be by those in the "corridors of power" who don't want to hear opinions conflicting with their own. Yet, as Churchill said, "Democracy is the worst of all political systems—except for all others which have been tried."

Stephan Konz
Manhattan, Kans.

I would like to express strong approval of the publication, if not the content, of the "controversial" items such as the Wald speech and the ABM articles. However, I would like to call attention to several controversies internal to the IEEE that might well be aired before we get too deeply involved in external matters.

The first is the reference direction for current. All military and other technician training defines current to have the direction of electron motion in a given field, whereas all professional education assumes a reference direction based on motion of a positive particle in the field. Further, "professional" current does not flow, derivation and usage to the contrary. At the very least, the IEEE should publicly acknowledge the situation and encourage debate in the hope of finding a solution. So far as I have been able to find out, no one with college training in electronics was consulted in the military "advance to the rear," whereas no representative of the military helped frame the conduction current definition in IEEE Standard No. 270, dated September 1966.

A second point for discussion is the vast surplus of engineers and the related need to terminate the undergraduate degree in engineering. Like law and medicine, engineering education might well become a professional course taken only after completion of more general undergraduate studies. Indeed, one is tempted to suggest that, if engineers had more exposure to history, political science,

sociology, and the techniques of open discussion necessary for full understanding of such subjects, President Willenbrock's letter in the September Forum might not have been necessary. There is little in the present engineering curriculum (which emphasizes memorization and regurgitation of facts, formulas, and techniques) to prepare someone to debate a point in politics, economics, or life in general. Thus it is not surprising that the IEEE receives a large number of letters demanding the elimination of controversial material that does not admit to solution by taking a Laplace transform.

Leo F. Goeller, Jr.
Hazlet, N.J.

Have we been "took"???

I mean, *IEEE Spectrum*, and the Editor, with much of IEEE as an organization!

The "issue" at M.I.T. is operation of the Lincoln and the Instrumentation Labs. (Both "engineering" labs, aren't they?) Does IEEE approve destruction of those labs??

More important: Does IEEE approve *FIRING* of Instrumentation Lab's Prof. Charles Stark Draper—glaring *ABOLITIONIST* of academic freedom at M.I.T.—or does "academic freedom" now only mean to please the "masters" in the Kremlin?

The "concerned" scientists, please note, were not identified (*Spectrum*, April 1969)! Was there any attempt to find out who they are, and what they really stand for? Shouldn't their anonymity be a tip-off that something was "sneaky"? Now, doesn't IEEE President Willenbrock, just about "close the trap" in Forum, *IEEE Spectrum*, September 1969???

Are Professor Draper, and any other M.I.T. Labs personnel, members of IEEE? If so, it seems to me the *least the Editor could have done* would have been to *ASK THEM* about the "concerned" scientists! While it is late, hopefully not too late, shouldn't IEEE members take a closer look? Usually I appreciate honest discussion of socio/economic/educational matters, but in the future, please beware that you don't swallow "hook and sinker" along with a "line"!

In any case, don't you think it high time for IEEE to rally to the support of Professor Draper, and the M.I.T. Laboratories????

L. H. Kimmons
Winston-Salem, N.C.

No, we haven't been "took." We saw a notice in Forum, in the April 1969 issue, that a group of M.I.T. faculty people—identified, actually, in the August issue—were organizing to produce change in the work carried on at M.I.T. Now we learn that they (in conjunction with student groups that regard them as conservatives) are getting results.

Dr. Draper is a Fellow of IEEE, and many of the Special Laboratories personnel at M.I.T. are members. One is Editor of *Spectrum*.

The group that submitted the April letter to Forum includes six professors of electrical engineering who are not members of IEEE, two who are members, and two who are Fellows.

The Editor

The "Proposed IEEE Policy Statement on the presentation of controversial material" seems entirely fair and reasonable, so much so that one might have taken for granted that a professional society of IEEE's caliber would hold essentially to this policy, with no statement necessary.

It is the professional responsibility of the engineer to inform himself about controversies arising from the activities of his profession: pollution, ABM, SST, safety in transportation, the "military–industrial complex," etc., etc. Indeed, the essence of a profession, as distinguished from a trade, is the ethic of responsibility to society at large. It would therefore seem a central function of a professional society such as IEEE to stimulate discussion among its members on matters where the concerns of the public and the profession overlap. *IEEE Spectrum* has been effectively carrying out this function.

Seventy years ago, "pure" scientists could with some justification claim the sanctity and immunity of the ivory tower. But *engineers* have never been in a position to thus withdraw from human concerns. There are no "pure" engineers. There are professional engineers who accept personal responsibility for their work, and there are others who, like those who practice the oldest "profession," make their services available on a strictly cash basis with no questions asked.

Carl Barus
Swarthmore College, Pa.

I am both amused and appalled at the letters appearing in the Forum of the November issue of *Spectrum*. Amused at the emotion and lack of objectivity on the part of those who pride themselves on being logical, using reason to solve problems. Appalled at the ignorance and stupidity of those who want to involve the profession, as a profession, in politics.

Engineers are respected and consulted on technical matters because they are thought to be objective, to forego emotion, to use reason to analyze problems, and to avoid personal bias in their evaluations. If engineers engage in politics as engineers, the credibility of the whole profession for objectivity is gone. A great disservice has been done engineers and scientists by those overbearing, conceited intellectuals who engaged in the ABM debate and went far beyond their technical competence in discussion. They did not confine themselves to technical opinions, but commented on military strategy, civil defense, psychology, sociology, economics, etc. They are now partisans, not objective experts.

Engineers who wish to participate in political decisions should speak as private citizens, not as engineers. This policy is not based on pompous puritanical attitudes or on fear of reprisals, but on the ground that experience of many people over many years has shown that if an expert, specialist, professional wishes to be heeded in his field of expertise he must, repeat *must*, avoid any taint of partisanship and emotional bias.

Keep our profession out of politics and keep politicians out of our profession.

R.M. Scott
Falls Church, Va.

Broadening the curriculum

No one can quarrel with the proposition that engineers, like other human beings, should be progressive, dynamic, forward looking, humane, moralistic, intellectual, etc. The controversy arises from the arbitrary definitions of these terms and their use as propaganda clubs by rigidly regimented bands of self-righteous, self-proclaimed intellectuals parrotting bogus "nonconformist" standardized slogans. By these arbitrary definitions, "progress" means a relapse into barbarism (jungle culture, mob rule, drug addiction, sexual debauchery); the right to dissent means the right to exclude, by violent means, all but "progressive" dogma from school curriculums; and "dialogue" means an incessant, monotonous, intolerant repetition of one-sided Orwellian propaganda monologues against the war in Vietnam, the Establishment, the Military–Industrial complex, the police, the doctors, the bankers, the sciences, the white-racist taxpayers, the space program, society, and numerous other assorted scapegoats. These attitudes are spawned by modern pseudo-scientific sociology whose total intellectual content is characterized by a resolute disregard for empirical evidence, elementary logic, and the quantitative evaluation of factors; and may be summarized by the following fallacies:

1. A nation can grovel, cringe, and bribe its way into international peace and domestic tranquility by unlimited craven capitulation to every foreign and domestic aggressor.

2. All crime is committed by saintly victims of society. It can be eliminated by heaping rewards and praise on the criminals, thus rehabilitating them into their inherent saintly state. The lives and property of an unlimited number of law-abiding taxpayers are expendable instruments of this expected rehabilitation but those of the criminals are sacred. Accordingly, only criminals and paupers have civil rights.

3. American culture, laws, courts, and schools are so despotically corrupt that justice can be achieved only by riotous mobs and our institutions must be completely replaced, immediately, by a Utopian society. The lessons of thousands of years of civilization must be jettisoned as worthless so that the new order can be designed in accordance with the miniscule experience and resourcefulness of the "intellectuals" who cannot cope with the problems and internal wars of their own universities, much less those of the whole world.

After ten years of intensive application of the above sociological quackery, we are in the midst of an epidemic, snowballing crime wave, expanding slums, substandard schools, bankrupt cities, unprecedented discontent, and a new dimension of race hatred ("polarization"), in the midst of prosperity. But the dogmatic sociologists refuse to learn from experience.

Is this the type of sociology that engineers should be forced to study? During the last decade, while the "humane," "enlightened" sociologists have created this mess, the scientific method has repeatedly demonstrated its marvelous progressive problem-solving capability. While it readily revises empirically discredited portions of old theory, it does not automatically totally jettison such old theories as Euclidean geometry, Newtonian mechanics, or Maxwell's wave theory in order to be "liberal." Sociologists should study the scientific methods. Engineering curriculums can withhold compulsory study of sociology from a crowded course until such study presents coherent approaches to problem-solving not instantly available to every television viewer. We urgently need problem solvers in both the technological as well as the sociological fields. Ranting revolutionaries are a glut on the market and the supply from the liberal arts schools is more than sufficient.

David Ginsberg
Washington, D.C.

Terminology booby trap

To avoid confusion between the Institution of Electronic and Radio Engineers (IERE) in Great Britain and the International Electric Research Exchange, which has the same initials, the latter organization will use its full name rather than initials. It hopes that this convention will be observed by others who have occasion to refer to it.

K. Masui
General Secretary
International Electric Research Exchange
Tokyo, Japan

I'm sure by now many of your readers have chuckled over your article on UHF transmission lines in the September issue. I don't believe UHF is the way to go for power transmission. However, the recent usage of the terms EHV and UHV seems to be a takeoff on the terms used to denote the radio-frequency spectrum. The radio-frequency terms have been well defined for years, but there seems to be a lack of definition of the new terms.

It seems quite possible to utilize almost all of the radio-frequency terms for electric potential if reasonable care is taken to define the range of potential for each term.

I don't believe the problem is immediate but to avoid future confusion it should be given some thought.

I enjoy reading *IEEE Spectrum*,

as I have for several years. However, I have one gripe. The physical size of *Spectrum* makes it difficult to hide away in a standard U.S. Army field pack!

Albert Helfrick
Fort Dix, N.J.

Fighting words

I am an engineer and am proud of the professional attitude and performance of the majority of my fellow engineers. I am not, by his definition, a colleague of William G. Naef, E.I.T. (November 1969, *IEEE Spectrum*). I have never attempted to become a state licensed engineer because:

1. I do not need a paper crutch to prove my professional capability.
2. I believe that the senior engineers in the organization are more qualified to judge my technical ability than is the state.
3. I have had better things to do with my time and money.

Mr. Naef, E.I.T., did a good job of categorizing engineers into five groups. My observations modify this arrangement slightly. I know several first-rate engineers who are not licensed. Other first-rate engineers of my acquaintance may be licensed without saying *anything* about it. I am sure that there must be many first-rate engineers who are licensed.

There is a sixth category that keeps trying to attach itself like a barnacle to the engineering profession. This category consists of those individuals who, lacking the creativity, logic, and broad technical knowledge to become an engineer, memorize enough formulas to pass a state licensing examination.

Philip S. Allen
Naval Missile Center
Point Mugu, Calif.

The letters by William Naef, E.I.T., were quite interesting. The subject of engineering licensing has long been of fascination to me.

The institution of "Professional" engineering is an archaic holdover from the horse and buggy days. The topics that must be studied for today's examination are not relevant to the world of today's engineer.

I know—I passed Parts I and II of the New York State licensing exam. When it came to take the remainder of the exam, I asked myself, "What for?" That portion of my mind that is devoted to engineering activities is too busy with today's problems to take time out to rehash some outdated college course I took seven years ago.

Engineering has had, as its reason for existence, an overriding concern for the problems of the present and future. To require an engineer to pass a ritualistic examination that is devoted to 50-year-old engineering problems is absurd.

I agree that there is a need for a criterion beyond a degree to judge engineers, especially where the public interest is at stake. The P.E. examinations, as presently instituted, are far from the answer. Is a certified P.E. knowledgeable in the hazards of electromagnetic radiation? Does a P.E. examination test a power engineer's ability to specify, build, and maintain a safe nuclear power plant?

Thanks to the standards of "professional" engineers in position of public responsibility, we have had the great Eastern blackout, air pollution, water pollution, and urban blight. Perhaps it is time to revitalize the engineering profession. I suggest that a good place to start is with an overhaul of the archaic and irrelevant institution of engineering licenses.

Alexander J. Kelly
Varian Solid State Div.
LEL Operation
Copiague, N.Y.

Memories

Haraden Pratt's reminiscences in November *Spectrum* reminded me.

At the close of World War II, I edited three volumes that made up Haraden's final report of his NDRC activities. At lunch one day at the Harvard Club in New York, I told him that I had never been able to see a finished bound copy of one of these volumes. The AEC clearance I had at the time was not good enough.

Haraden surprised me by saying that he too had been unable to see the finished report. Both of us had read my edited copy and I at least had read galley and page proof.

This amusing episode brought on further reminiscences.

Near the end of World War I, I was at Waite High School in Toledo getting students ready for the radio exams of the time to fit them for somewhat special work in the war.

During this war all amateur radio activity was taboo. All transmitters and receivers were sealed up tight.

Back of my chair at Waite was a door that led to a room with no other doors and no windows. One day one of my students brought in a basket full of wartime tubes. These tubes and the secret room presented more temptation than Willis Wing, with a similar job at Scott High, and I could stand. Winding up some big coils and hitching them to the tubes, we listened in on the long waves floating through the ether. As all of it was in war code we got nothing out of this adventure except code practice.

One night, however, we heard the German station at Nauen calling NAA, over and over. Finally we copied a message in plain English from the women of Germany asking the women of America to stop the war.

Now Willis and I were in a fix. We were listening illegally. Did NAA get the message? What should we do? We decided to sit tight and see what happened. Continuing to listen in, several nights later we heard NAA call Nauen and, sending blind, send a reply. It was from Mrs. Woodrow Wilson saying that the women of Germany knew how to stop the war—unconditional surrender.

We watched the newspapers but never saw anything that related.

As I told this to Haraden at the Harvard Club, he had a strange look on his face, and as he was a very serious person and as he had some authority because of his White House job, I thought I might be in trouble.

"Well Keith," he said, "I thought I was the only person who knew about those messages."

"Haraden," I said, "How did you know about them?"

Haraden said, "I was at NAA when they went through."

Keith Henney
Snowville, N.H.

Plurals

Armed no doubt with the latest datums on the subject, the IEEE Publications Board has decided to outlaw Latin and Greek plurals. Such nonsense is, I suppose, no more than one can expect from alumnuses of today's engineering schools, which in dismissing the classics as trivial are laying the basises for a further deterioration of our language.

Frederick Van Veen
Concord, Mass.

The Publications Board has not decided to outlaw Latin and Greek plurals. It has only decided to outlaw the Latin and Greek plurals that are commonly misused. That does not mean most of them.

The Editor

Moratorium on hydro plants?

"Environmental ugliness and the rape of nature can be forgiven when they result from poverty, but not when they occur in the midst of plenty and indeed are produced by wealth." [1] Unfortunately these words apply to too many of the activities in which we, the U.S. engineers of 1970, the professed custodians of our nation's resources, presently are involved.

Consider the ongoing process of bulldozing into oblivion our river valleys. I shall not refer here to such explicitly wasteful and unnecessary boondoggles as the Florida Barge Canal and the like. I shall confine myself to the destruction of our remaining free-flowing rivers for the primary purpose of taming their energy.

We have presently developed about half of the available U.S. hydro power. Not counting Alaska, about 50 GW of potential hydro power remains. The relevant question is: Can it be considered good engineering judgment to develop this power in view of the environmental costs involved?

I believe we must answer this question in the negative and shall try to give the reasons for this opinion.

The total 1969 U.S. installed electric capacity equals approximately 300 GW.[2] The growth has been exponential since the early thirties, with a doubling time of ten years, corresponding to an annual growth rate of about 8 percent.

Presently, hydro power accounts for only 17 percent of our total electric energy consumption. The bulk, or 82 percent, is obtained from fossil-fueled steam plants, with only about one percent coming from nuclear sources.

Hydro power therefore already constitutes a relatively unimportant part of the overall power picture. If we were to add the total remaining 50 MW of hydro power, we would thereby meet only two years of present power increase. This illuminates clearly the futility of the situation.

But let us look a few years into the future. After all, the proposed hydro plants will take some time to construct. By the year 2000 (and this is not further ahead into the mists of time that Pearl Harbor is looking backwards), we must increase the total generation to 2400 GW[3] in order to meet the expected demand. With the arrival on the scene of fast-breeder reactors, electric energy will probably outcompete other energy forms in the area of home heating, transportation, etc., and the figure 2400 GW is therefore very conservative. But using this conservative figure we find that the remaining hydro power would correspond to less than 2 percent of the total installed power by the year 2000.

Electric energy can be had only at an environmental cost. The cost associated with steam generation (air and thermal pollution) can be reduced to an arbitrary low level by sufficient capital outlay. The environmental cost associated with hydro development cannot be minimized by any amount of money. If present development plans are not canceled, then a U.S. child born today will never see a free-flowing U.S. river except those very few saved by our conservation groups and those utterly polluted ones used for barge traffic.

For a country possessing neither fossil fuel nor a nuclear technology, development of the hydro resources becomes a necessity. For the U.S., however, where alternative primary sources can be found, a similar development takes on the character of a needless waste. It is high time for a moratorium on hydro plants.

Olle I. Elgerd
Gainesville, Fla.

1. Dubos, R., *So Human an Animal.* New York: Scribner.
2. "Statistical Yearbook of the Electric Utility Industry," Edison Electric Institute, New York, N.Y.
3. Hicks, B. C., "The future of energy supply," *IEEE Spectrum*, vol. 3, pp. 82–84, Oct. 1966.

Environmental Design

The Interprofessional Commission on Environmental Design (ICED) is an organization intended to fill a long-existing and urgent need. Environment, by definition, denotes the aggregate of all conditions, circumstances, and influences affecting the development of man's physical surroundings.

Development and maintenance of the environment is an extremely involved matter, with legal, ethical, medical, technological, economic, social, cultural, esthetic, and political considerations. Technical phases are of major significance. It is of utmost importance that such theory and practice be widely applied.

It is not to be expected that any individual will attain the knowledge and skills requisite for best results in all of the many factors entering into our total environment. A leader in the design of esthetic and well-planned buildings is unlikely to be expert in the development of adequate water supply or electric power. One competent in the latter areas may not be qualified to plan an efficient transportation system or properly provide for recreational facilities. Teamwork to blend the capabilities of the professional design groups is essential. ICED exists to strengthen and identify this teamwork.

ICED took tentative form on March 8, 1963, at an informal meeting of representatives of the American Institute of Architects, the American Institute of Planners, the American Society of Civil Engineers, and the American Society of Landscape Architects. The governing bodies of the four participating organizations officially approved the formation of ICED and later acted to include the Consulting Engineers Council and the National Society of Professional Engineers as constituent societies. Thus, ICED now consists of representatives from six national organizations.

The concept on which ICED is based is expressed in the preamble of a resolution adopted at the organization meeting. This sets forth the precepts that:

"The responsibility for conceiving, designing, and producing an optimum environment represents a challenge and an opportunity for the design professions and more specifically for the national professional societies which represent the design professions. . .

"The achievement of the desired goals in environmental design requires the highest degree of collaboration between these national professional societies on such matters as education, public affairs, ethical practice, and similar subject. . ."

Membership consists of two representatives from each society. One of the two must be an officer or member of the board of his society—preferably the president or president-elect—at the time of his appointment. The other always will be the chief staff executive of the society. Terms of office are established at the pleasure of each appointing society, with the intent that they be not less than two years and a preference for three.

Meetings are scheduled in March and September of each year. At tne September meeting members elect a chairman from among their own number to serve for the following year. The secretary always will be the staff executive of the chairman's society.

ICED will not, in its own name, release any communication to represent a position on legislative or other matters. It will endeavor to keep each constituent society informed as to significant legislation, or governmental action, leaving each to act independently as it may determine.

ICED affords a framework for the voluntary development of the interdisciplinary trust and understanding that will engender the most effective coordination of the design professions in their common goal—the advancement of the public interest through improved environment. Thus, as each profession serves its best function, duplication and conflict will be minimized, as will the concomitant "jurisdictional" differences and controversies that sometimes arise. Where such problems do occur, ICED will encourage their rational resolution without jeopardy to the public image of all the disciplines represented.

All indications are that in the future, even more than now, there will be need for well-balanced teams of professional people properly educated and qualified to handle the many different problems of design for total physical environment. ICED is dedicated to the task of furthering the fulfillment of that need.

Paul Robbins
Washington, D.C.

CHANGE OF ADDRESS

Change of address can be made effective with any particular issue if received by IEEE Headquarters by the 10th of the preceding month. Please attach magazine address label here, print your new address below. Be certain that you include your IEEE member number and new ZIP Code. Mail this form to: Coding Department, The Institute of Electrical and Electronics Engineers, Inc., 345 E. 47 St., New York, N.Y. 10017.

Name

IEEE Member Number

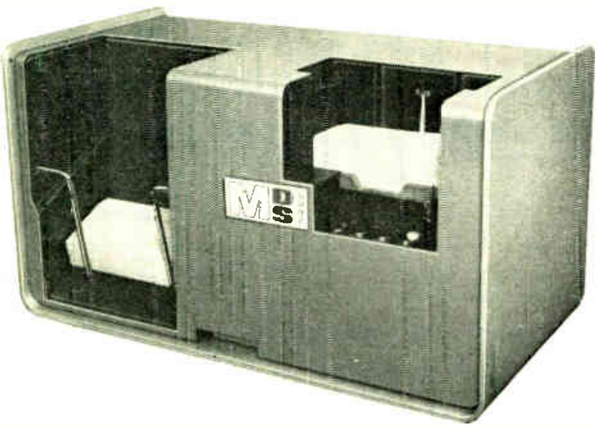
Address

City State Zip Code

Many of our best customers were Card Reader "skeptics"... until they tried the 6002.

MDS 6002 CARD READER

Specify the speed you prefer . . . 225 or 400 cards/minute . . . no extra cost. The 6002 requires only 2 sq. ft. of space. Ideal for peripheral use with small computers . . . or in remote terminals and data communications systems.



The 6002 is the "runt" in the MDS Card Reader line . . . 12½" H x 12½" D x 23" W. But it packs a terrific wallop!

Although the 6002 is a low-cost unit, it contains features ordinarily found only in more expensive Card Readers.

Input and output hoppers each hold 500 cards, face up. Cards can be replenished or removed during operation. No adjustments necessary for different types of cards.

Picker/Reader Head is the same as the head used in larger MDS units. It incorporates the picker mechanism, a pneumatic picker throat, photoelectric read station and metering capstan in a single, precision-built assembly. This permits simple interface and flexible operation when combined with end-fed reading and pick-on-demand card processing.

Cards are driven through the read station by metering capstans which also provide 16 precise clock pulses for each column of card motion. These signals . . . and those from sensors which detect the leading and trailing edges of the card . . . enable highly accurate column registration (±0.0054"), as well as a check of known vs. measured card length.

MDS sealed units and modular design hold servicing to a minimum.

The 6002 is available with or without logic circuitry (without, it is designated as 6001) . . . and with options accommodating 51-column cards.

Ask about other MDS Card Readers.

Phone or write for literature. Your Man from MDS will gladly help you make your decision.

HELP FOR OEM DECISION-MAKERS

MOHAWK DATA SCIENCES CORPORATION
OEM MARKETING
Palisade Street, Herkimer, N.Y. 13350
Phone 315/866-6040

REGIONAL MANAGER OFFICES:
Orlando, Fla., Phone 305/841-3091 • Waltham, Mass., Phone 617/891-5870 • Des Plaines, Ill., Phone 312/298-4141 • Los Angeles, Calif., Phone 213/685-5165
BUFFERED TAPE UNITS • DIGITAL STRIP PRINTERS • LINE PRINTERS • CARD PUNCHES AND READERS • PAPER TAPE PUNCHES AND READERS

Circle No. 9 on Reader Service Card.

Spectral lines

Thoughts on being nonnational. Although good science and good engineering are the same on one side of a political boundary as they are on the other, scientific and engineering societies are usually organized to function within some politically defined region—typically a state or country. Apart from organizations such as the International Scientific Radio Union (URSI), which function internationally but are small and exclusive, IEEE is in a class by itself; to it, national boundaries are recognized only for their helpfulness in delineating organizational units called Regions and Sections, and even these in general are not the same as countries. It is on this basis that the Institute claims to be nonnational.

Is the claim justified? The national engineering societies usually accept foreigners as members. Are we really different, or is the nonnational label just a device for allowing an ordinary society to expand its membership artificially, in a kind of imperialism?

All of my experiences within IEEE indicate that we really are different; the nonnational character is a reality, not a pose. That engineers outside the U.S. believe in IEEE's nonnational character is attested by the rate of growth of the non-U.S. membership, which exceeds the rate of increase of U.S. members.

The nonnational stance has advantages for all members. It provides a strong technical affiliation for electrical engineers in underindustrialized nations, and for those in the highly developed countries outside the U.S., it provides not only an open gateway to U.S. technology, but also an avenue for coming together across national boundaries, within Europe or elsewhere. For members in the U.S., the benefits include increased awareness of engineering developments in other countries, and increased opportunity for international contacts on a personal basis. The vigor of the organization enhances the status of electrical engineering everywhere.

The coin has another side, however. The merger that produced the IEEE dissolved the U.S. national society of electrical engineers. The architects of the new institute were fully aware, therefore, that it has to be flexible enough to cover the national interests of engineers in the U.S. For example, it needs to contribute advice to the U.S. Government on the use and conservation of the radio spectrum, and it needs to play a part in the Engineers' Council for Professional Development. In many other countries there are national societies that take care of the national problems in electrical and electronics engineering, but in some there are not, and to these nations the proper geographical unit of the IEEE can be of use in many of the same ways.

In recognition of these facts, Policy Statement No. 22 of the Board of Directors states, "Appropriate IEEE organizational units may, upon request, appoint representatives to national agencies of a technical nature in any country, subject to approval of the Executive Committee. Appointees will represent IEEE only insofar as its activity in that country is concerned." The same Policy Statement asserts unequivocally, "It is the policy of IEEE to cooperate and not to compete with existing national societies."

The first of these excerpts suggests the possibility that different units of IEEE, in different national or geographical areas, may adopt differing positions on the same question. Only by recognizing this possibility can IEEE remain nonnational without renouncing functions that are essential to the well-being of engineers in some of the Regions. The possibility has been recognized explicitly by the Board of Directors. Consciously implied is the recognition that to be of maximum utility, IEEE must not walk away from a problem that concerns electrical engineers in only one nation, unless there is a national society through which the problem can be attacked.

In effect, Regions 1 through 6 comprise the U.S. national society for electrical engineering. If it is deemed advisable for such a society to interest itself in such questions as professional registration, or portable pensions that will win for electrical engineers in the U.S. the ability to change employers without sacrificing retirement income, the appropriate portion of IEEE can address itself to the problem. The same is true for other nations. Canada (Region 7), for example, has a regional committee of IEEE that is the voice of Canada's electrical engineers on questions of spectrum utilization.

Europe (Region 8) presents a quite different situation. There, national questions can be handled by existing national societies, and the chief usefulness of IEEE is its bringing together of engineers with similar interests but differing nationality.

Since it is not true that nationality has no role or significance in IEEE, Director Emeritus Dr. Goldsmith has asserted that "nonnational" is inadequate to describe the Institute. He suggests "transnational" as a more suitable descriptor. The suggestion is sound, and I hope that it gains favor, but choosing the word is less important than understanding the situation. *J. J. G. McCue*

Aerospace nomads: How professionals move

His age, his education, his seniority are all factors that help the engineering professional to decide when, or whether, to move to another job—or even to another geographical area

Richard P. Howell New Management Center

If it became necessary to locate a large, high-technology defense facility in X location, where would the technical professional work force be obtained? Attempts to answer this question, posed by the Assistant Director of Defense Research and Engineering (Engineering Management) to a Stanford Research Institute team, opened a Pandora's box. Data collected in pursuit of the answer enabled the team to develop several inferences bearing upon the movement of this technical, intellectual resource between geographical areas and between employers. After describing briefly the source and processing of the data, this article describes the team's findings concerning intergeographical and interemployer movement of technical professionals, in that order.

In their effort to answer the question as to the movement of scientists and engineers from area to area and from employer to employer, the Stanford Research Institute team collected their data in three sequential phases. Phase I, an exploratory phase, reached into the personnel files of missile-producing facilities located during the prior decade in Denver, Tucson, and Orlando. The following information was acquired on 5233 currently employed and terminated professionals: their previous location and name of employer, educational level, and length of time with present company. Additional information on the locations and names of new employees was acquired for the 2255 terminated employees included in the sample.

In Phase II, exact copies of application forms filled in by technical professionals at time of hire were the data source. Twenty-two Los Angeles and Boston facilities participated. The participating companies also provided the starting and present salaries of the professionals under study. (Names were removed from the application forms to protect the individual's privacy.) In this second phase, data were compiled on 30 163 currently employed professionals and on 3599 terminated individuals. These 22 cooperating facilities represented missile, aircraft, propellant, electronic, atomic, space, and basic science industries.

In Phase III, copies of application forms were obtained in the same manner as in Phase II. Four facilities in the Twin Cities (Minneapolis, St. Paul) provided 1966 application forms of engineers and scientists currently engaged principally in the computer and control fields. The sum total of the three phases amounted to data on more than 40 000 engineers and scientists.

Geographical movement

In Phase I, our work to determine the key to the source of technical, professional personnel consisted of examining four variables. Since common knowledge indicated that the professional work force was moving from one defense R&D contractor to another (as will be discussed under "project piston"), the first variable applied was the level of R&D contracts from the nine census divisions. Preliminary analysis indicated that certain preferred routes to employment were taken by professionals; therefore, the second variable included in the analysis was the geographical source of the existing population of an area. Since distance appeared an important variable, though with nonlinear effect, the third variable considered was the reciprocal of distance. Finally, as the last variable, the size of the relevant labor pool was assumed to be of considerable influence.

Multiple regression analysis was undertaken to compare the relative effects of these variables on the industry's professional-work-force movements. Two analyses were conducted in the first, exploratory phase. In one, the data on the four variables were correlated with the data obtained on the given destinations of salaried workers terminating from the three companies studied. In the second, the data on the four variables were correlated with the data obtained on the geographic sources of the retained salaried work forces of the three companies studied. The general form of the equation used in both

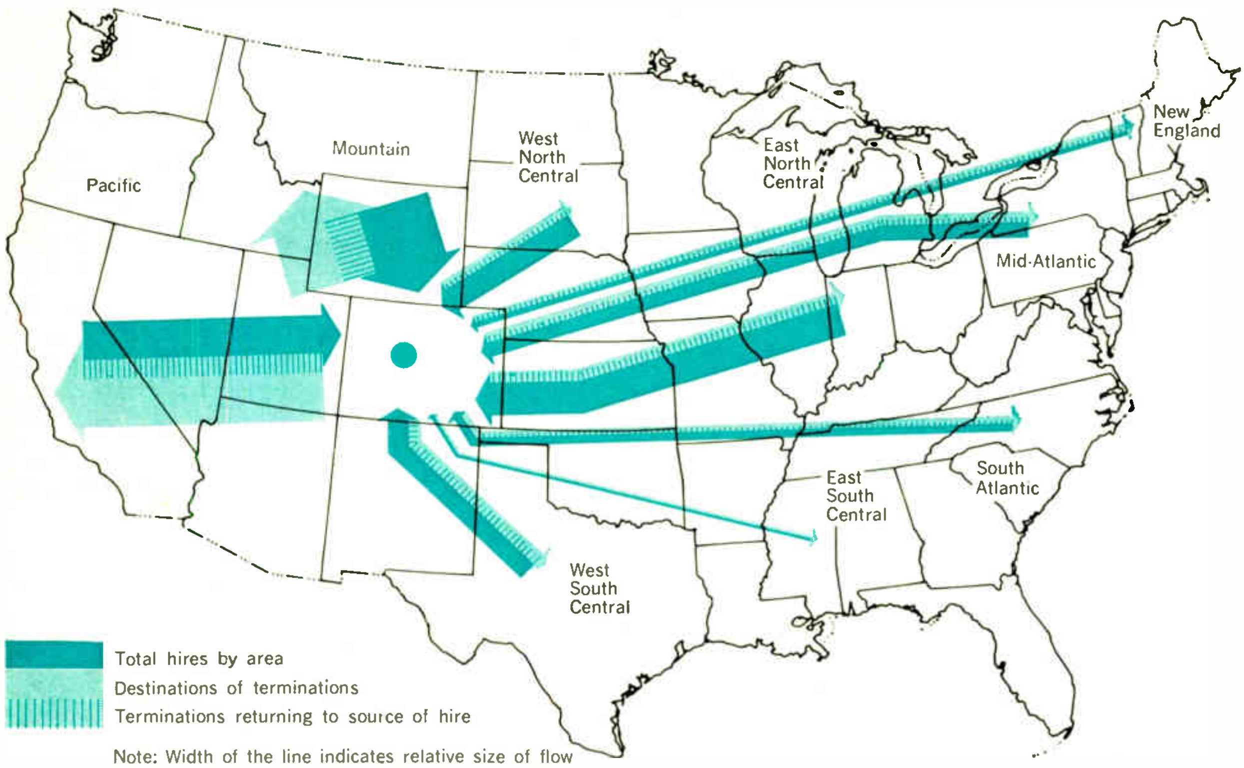


FIGURE 1. Flow of salaried workers, Company 1.

cases was

$$Y_1 = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4$$

where, in the first analysis,

Y_1 = proportion of workers that terminate from a company and give a particular census division as their destination

X_1 = proportion of defense R&D prime contracts (measured in dollars) that have been awarded to the census division

X_2 = proportion of out-migrants that have moved to the census division from the metropolitan area in which the company is located

X_3 = reciprocal of the distance between the company area and the census division (i.e., the larger the distance, the smaller the number giving the destination). This number was multiplied by 500 for convenience in computation

X_4 = size of the relevant labor pool in the census division

In the second analysis:

Y_2 = proportion of the company's retained work force hired from a particular census division

X_1 = reciprocal of X_1 as given above (i.e., the larger the contract volume in a census division, the smaller the flow of workers from there)

X_2 = proportion of the in-migrants to the metropolitan area of the company from the census division

X_3 = reciprocal of the distance between the company area and the census division (i.e., the larger the distance, the smaller the number giving the destination). This number was multiplied by 500 for convenience in computation

X_4 = size of the relevant labor pool in the census division

The equation from the first analysis is as follows:

$$Y_1 = -5.497 + 0.237X_1 + 0.102X_2 + 0.038X_3 + 0.0906X_4$$

The multiple regression correlation coefficient obtained was 0.975 and was significant at the 0.1 percent level. Of greater interest, however, were the partial correlation coefficients indicating the fits obtained by each of the variables. These were

$r_1 = 0.510$ (R & D contracts)
$r_2 = 0.950$ (out-migration)
$r_3 = 0.720\left(\frac{1}{\text{distance}}\right)$
$r_4 = 0.113$ (relevant labor pool)

These partial correlation coefficients indicate that although the movements of the salaried workers are associated with the relative volume of contract dollars, they are much more significantly correlated with general population movements. Reciprocal of distance is shown as having a fairly high correlation with destinations given.

A two-at-a-time multiple correlation using X_1 (R&D contracts) and X_3 (1/distance) resulted in a correlation coefficient of 0.920, indicating that the closer the demand for R&D labor, the higher the correlation with destinations given. A two-at-a-time correlation using X_2 (out-migration) and X_3 (1/distance) resulted in an even higher coefficient of 0.964; however, a bivariate correlation of X_2 and X_3 yielded a correlation coefficient of 0.830, indicating that these two variables may be interdependent. The best two-at-a-time correlation was that for X_1 (R&D contracts) and X_2 (out-migration), which yielded a correlation coefficient of 0.974.

The equation from the second analysis is as follows:

$$Y_2 = 0.660 + 0.013X_1 - 2.762X_2 + 0.211X_3 + 0.705X_4$$

The multiple regression coefficient obtained was 0.969. The partial correlation coefficients obtained with each of the variables were

$$r_1 = 0.389 \frac{1}{\text{(R\&D contracts)}}$$

$$r_2 = 0.915 \text{ (in-migration)}$$

$$r_3 = 0.707 \frac{1}{\text{(distance)}}$$

$$r_4 = 0.117 \text{ (relevant labor pool)}$$

Again, the movements of workers into the retained work force were very highly correlated with general population movements. The volume of R&D contracts showed a negative correlation, which indicates that the demand for workers in a census division to some extent may act to hold them there. Distance was again shown to have a fairly high correlation with movement of workers.

A two-at-a-time multiple correlation of X_1 (1/R&D contracts) and X_3 (1/distance) resulted in a correlation coefficient of 0.809, indicating that the more distant the originating area and the greater the demand in that area, the smaller the percentage of the work force obtained from there. The best two-at-a-time multiple correlation was that of X_2 (in-migration) and X_3 (1/distance), which yielded a correlation coefficient of 0.977; but again, a bivariate correlation of X_2 and X_3 yielded a correlation of 0.849, showing the interdependence of these two variables.

We inferred from this Phase I analysis that the technical professional work force moves, like other people, along migratory paths—some established a hundred years earlier. In contrast to the opinions held by many—ourselves included—the existence of large R&D expenditures or of pertinent labor pools does not serve as a good predictor as to "whence come engineers and scientists."

Figure 1 illustrates how Denver acts as a "staging" area where engineers, following the Overland Trail, pause before proceeding to the West. Note in this figure that when the technical professional is hired against a migratory stream, the tendency is to return at termination whence he came (i.e., the shaded area is as large as the solid area). Note, too, that most of those professionals who, upon departing, travel against the migratory path are returning to areas whence they came originally. In both of these instances we see a sort of "homing" behavior that is more pronounced when "home" is along a migratory stream.

Geographical movement theory substantiated. In the second phase, our inference developed in Phase I—that technical professionals move like the general population—was substantiated. Our much greater sample size made it practical to perform the analysis by 33 geographical subdivisions treated, instead of by only the nine census divisions analyzed in Phase I.

From the data covering engineers and scientists in 22 facilities in Los Angeles and Boston, the following equation concerning the source of engineers and scientists was derived:

$$Y = 1.03 + 0.68X$$

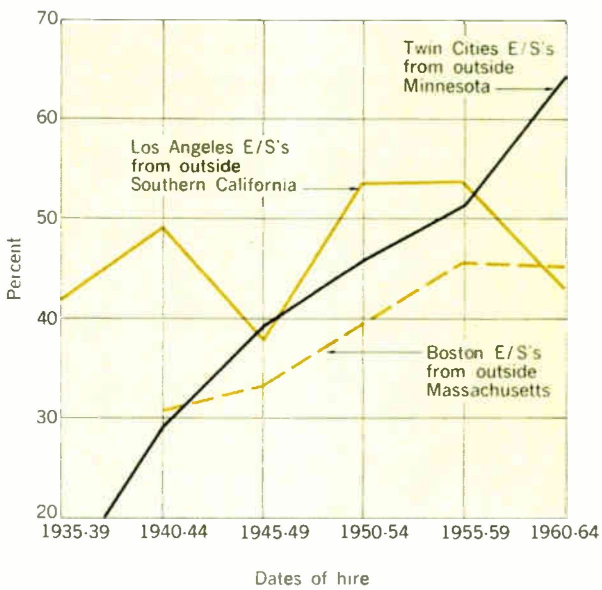
where X equals the percent in-migration of the general population and Y the percent engineers and scientists from each of the 33 areas analyzed. The correlation coefficient obtained was 0.99, even higher than that obtained from the Denver, Tucson, and Orlando study. For the data from the Los Angeles establishments, the correlation coefficient was 0.99; for the data from the Boston establishments, the coefficient was 0.98.

These results again point out that the movements of the technical professionals in the defense R&D industry are significantly correlated with the migration patterns of the general population. Moreover, these results from data on the geographic sources of hires, together with the data on destinations of terminating professionals in the Los Angeles and Boston work forces, suggest that there exist long-term geographic preferences of the defense R&D industry's professionals originating from any given area. Their preferences, we conclude, are closely related to the preferences of the general population.

In order to understand the patterns of flow that are relevant to the technical professionals in the defense R&D work forces of Los Angeles and Boston, the last three job moves of these professionals between areas were analyzed. The results of this analysis were combined with the results on sources and destinations.

Among the 29 253 professionals in the Los Angeles

FIGURE 2. Engineer/scientist hires from outside the home area (Twin Cities, Los Angeles, and Boston defense R&D establishments).



and Boston establishments who listed previous jobs, 49 217 job moves were obtained. (The 29 253 included 26 342 in Los Angeles with 44 614 moves, and 2911 in Boston with 4603 moves.) A study of these data showed that:

1. By far, the largest single fraction of moves among both the Los Angeles and Boston professionals occurred within the local areas. In Los Angeles there were 25 081 moves within the southern California area, representing 56.2 percent of the total moves analyzed. The next largest number of moves was from northern California to southern California (876), followed by 769 moves from Illinois to southern California. Similarly, in Boston there were 2315 intra-Massachusetts moves, representing 50.3 percent of the total analyzed moves. This was followed by 149 from metropolitan New York and 146 from Connecticut and Rhode Island.

2. Further reinforcing this intra-area pattern were the data showing that intra-area moves represented the largest single bloc of previous moves to, or within, the great majority of the 33 areas included in the job-move analysis. For example, of the 593 moves by Los Angeles professionals that involved Illinois, 412 were intra-Illinois. Analysis of the past job moves of the Boston professionals indicated a similar pattern, but to a lesser extent, primarily because of the relatively smaller sample.

3. The Boston professionals were drawn principally from the New England area and upstate New York. There has been also a definite stream up through the megalopolitan corridor from the Washington, D.C., area through Pennsylvania, New Jersey, New York City, and Connecticut to Boston. This stream is identified by the relative numbers of the interstate movements "on the way" to Boston, as well as by the direct moves from each of the states in the corridor to Boston.

4. There have been strong flows to southern California from the Middle Atlantic and North Central states, as well as from within the Western census region. The interstate movements "on the way" to Los Angeles are relatively smaller than those to Boston because of the smaller number of areas with intervening opportunities for the skills of the defense R & D industry.

5. There are strong indications of a "boomerang" or "homing" flow from Los Angeles and Boston to other areas and back again. Of the analyzed interarea moves of the Los Angeles professionals, approximately 19 percent had come from the Los Angeles area earlier and were returning. Similarly, with the Boston professionals, just under 18 percent of the interarea moves were from Boston with a consequent return.

Both Los Angeles and Boston appear to be termination points or "preferred" destinations for a number of streams of flow. The data on destinations given by terminated professionals, though limited, showed the local census division as the predominant destination in each case, the percentages being 73.7 percent and 63.3 percent, respectively.

In Phase III of our study, the conclusions reached earlier were again supported in an analysis of the source of engineers in Minneapolis and St. Paul high-technology companies. In this instance, the equation of best fit relating proportional source of technical professionals to source of general population was

$$Y = X$$

After expressing the usual caveat that correlation does not prove causation, we feel justified for the analysis of the three phases in making the following statement: "The proportional movement of the general population has proved successful in three instances in explaining at least 80 percent of the proportional movement of scientists and engineers at work in defense R & D. Although the proportional movement of the general population will change with the passage of time, such changes are slow to come about, and hence such general movement is useful in predicting the source and destination of engineers and scientists on the move."

Attainments of correlation coefficients of 0.90 and above are rare in matters involving human behavior. We believe, therefore, that a relationship has been well established. A generalized model fitting any and all cases, however, has not yet been constructed. It is my belief that, if a useful generalized model is attained, it will include a time variable that measures the maturity of the R & D complex.

Figure 2 shows, for example, how the well-established R & D complex will reach out for about two in five of its engineers and scientists, whereas a rapidly expanding, but relatively new, complex may have to import two

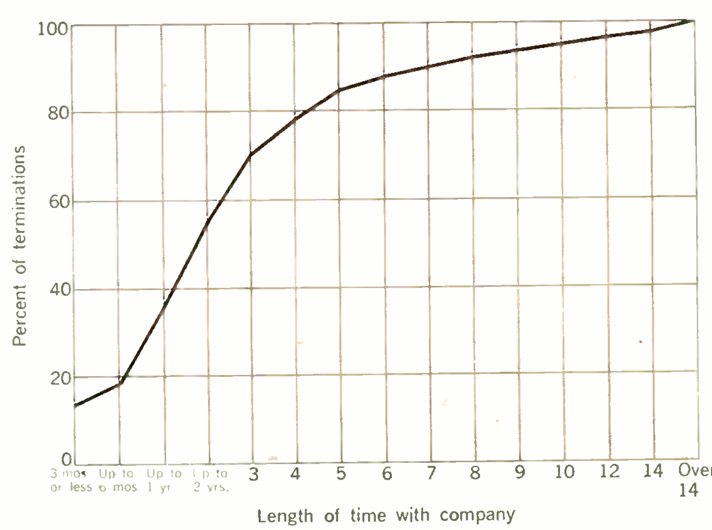
I. Turnover rates for selected defense R&D establishments*

Year	Establishment, percent†										
	A	B	C	D	E	F	G	H	I	J	K
1964	na	na	19	24	na	26	na	na	na	na	na
1963	19	20	17	39	12	25	10	14	6	28	26
1962	17	21	27	18	18	18	12	11	11	24	8
1961	18	25	22		19	28	11	10	19	14	10
1960	20	33	18			42				13	
1959	21	25	34			40				11	

* Turnover rate defined as: 100 × terminations ÷ $\frac{\text{Professionals beginning + ending}}{2}$

† na = not available

FIGURE 3. Cumulative terminations of engineers/scientists, by length of time with company (Los Angeles and Boston defense R&D establishments).



II. Institutional sources of professional personnel defense R&D companies (percent of work force)

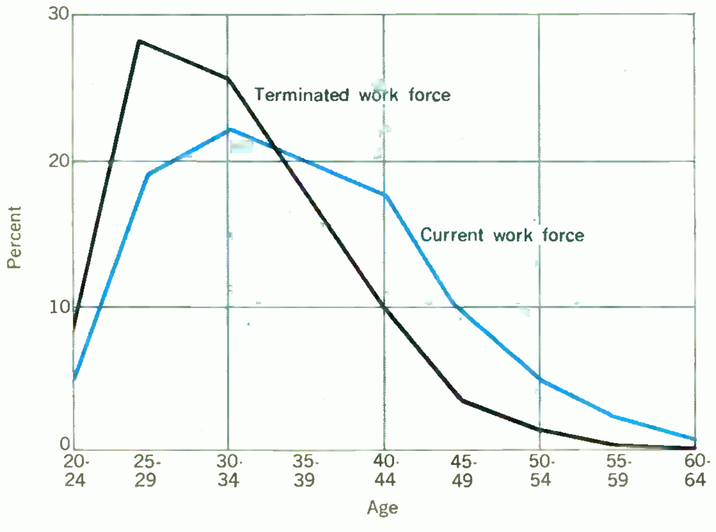
	Sources of Professionals			
	500 Top Defense R&D Contractors	Government and Military	Educational Institutions	Other
Phase I				
Company 1	50.0	11.0	6.6	32.3
Company 2	59.0	9.6	3.7	27.7
Company 3	57.0	8.3	7.4	27.4
Phase II				
Los Angeles/ Boston companies	42.9	3.5	26.1	27.5
Phase III				
Twin Cities companies	33.2	3.5	42.5	20.8

III. Institutional destination of professional personnel defense R&D companies* (percent of terminations)

	Destinations of Professionals			
	500 Top Defense R&D Contractors	Government and Military	Educational Institutions	Other
Phase I				
Company 1	60.7	6.6	10.5	22.1
Company 2	55.7	4.9	6.5	33.0
Phase II				
Los Angeles/ Boston companies	73.7	4.8	15.7	5.8

* Data not available or too limited for analysis of company 3, Phase I, and all companies, Phase III.

FIGURE 4. Age distribution of engineers/scientists: current work force and terminated work force (Los Angeles and Boston defense R&D establishments).



thirds or more of the professionals hired (Tucson, Denver, and Orlando found it necessary to reach out for over three fourths of their engineers and scientists, even after eight years).

Interinstitutional movement

A model describing the movement of professionals between establishments will almost certainly contain more variables than those models outlined in the foregoing. Among the interinstitutional movement variables focused upon in the three phases of the study were (1) establishment differences, (2) length of time with present employer, (3) type of previous establishment employing the professional, (4) professional's age and level of education, (5) route taken in applying for work, (6) geographica- location of employer, (7) procurement policies of governl ment agencies, and (8) general state of the R&D economy. These variables will be discussed in order and some general inferences will be drawn.

Personnel practices and policies vary from one company to another to the extent that turnover rates will also vary markedly. Although turnover records were not always maintained separately for professionals, Table I exemplifies how the rate varies from one company to the next and within a company from one period to another. Higher turnover rates entail significant costs. Although estimates of costs will vary, a reasonable estimate would place the cost of recruiting (moving when necessary) and orienting a professional at somewhat over $2000 and possibly as much as $5000.

Termination aggravated by mismatches between professionals and job content, conflict of personalities, dissatisfactions with community or environment, etc., occur for the most part before two years have passed on the job. Over 70 percent of those terminating from the Los Angeles–Boston work forces had seniority of less than three years (see Fig. 3). If we assume that an equal number of professionals are in each seniority group employed by the companies participating in Phase II, then the highest rate of termination, as shown in Fig. 3, will occur in the first year of employment, with subsequent decreases in rate occurring at the end of three and five years.

In tabulating the source of a company's technical professional work force by type of previous employer, we found it also related to the maturity of the complex in which the present employer is located. As demonstrated in Table II, facilities in an emerging complex will hire the largest single percentage of their professionals directly from educational institutions whereas those located in a mature complex will acquire their largest single percentage of professionals from other defense R&D companies. But even in the mature complex, about one in four technical professionals will be hired directly from a college or university.

By contrast, the companies studied in Phase I, each of which was organized on a crash basis, found it necessary to recruit most of their professionals from other defense R&D facilities, and to hire a greater proportion from government and nondefense R&D establishments. Table II shows that the number recruited for crash programs in Phase I from colleges and universities was small. Evidently, newly forming companies requiring many professionals find it advantageous to hire those with experience.

Table III identifies the institutional destination of

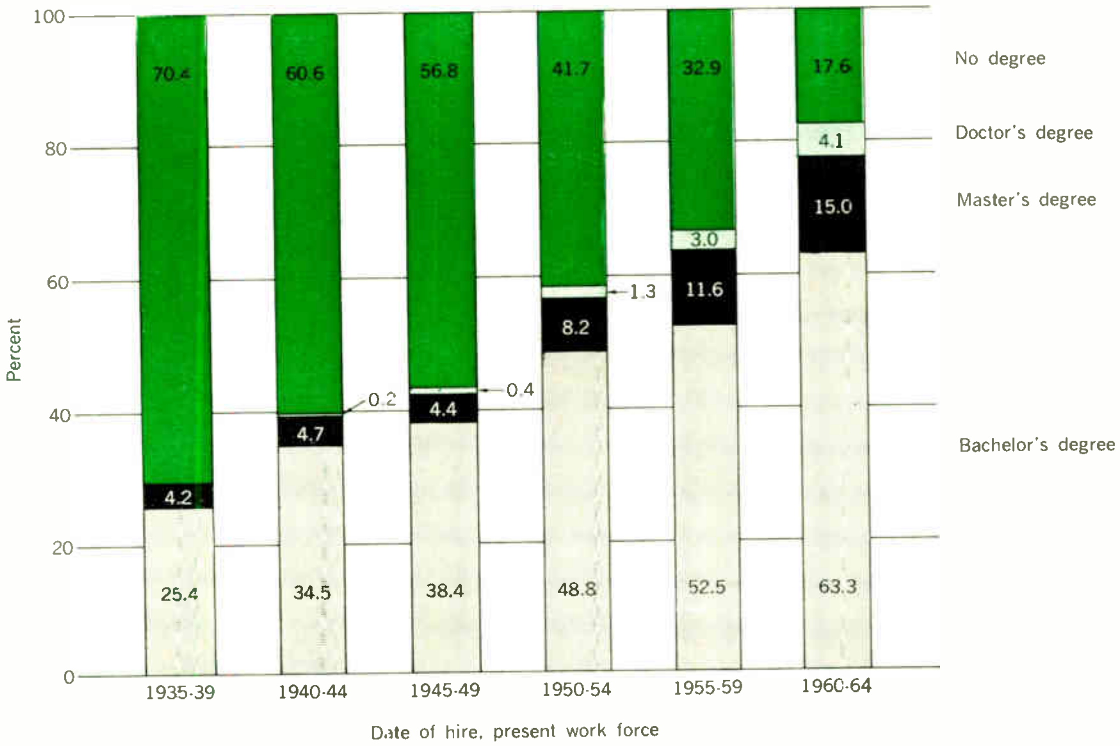


FIGURE 5. Educational composition of engineer/scientist work force, by period hired (Los Angeles and Boston defense R&D establishments).

terminated professionals from two of the Phase I companies and from the Phase II companies that supplied this information. It would appear, by relating Table III to Table II, that the net institutional change of the companies studied in Phase I would be small (i.e., they lost professionals to the same types of institutions and in about the same proportion as they obtained them originally). In contrast, the Los Angeles and Boston companies experienced a significant net gain from companies not involved in defense R&D, and from colleges and universities.

Another way of examining the data would be to determine the proportion of types of institutions represented in the current work force compared with the terminated work force. It was found that among the Los Angeles and the Boston companies studied, although 42.9 percent of the current work force was hired from jobs with other defense R&D establishments, only 39.5 percent of the terminated work force had originated with other defense facilities. This indicates that such experienced professionals hired from companies in the same industry are more likely to be retained. By the same method, it can be shown that professionals hired from educational institutions have higher-than-average turnover rates. As will be shown, the question of whether this can be attributed to age, rather than institutional source, is moot.

The young, single, technical professional represents the epitome of interinstitutional transferability. Demands for his services are such that he can write his own ticket to almost any geographical area and type of institution. The more advanced his degree, the more marketable are his services.

It was no surprise, then, to find the difference in age profile of the terminated work force vs. the current work force depicted in Fig. 4. One would expect a greater mobility among the younger work force. Moreover, it was found that age and average seniority of the current professional work forces studied in Phases II and III had a linear relationship—at least until age 50. This relationship will be discussed under "Geographical differences."

Since we equated degree level with demand, we expected that those holding doctors' degrees would be in great demand, and would experience the highest turnover rate, followed by those holding masters' and bachelors' degrees and the nondegreed, in that order. In Phase I, however, the order of turnover was, first, doctors' degrees, followed by nondegreed professionals and holders of bachelors' degrees, with master-degree holders showing the lowest turnover of all.

In Phase II the order was repeated in Boston, but, surprisingly, in Los Angeles the exact opposite of what we originally expected occurred—i.e., the nondegreed had the highest turnover rate, followed by bachelors, masters, and doctors in that order.

What inferences can be drawn concerning the impact of education on turnover from these interestablishment movement rates? One might be that a basic economic theory (supply–demand differences) does not play as important a role in moving professionals as we might expect, particularly in a mature R&D complex. Another might be that the periodic layoffs, the "project pistons" that take place in this industry, may be a self-renewing purging mechanism, which rids itself of the more marginal workers, evidently the nondegreed. Such a hypothesis would not support the view that *no* nondegreed technical professionals could "cut the mustard," but would stipulate that, more and more, to participate in the high-technology industry, a degree, or better still an advanced degree, is rapidly becoming a requisite. This stand is reflected also in the hiring policies followed by employers in this industry, as reflected in Fig. 5 where it is shown that fewer and fewer nondegreed persons are being hired in professional capacities.

IV. Turnover related to route of application for Los Angeles–Boston scientists and engineers

Stated Reason for Applying for Work	Currently Employed N = 3045	Terminated Professionals N = 413
Company-initiated through third party		
Ad in newspaper	10.6	19.1
Ad in magazine	2.2	4.4
Ad in trade journal	1.8	1.9
Placement service	5.1	14.3
Other-company-initiated		
College recruitment	4.8	10.2
Other recruitment	2.8	0.7
Employee-initiated		
Personal acquaintance in company	51.1	28.6
Company's work in field	17.0	12.1
Advancement opportunity	2.8	3.9
Other	1.9	4.9

The path followed by the technical professional in seeking employment is also of interest—especially as it relates to subsequent interinstitutional movement. In the second phase of our study, it was found that a few companies inquired on the application form, "Why did you apply to the *X* company?" Although the size of the sample, particularly for terminated professionals, was smaller than we would wish, nevertheless some inferences have been drawn from the responses.

Table IV shows that a surprisingly high percentage of those applying for work do so on the advice or counsel of a friend. However, the proportional representation of this category among the terminated is much smaller, indicating that professionals who apply for work through friends and acquaintances will stay with their employer.

It seems evident, also from the table, that an organization wishing to reduce turnover will make its professional needs known through its existing employees, through publicity, or by others means, in a way to induce professionals to take the initiative to apply.

The turnover rates in the mature complexes of Los Angeles and Boston were found to be greater than in the Twin Cities. This difference in retention is pictured in Fig. 6. It can be deduced from this illustration that the average seniority in the Los Angeles–Boston complexes is about one third of a year for each year that the professional's age exceeds 20, whereas in the Twin Cities this average is over two fifths of a year for each year over 20.

The extent to which these rates might reflect attitude

FIGURE 6. Age–seniority comparisons of engineers and scientists: Twin Cities vs. Los Angeles–Boston defense R&D establishments.

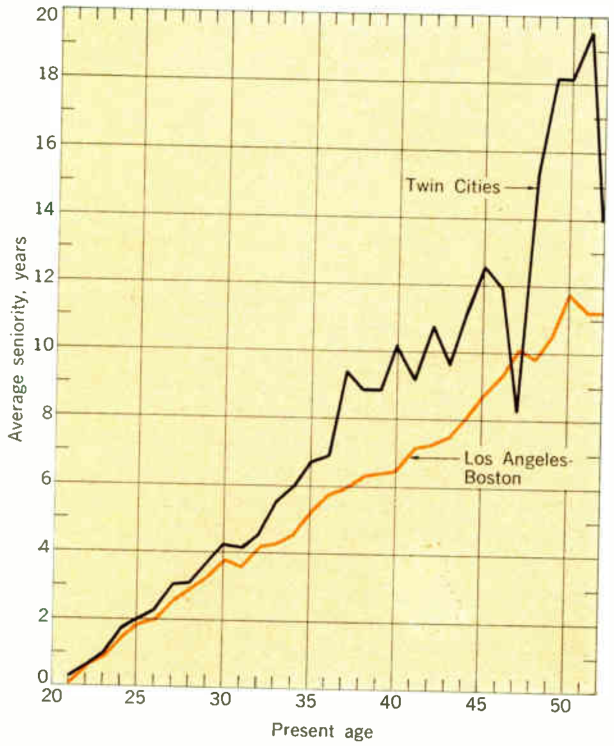
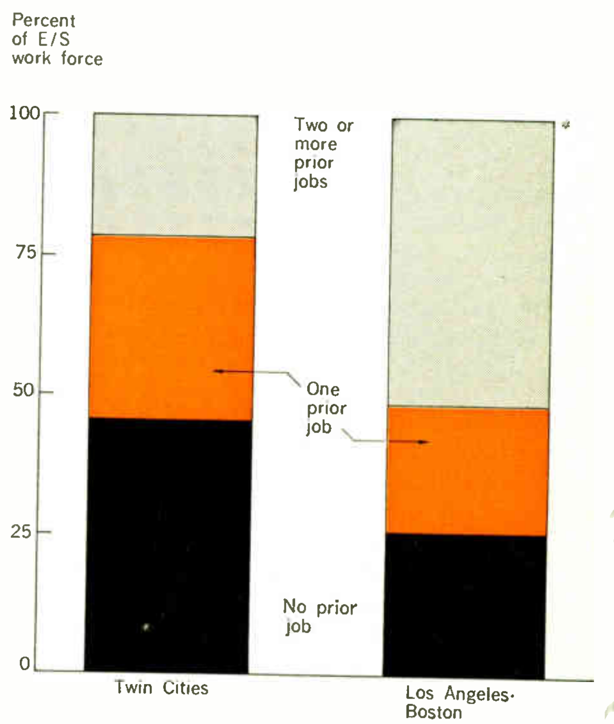


FIGURE 7. Number of prior jobs held by engineers and scientists in the Twin Cities and Los Angeles–Boston defense R&D establishments.



*In Los Angeles and Boston, nearly 24 percent had held four or more prior jobs differences between areas cannot be determined from our data. Such area attitude differences, if they indeed exist, could be on the part of the employers, who might have greater or less paternal feelings, or on the part of the employees, who could bear stronger feelings of loyalty in one area relative to another. On the other hand, because of the great many more high-technology establishments, the points of opportunity in Los Angeles and Boston may induce greater mobility, or perhaps the "project piston" (to be discussed in the following) is more actively at work in these areas. Whatever the cause, the effect is real, particularly when the individual engineer or scientist is considered. Figure 7 shows the much greater amount of interinstitutional movement that has been experienced, at least for the present, in the mature complexes.

Government procurement policies, in protecting the public interests, are meant to be cold, calculating, and objective. Hence, analysis of R&D contracts would show that the government lets large projects to first one facility and then another, with consideration of the effect of such procurement on scientists and engineers as secondary to costs, schedules, and technical specifications. Consequently, the fluctuations in any one employer's revenues for R&D may be great. Certain practices have been followed by employers in their attempts to cope with project fluctuations that affect the movement of professionals. For example, it may be possible over short periods of time for an employer to "stockpile" professionals in anticipation of oncoming work. If such work does not materialize and other projects trail off, sooner or later employers find it necessary to lay off large numbers of professionals, only to be forced sometimes to rehire large numbers at a later date. When it is necessary to release large numbers, positive steps are sometimes taken to help those released relocate in comparable work. This pulling in of large numbers to perform project work and subsequently pushing large numbers out as work is completed has been referred to by Prof. Albert Shapero as the "project piston." An actual case is shown in Fig. 8.

The effects of the project piston on an individual scientist or engineer can be quite unsettling. Yet there are those who claim that the constant "stirring up" can have a beneficial effect on the professional and his industry alike. Benefits to the industry accrue through the cross-fertilization that is said to occur as a result of intercompany transfers. Some attribute our preeminence in high technology as compared with European technical industry (where turnover is nil) to the favorable know-how and practices that transfer with the mobile professional. With each move, the individual professional may benefit by being required to perform new duties that either stretch latent capabilities or require further self-imposed or company-imposed broadening education to make adequate performance possible. Recent critical shortages of structural engineers, for example, resulted in company-sponsored classes to educate civil engineers and mathematicians to perform in a new capacity as structural engineers.

Two views might be taken of the cost of moving a professional from one establishment to another. One might be an overview of the industry in which the project pistons are working away, usually out of synchronization but sometimes synergistically. Table V demonstrates how first one establishment and then another will increase the salary incentive offered to new professionals, as employers desperately try to meet the needs of their project pistons. The incentive increases act like a giant ratchet, edging salaries upward as employers attempt to entice new employees or to retain their existing professionals by meeting salaries offered by competition. The extent that this interinstitutional movement affects overall salaries is illustrated in Fig. 9. Recall that the turnover rates were lower in the Twin Cities and lower among older professionals. Salary differentials stemming from both these causes are reflected, we believe, in this illustration. It is evident that the area salary differentials are

FIGURE 8. Typical requirements for engineering personnel on a major missile development program (source: McGuire, J. I., Jr., "Case study of the Titan II ICBM Program," United Research, Inc., Washington, D.C., Jan. 29, 1964).

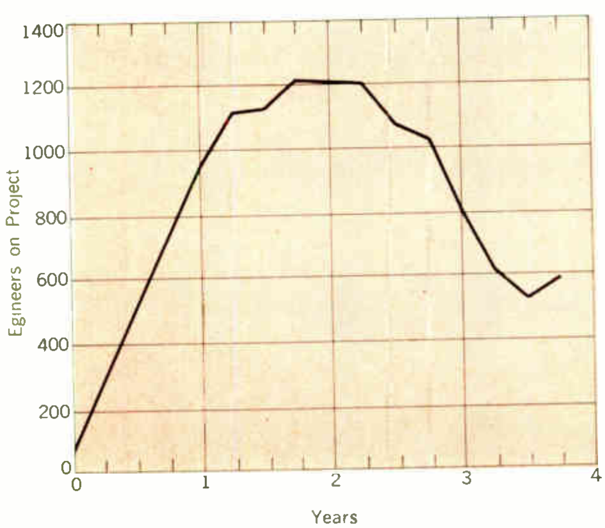
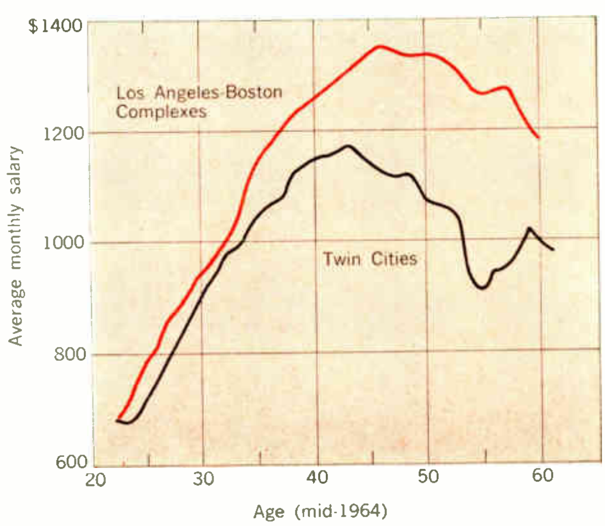


FIGURE 9. Average monthly salary of engineers and scientists with bachelors' degrees, by age, in Twin Cities and Los Angeles–Boston defense R&D establishments. Five-year averages weighted binomially.



V. Average percent of salary increase given engineers/scientists to start with present company, by year hired and for selected industries

	Missile and Space Establishments		Aircraft Establishments			Electronics: Los Angeles Establishments		Electronics: Boston Establishments	
Year	1	2	1	2	3	1	2	3	4
1964	6.8	7.9	18.9	8.2	25.8	9.4	6.5	8.9	10.2
	(524)	(256)	(102)	(31)	(11)	(102)	(106)	(61)	(29)
1963	7.2	9.7	16.7	3.4	14.5	16.3	9.2	14.6	13.5
	(759)	(288)	(158)	(87)	(54)	(158)	(164)	(58)	(109)
1962	8.6	12.8	15.2	9.3	8.9	9.8	10.2	12.7	12.6
	(1399)	(376)	(362)	(69)	(69)	(362)	(102)	(73)	(147)
1961	6.6	11.0	7.1	7.8	9.8	9.1	8.3	11.2	12.7
	(284)	(154)	(291)	(18)	(104)	(291)	(69)	(44)	(64)
1060	4.6	10.5	10.5	5.4	6.1	4.6	5.0	16.4	13.0
	(141)	(148)	(251)	(13)	(43)	(251)	(72)	(22)	(48)
1959	9.8	24.0	28.0	24.4	12.7	13.9	23.3	15.1	12.6
	(171)	(66)	(200)	(6)	(41)	(200)	(64)	(51)	(37)
1958	12.4	3.3	44.4	15.5	2.4	11.0	7.4	6.5	14.3
	(113)	(94)	(105)	(14)	(13)	(105)	(34)	(75)	(27)
1957	1.6	4.1	10.1	6.0	0	17.4	5.0	9.3	14.4
	(18)	(127)	(24)	(20)	(17)	(24)	(14)	(55)	(36)
1956	16.5	8.8	25.8	13.4	9.8	36.8	11.4	14.8	18.1
	(94)	(89)	(71)	(24)	(44)	(71)	(16)	(47)	(35)
1955	6.7	8.8	7.6	17.3	25.3	11.8	20.5	4.9	9.0
	(89)	(69)	(77)	(20)	(72)	(77)	(10)	(31)	(5)

Note: Shaded areas indicate time periods in which for a given company the "inducement" salary increase percentage decreased.
( ) = Number in sample.

VI. Average monthly salary of engineers/scientists with the bachelor's degree as related to number of full-time jobs held (Los Angeles and Boston defense R&D establishments)

Number of Full-Time Jobs Held	Age 20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64
1	810	936	1118	1450	1576	1561	1513	1620	1200
(N)	(103)	(418)	(215)	(101)	(64)	(14)	(6)	(1)	(1)
2	829	955	1153	1369	1511	1505	1580	1481	1117
(N)	(11)	(277)	(254)	(142)	(117)	(78)	(27)	(8)	(3)
3	663	950	1132	1372	1496	1548	1533	1411	1157
(N)	(3)	(95)	(179)	(155)	(142)	(70)	(18)	(9)	(3)
4		1045	1121	1289	1425	1487	1337	1417	1282
(N)	(0)	(26)	(88)	(122)	(126)	(50)	(14)	(10)	(5)
5		1044	1161	1342	1441	1342	1408	1466	1178
(N)	(0)	(5)	(47)	(87)	(95)	(46)	(23)	(10)	(4)
6		1315	1136	1318	1464	1490	1336	1530	1253
(N)	(0)	(4)	(20)	(36)	(67)	(40)	(16)	(7)	(3)
7			1441	1306	1291	1468	1214	1410	1330
(N)	(0)	(0)	(8)	(20)	(31)	(28)	(11)	(7)	(2)
8			1102	1263	1399	1397	1363	1327	1350
(N)	(0)	(0)	(5)	(15)	(28)	(19)	(9)	(7)	(5)
9				1355	1513	1270	1395		1276
(N)	(0)	(0)	(0)	(2)	(4)	(3)	(2)	(0)	(5)
10								1080	
(N)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(1)	(0)
11 or more					1100	1363	1174	1130	1460
(N)	(0)	(0)	(0)	(0)	(1)	(3)	(1)	(1)	(1)

Note: Shaded area indicates limit of data considered relevant in view of sample size.

less among the younger, more mobile professionals, and that, regardless of area, there is a trailing off of average salaries with age.

The other view that might be taken is that of the professional. Examples are sometimes cited in which an individual will leave an organization and, after two or three additional moves, return to his first employer at a salary substantially above that of his peers who stayed. The facts, however, do not support the theory that such examples are common. It is shown in Table VI that, although it may be true that the young professional who has had more than one job experience will be better remunerated, in the long run those remaining loyal to one or a few employers will be ahead. On a life-long basis, the "loyal" professional will have aggregated about 4 percent more in pay and will not have had the out-of-pocket expenditures that sometimes accompany job moves nor the loss of vested retirement funds that may be suffered in the process of making a change. (However, "present value" analysis, not attempted here, might temper this conclusion somewhat.)

Conclusion

Examining the movement of over 40 000 scientists and engineers at work in six geographical areas and in high-technology industry has led to the following conclusions or strong inferences:

- Technical professionals are influenced by the same motivations to move geographically as the general population. Hence their movements are closely synchronized with well-established migratory paths.
- The extent of interarea movement of professionals is linked to the age and size of an R & D complex to which professionals are attracted. Although the absolute numbers may be greater, the proportion of in-migratory professionals will be less in a mature complex.
- Variation in movement between employers varies among employers. The extent of this variation in turnover rates may reflect differences in internal policies or of failure or success in winning overlapping contracts.
- Movement between employers (and probably between areas) is related to the age of the professional. The evidence indicates that if a lower turnover rate is a goal, older, more experienced professionals should be hired.
- As a professional's seniority increases, the chance of his leaving an employer becomes less.
- Professionals tend to remain within an industry. Moreover, those having had prior experience in the same industry will experience lower turnover rates on the average than those hired from a different industry.
- The higher the level of education, the less the probability of interinstitutional movement. Demand for the most highly educated may alter this generalization, particularly in newer or rapidly expanding aerospace complexes.
- Professionals who take the initiative in applying for work appear to be more satisfied with their work, as reflected in lower turnover rates.
- Larger, older aerospace complexes on the average employ professionals who have had more different job experiences than a nascent high-technology area.
- Contracting policies of the U.S. Government are conducive to professional movement among competing establishments. The process increases costs but may also improve capabilities of the industry, and of the individuals moved.
- Changing jobs may provide some short-term salary advantages to the individual professional, but in the long run, keeping interinstitutional movements to a minimum may pay off.
- The role of personal ties in the movement of technical professionals is greater than one might expect. Such ties may be the key to interarea migration as technical professionals follow the leaders down migration streams. Personal ties have also been shown to be a major, basic ingredient in interfacility movement, and those following friends to an employer will tend to be loyal employees. Lastly, the homing tendency may also be a manifestation of the power of personal ties—in this instance causing complete reversal of earlier moves.

BIBLIOGRAPHY

Area Manpower Guide (book), BES no. R-174, U.S. Department of Labor, Bureau of Employment Security, 1957.

Bosley, J. J., and Shapero, A., "A preliminary analysis of inter-specialty mobility of technical/professional manpower resources," Stanford Research Institute, Menlo Park, Calif., Sept. 1967.

Draheim, K., Howell, R., and Shapero, A., "The development of a potential defense R & D complex," Stanford Research Institute, Menlo Park, Calif., July 1966.

Hearings before Subcommittee on Government Research, U.S. Senate, Committee on Government Operations, May 11, 1967.

Howell, R. P., "Are defense engineers switchable," *Eng. Opportunities*, May 1968.

Howell, R. P., "Engineers on the move," *Eng. Manpower Bull.*, no. 11, June 1968.

Howell, R. P., "On professional salaries," *IEEE Spectrum*, vol. 6, pp. 22–29, Feb. 1969.

"Mobility for state and state economic areas," *U.S. Census of Population*, DC(2)2B, U.S. Department of Commerce, Bureau of Census, 1960.

Robinson, E. A. G., *The Structure of Competitive Industry*. Chicago: University of Chicago Press, 1958.

Shapero, A., "Life styles of engineering," *Space Aeron.*, Mar. 1969.

Shapero A., Howell, R., and Tombaugh, J., "An exploratory study of the structure and dynamics of the R&D industry," Stanford Research Institute, Menlo Park, Calif., June 1964.

Shapero, A., Howell, R., and Tombaugh, J., "The structure and dynamics of the R&D industry," Stanford Research Institute, Menlo Park, Calif., Nov. 1965.

Shapero, A., "Managing technical and intellectual resources," *Business Horizons*, Apr. 1969.

Webb, E. J., *et al.*, *Unobtrusive Measures: Non-reactive Research in the Social Sciences.* Chicago: Rand, 1966.

Richard P. Howell (SM) is a senior associate of the New Management Center of Palo Alto, Calif., which is primarily engaged in reducing the gap between the state of management technology and its application. Under contract to private groups, the firm specializes in education, research, and planning for management. Mr. Howell received the B.S. degree in mathematics from the University of Washington in 1942 and, after a tour as a naval officer during World War II, was employed as an analyst in industry and later as a consultant in industrial management. Prior to joining the New Management Center, he was with Stanford Research Institute, in the Management Sciences Division, for more than 13 years. He has had a number of papers published.

The operational amplifier in linear active networks

The analysis of some of the pertinent network characteristics of a nonideal operational amplifier, based on the active two-port network that it represents, provides a useful starting point for designing linear active networks

G. S. Moschytz *Bell Telephone Laboratories, Inc.*

In this article the two-port characteristics, the gain stability, and the sensitivity of the nonideal operational amplifier are presented. A unified approach is taken to cover both the inverting and the noninverting modes of amplifier operation. Some of the sensitivity aspects pertaining to network synthesis using operational amplifiers are also discussed.

High-performance silicon integrated operational amplifiers are presently available commercially at very reasonable (and still decreasing) prices. It is therefore no wonder that these devices are finding widespread usage in the synthesis of linear active networks. Very often the operational amplifier is represented analytically by idealized characteristics that differ appreciably from those of commercial monolithic integrated units; consequently, there is a need for an evaluation of the appropriate network characteristics of the nonideal operational amplifier.

Characterization of the operational amplifier

Two-port characteristics. A nonideal operational amplifier in the inverting mode and its equivalent circuit diagram are shown in Fig. 1(A). R_i and R_o are the open-loop input and output impedances, R_b the intrinsic feedback impedance, and A_o the open-loop differential voltage gain. Actually, these parameters are frequency-dependent, but for the purpose of this analysis, they are considered constant. The remaining three resistors are external to the amplifier. R_G and R_F determine the closed-loop gain, and R_n is required to minimize the offset voltage at the output.* V_i refers to the voltage at the inverting input terminal, and V_n refers to the voltage at the noninverting input terminal of the amplifier. Similarly, the noninverting mode of a nonideal operational amplifier and its equivalent circuit diagram are shown in Fig. 1(B). It contains the identical components as those in the inverting mode.

In terms of the chain or ($\mathfrak{ABCD}$) matrix elements of the two-port shown in Fig. 2, the voltage and current gain, input impedance, and output impedance of the operational amplifier can be assumed to be independent of both the source impedance R_S and the load impedance R_L. The former can be included in R_G or R_n [see Fig. 1(A) and 1(B), respectively]. The latter can be considered much larger than the amplifier output impedance and, therefore, negligible. Thus,

$$G = \left.\frac{V_{out}}{V_{in}}\right|_{I_{out}=0} = \frac{1}{\mathfrak{A}} \tag{1}$$

$$-\left.\frac{I_2}{I_1}\right|_{V_2=0} = \frac{1}{\mathfrak{D}} \tag{2}$$

$$Z_{in} = \left.\frac{V_{in}}{I_1}\right|_{I_{out}=0} = \frac{\mathfrak{A}}{\mathfrak{C}} \tag{3}$$

$$Z_0 = \left.\frac{V_2}{I_2}\right|_{V_1=0} = \frac{\mathfrak{B}}{\mathfrak{A}} \tag{4}$$

To obtain the chain matrix elements of the two amplifier configurations shown in Fig. 1, the aforementioned gains and impedances can be calculated directly and Eqs. (1) through (4) can be solved for the matrix elements. Neglecting the intrinsic feedback impedance R_b, which is generally very large, the chain matrixes for the amplifier in the inverting and noninverting modes, respectively, are given by Eqs. (5a) and (5b) at the top of page 43.

The expressions obtained in calculating Eqs. (1) through (4) suggest a flow-graph representation of the operational amplifier that is the same for either mode of operation. This is described in the following paragraphs.

Feedback representation. It is useful to look at the operational amplifier as a feedback amplifier with a flow graph, as shown in Fig. 3. The corresponding transmission function is

$$G = \frac{V_{out}}{V_{in}} = \alpha \cdot \frac{A\beta}{1 + A\beta} \tag{6}$$

Using the subscripts I and N to differentiate between the inverting and noninverting modes, respectively, we obtain from Eqs. (1), (5a), and (5b) the coefficients α_I, A_I, β_I, α_N, A_N, and β_N for Eq. (6) as follows:

α is the closed-loop gain that would be obtained with an ideal operational amplifier characterized by

$$A_o = \infty \qquad R_i = \infty \qquad R_o = 0 \tag{7}$$

For the inverting mode, we obtain

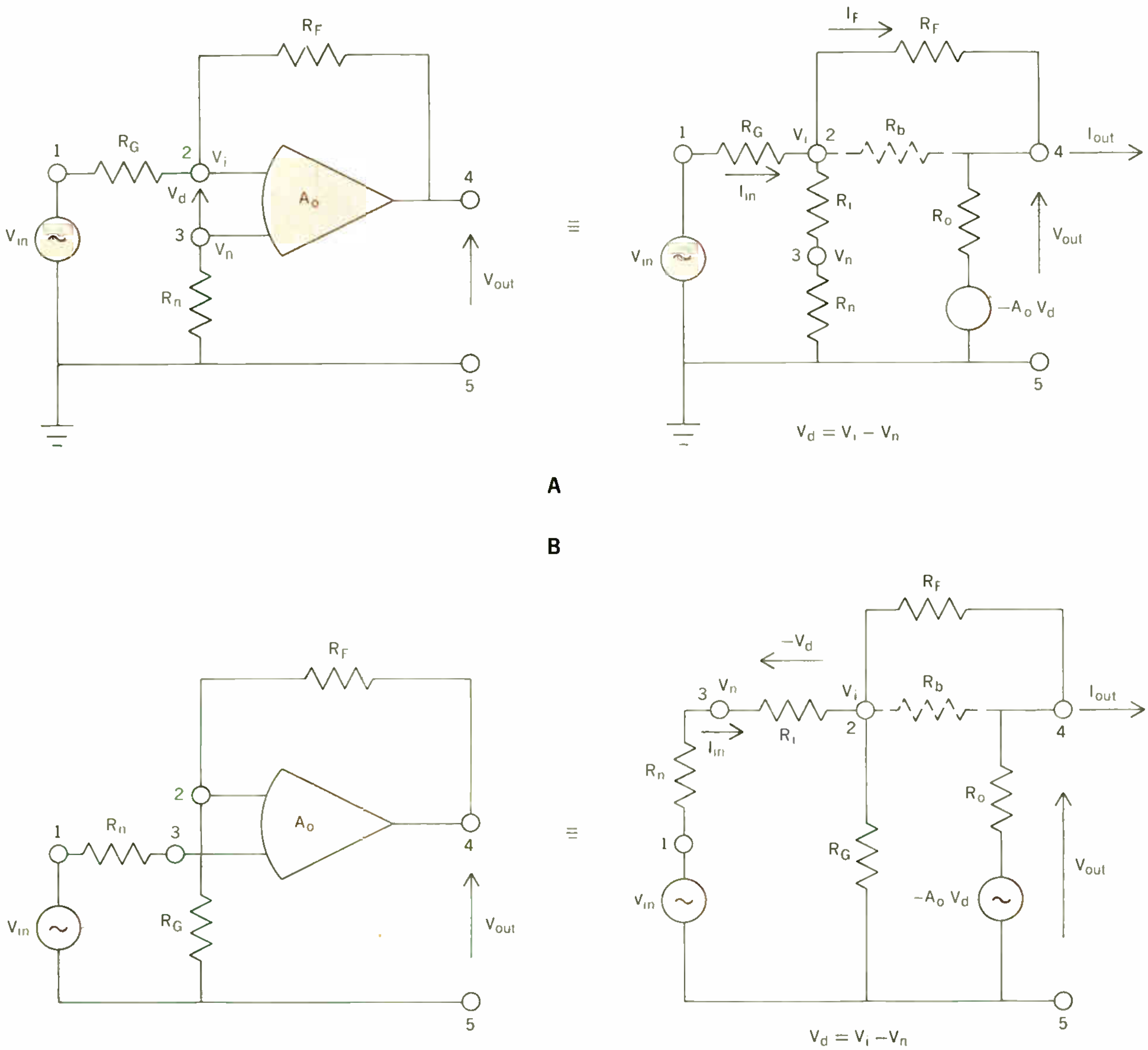
$$\alpha_I = -\frac{R_F}{R_G} \tag{8}$$

* The bias currents at each input terminal generate a dc error or offset voltage at the output if the dc paths to ground at each input are not equal. Thus, for minimum offset voltage at the output, R_n must equal the parallel combination of R_G with the series combination of R_F and R_o. (The load impedance can generally be neglected compared with the small value of R_o.)

$$[F]_I = \begin{bmatrix} \dfrac{(R_F + R_o)(R_i + R_n) + R_G(R_F + R_o + R_i + R_n) + A_oR_GR_i}{R_o(R_i + R_n) - A_oR_iR_F} & \dfrac{R_o[R_G(R_F + R_i + R_n) + R_F(R_i + R_n)]}{R_o(R_i + R_n) - A_oR_iR_F} \\ \\ \dfrac{R_F + R_o + R_n + R_i(A_o + 1)}{R_o(R_i + R_n) - A_oR_iR_F} & \dfrac{R_o(R_F + R_i + R_n)}{R_o(R_i + R_n) - A_oR_iR_F} \end{bmatrix} \quad (5a)$$

$$[F]_N = \begin{bmatrix} \dfrac{(R_i + R_n)(R_G + R_F + R_o) + R_G(R_F + R_o) + A_oR_iR_G}{R_GR_o + A_oR_i(R_G + R_F)} & \dfrac{R_o[(R_i + R_n)(R_F + R_G) + R_FR_G]}{R_GR_o + A_oR_i(R_G + R_F)} \\ \\ \dfrac{R_G + R_F + R_o}{R_GR_o + A_oR_i(R_G + R_F)} & \dfrac{R_o(R_F + R_G)}{R_GR_o + A_oR_i(R_G + R_F)} \end{bmatrix} \quad (5b)$$

FIGURE 1. Nonideal operational amplifier with the corresponding equivalent circuit diagram. A—Inverting mode. B—Noninverting mode.



and for the noninverting mode,

$$\alpha_N = \frac{R_F + R_G}{R_G} \tag{9}$$

A is the forward gain. For the inverting mode [see Fig. 1(A)], it corresponds to the output voltage when a unity voltage signal is applied to the input of the operational amplifier (terminals 2 and 5) and V_{in} is set equal to zero. For the noninverting mode [see Fig. 1(B)], it corresponds to the output voltage when a unity voltage is generated across the amplifier input (terminals 2 and 3) and the signal source V_{in} is removed (i.e., terminal 1 open). We obtain by inspection:

$$A_I = A_o \frac{1 - \frac{1}{A_o}\frac{R_o}{R_F}\left(1 + \frac{R_n}{R_i}\right)}{\left(1 + \frac{R_o}{R_F}\right)\left(1 + \frac{R_n}{R_i}\right)} \tag{10}$$

and

$$A_N = A_o \frac{1 + \frac{\beta_o}{A_o}\frac{R_o}{R_i}}{1 + \frac{R_o}{R_F}(1 - \beta_o)} \tag{11}$$

where β_o is as defined further on; see Eq. (14).

β is the feedback factor. In the inverting mode [see Fig. 1(A)], it corresponds to the voltage fraction at the input terminals 2 and 5 when a unity voltage signal is applied across the output terminals and the input signal source V_{in} is short-circuited. In the noninverting mode, Fig. 1(B), it corresponds to the voltage fraction appearing across the input terminals 2 and 3 under the same conditions at the output and with V_{in} short-circuited. Thus,

$$\beta_I = \frac{\beta_o}{1 + \frac{R_p}{R_i + R_n}} \tag{12}$$

and

$$\beta_N = \frac{\beta_o}{1 + \frac{R_n + \beta_o R_F}{R_i}} \tag{13}$$

where β_o is the feedback factor of the ideal amplifier characterized by Eq. (7)—namely,

$$\beta_o = \frac{R_G}{R_G + R_F} \tag{14}$$

and $$R_p = \frac{R_G R_F}{R_G + R_F} = \beta_o R_F \tag{15}$$

The ratio of actual to idealized closed-loop gain follows from (6) as

$$\frac{G}{\alpha} = \frac{1}{1 + \frac{1}{A\beta}} \tag{16}$$

Since $A\beta$, the amplifier-loop gain, must be much larger than unity for gain stability, Eq. (16) can be expanded into a form of Taylor series as follows:

$$\frac{G}{\alpha} = 1 - \frac{1}{A\beta} + \left(\frac{1}{A\beta}\right)^2 - \left(\frac{1}{A\beta}\right)^3 + \cdots = 1 - E \tag{17}$$

where E is the gain error. In general, E can be approximated by the linear term in (17). With the expressions derived thus far, the gain error E, as well as the amplifier characteristics defined by Eqs. (1) to (4), can now be calculated directly.

Assuming a minimum closed-loop gain of unity,* the limits on β_o are

$$0 \leq \beta_o \leq 0.5 \tag{18}$$

To minimize dc offset, the external resistor R_n is added to the noninverting input terminal of the amplifier (see previous footnote). It must satisfy the condition

$$R_n = \frac{R_G(R_F + R_o)}{R_G + R_F R_o} \approx R_p = \beta_o R_F \tag{19}$$

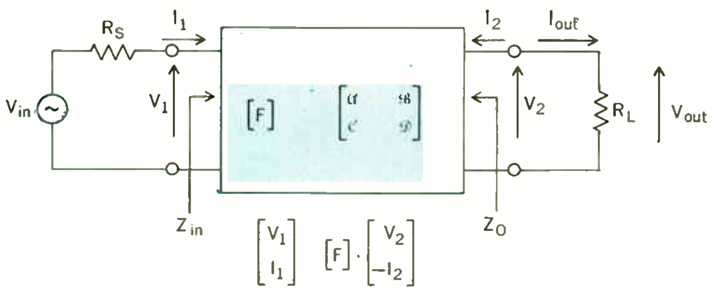
To simplify further expressions it is useful to introduce the substitutions

$$\epsilon_i = \frac{R_F}{R_i} \tag{20}$$

$$\epsilon_o = \frac{R_o}{R_F} \tag{21}$$

$$\epsilon_o' = \frac{R_o}{R_F}(1 - \beta_o) \tag{22}$$

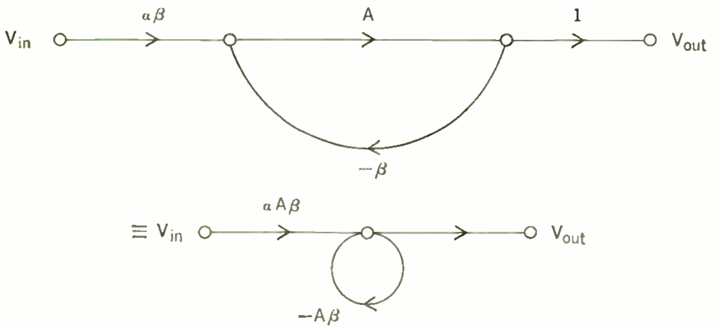
In practice the following inequalities hold:



$$\frac{V_2}{V_1} = \frac{R_L}{\mathcal{B} + \mathcal{A}R_L} \qquad -\frac{I_2}{I_1} = \frac{1}{\mathcal{D} + \mathcal{C}R_L} \qquad Z_{in} = \frac{\mathcal{A}R_L + \mathcal{B}}{\mathcal{C}R_L + \mathcal{D}} \qquad Z_O = \frac{\mathcal{D}R_S + \mathcal{B}}{\mathcal{C}R_S + \mathcal{A}}$$

FIGURE 2. Linear two-port given by its chain matrix. R_S = source impedance; R_L = load impedance.

FIGURE 3. Operational-amplifier flow graph. α = ideal closed-loop gain; A = forward gain; β = feedback factor.



* From Eq. (9) it is clear that unity gain is actually the lowest gain possible in the noninverting mode. In the inverting mode, a closed-loop gain of less than unity can be obtained—that is, $R_F < R_G$ in Eq. (8)—but is rarely required.

$$A_o \gg 1 \tag{23}$$

$$\epsilon_o' < \epsilon_o \ll 1 \tag{24}$$

In these terms, the dc characteristics of the operational amplifier can be derived and, in view of the inequalities (23) and (24), can be simplified considerably, resulting in the expressions listed in Table I. The feedback resistor for minimum gain error (line 7) was obtained by setting the derivative of the gain error (line 5) with respect to R_F equal to zero and solving for R_F. From lines 4 and 5,

$$\frac{E_N}{E_I} \approx \frac{1+\epsilon_o'}{1+\epsilon_o} \leq 1 \tag{25}$$

That is, the noninverting gain error is always less than or equal to the inverting gain error. As an illustrative example, the gain error E has been plotted (Fig. 4) as a function of R_F and α for the inverting and noninverting mode of a typical operational amplifier.

AC characterization. In the preceding discussion A_o, the open-loop differential voltage gain of the amplifier, was considered to be constant. In reality, however, it is frequency-dependent; in fact, it can be approximated by a rational function of poles as follows:

$$A_o(s) = \frac{A_o}{(s+\omega_1)(s+\omega_2)(s+\omega_3)} \tag{26}$$

This representation corresponds to the Bode plot of the open-loop gain, where ω_1, ω_2, and ω_3 are the corner frequencies in the band of interest (from zero to unity-gain crossover), and the slope increases by 6 dB per octave at each corner. A typical representation of the Bode plot for the open-loop gain corresponding to Eq. (26) and the closed-loop gain corresponding to Eq. (6) is shown in Fig. 5. This representation graphically shows the open-loop gain A, the closed-loop gain G, and the loop gain $A\beta$. As long as the loop gain $A\beta$ is large, the closed-loop gain is approximately equal to α—that is, the error E discussed previously is negligibly small. It is clear, however, from this figure that the loop gain is now frequency-dependent and, in fact, decreases to 0 dB at the frequency ω_α at which the two plots intersect. To satisfy Bode's criterion for absolute stability of a feedback amplifier,* the frequency response of the operational amplifier must generally be limited still further by external frequency-stabilizing RC circuits. This is shown for a typical situation by the dashed lines in Fig. 5. The loop gain has thereby been modified to roll off at the frequency Ω, which is lower than the first natural-break frequency (ω_1) of the amplifier.

Numerous methods of incorporating frequency-stabilizing RC networks to ensure the required 6 dB-per-octave rate of closure are available in the literature.[3-5] One approach is to modify the frequency response of the open-loop gain by incorporating lead–lag networks in the forward path of the amplifier. Another is to modify the frequency response of the closed-loop gain by incorporating frequency-dependent networks in the feedback network. Either way, the resulting loop gain generally has a single pole† (on the negative real axis) and can be

* According to Bode's stability criterion, the rate of closure between the open-loop and the closed-loop frequency response must be less than 12 dB per octave. To guarantee sufficient phase margin the rate of closure is usually chosen to be 6 dB per octave.

† The number of loop-gain poles can be any number N as long as there are $N - 1$ zeros such that the roll-off in the vicinity of unity crossover is -6 dB per octave. Whereas multiple poles and zeros widen the loop-gain bandwidth, the compensating networks generating them depend on the value of loop gain available. They are useful, therefore, only when an amplifier is required to provide a predetermined and constant amount of gain. When the required gain is variable, a compensating network is used to generate a single pole such that the -6-dB-per-octave roll-off covers at least the entire range of required gain anticipated. To provide adequate phase margin, it extends either from the minimum loop gain required for gain stability or, more generally, from zero loop gain to dc open-loop gain (corresponding to unity closed-loop gain). The latter is the most commonly employed frequency-compensating scheme and has been assumed in the text.

I. Operational amplifier dc characteristics

	Inverting Mode	Noninverting Mode
1. Ideal closed-loop gain	$\alpha_I = -\frac{R_F}{R_G}$	$\alpha_N = \frac{R_F+R_G}{R_G}$
2. Forward gain	$A_I \approx \frac{A_o}{(1+\epsilon_o)(1+\beta_o\epsilon_i)}$	$A_N \approx \frac{A_o}{1+\epsilon_o'}$
3. Feedback factor	$\beta_I = \beta_o\frac{1+\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}$	$\beta_N = \frac{\beta_o}{1+2\beta_o\epsilon_i}$
4. Loop gain	$A_I\beta_I \approx \frac{A_o\beta_o}{(1+\epsilon_o)(1+2\beta_o\epsilon_i)}$	$A_N\beta_N \approx \frac{A_o\beta_o}{(1+\epsilon_o')(1+2\beta_o\epsilon_i)}$
5. Gain error	$E_I = \frac{1}{A_I\beta_I} - \left(\frac{1}{A_I\beta_I}\right)^2 + \ldots \approx \frac{1}{A_I\beta_I}$	$E_N = \frac{1}{A_N\beta_N} - \left(\frac{1}{A_N\beta_N}\right)^2 + \ldots \approx \frac{1}{A_N\beta_N}$
6. Closed-loop gain	$G_I = \alpha_I\frac{A_I\beta_I}{1+A_I\beta_I} \approx \alpha_I(1-E_I)$	$G_N = \alpha_N\frac{A_N\beta_N}{1+A_N\beta_N} \approx \alpha_N(1-E_n)$
7. Feedback resistor for minimum gain error	$R_F\|_{E_{I_{min}}} = \left[\frac{R_iR_o}{2}(1-\alpha_I)\right]^{1/2}$	$R_F\|_{E_{N_{min}}} = \left[\frac{R_iR_o}{2}(\alpha_N-1)\right]^{1/2}$
8. Input impedance	$(Z_{in})_I \approx R_G\left(1+\frac{\alpha_I}{A_o}\right)$	$(Z_{in})_N \approx A_N\beta_N R_i$
9. Output impedance	$(Z_o)_I \approx \frac{R_o}{A_I\beta_I}$	$(Z_o)_N \approx \frac{R_o}{A_N\beta_N}$

expressed in the form

$$A\beta(s) = A\beta \frac{\Omega}{s + \Omega} \tag{27}$$

The corresponding loop-gain magnitude thus becomes

$$|A\beta(j\omega)| = A\beta \cdot F_\omega \tag{28}$$

where

$$F_\omega = \frac{1}{\sqrt{1 + (\omega/\Omega)^2}} \tag{29}$$

F_ω takes account of the frequency characteristics of the loop gain (in this case, a single pole) and must, of course, be appropriately modified if another pole–zero configuration is used.

Substituting (27) into (6) we obtain, for the closed-loop gain,

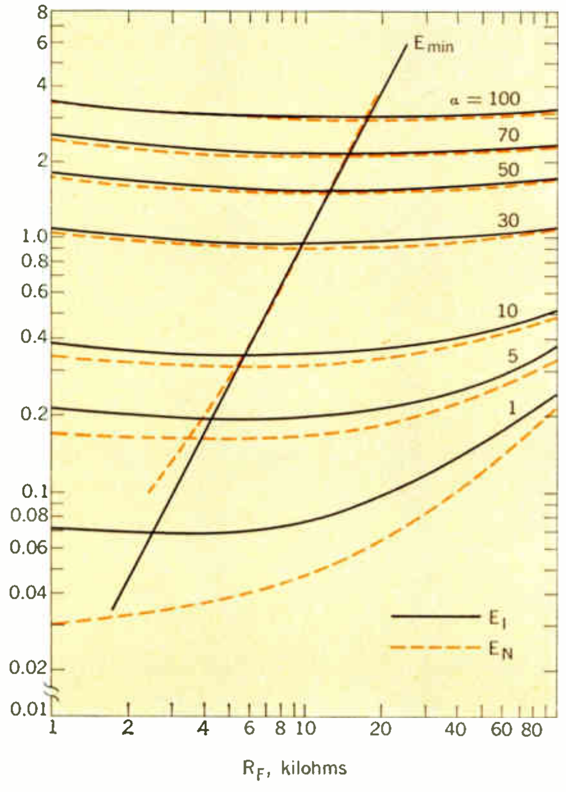
$$G = \alpha \cdot \frac{A\beta\Omega}{s + \Omega(1 + A\beta)} \tag{30}$$

where, for a given value of Ω, the cutoff frequency is proportional to the loop gain. The gain error is also modified now by F_ω, and, according to Table I, line 5,

$$E \approx \frac{1}{A\beta F_\omega} \tag{31}$$

Assuming the most common case given by (29), the correction factor $1/F_\omega$ is plotted for convenience in Fig. 6. It must also be considered for all other amplifier parameters, depending on loop gain, such as the input and output impedance (see Table I, lines 8 and 9).

FIGURE 4. Gain error E for inverting and noninverting modes of typical operational amplifier.



Gain stability. One of the main reasons for the popularity of operational amplifiers in active-network synthesis is that, because of their high open-loop gain, they can provide highly stable closed-loop gain. As will be shown here, the degree of gain stability is dependent mainly on the amount of available loop gain and on the stability of the feedback network.

From Table I (lines 4 and 6) we obtain for the closed-loop gain of the inverting and noninverting operational amplifiers, respectively,

$$G_I = \frac{A_o\beta_o - A_o}{1 + \epsilon_o + 2\beta_o\epsilon_i + 2\beta_o\epsilon_i\epsilon_o + A_o\beta_o} \tag{32}$$

and

$$G_N = \frac{A_o}{1 + \epsilon_o + 2\beta_o\epsilon_i + 2\beta_o\epsilon_i\epsilon_o - \epsilon_o\beta_o - 2\beta_o^2\epsilon_i\epsilon_o + A_o\beta_o} \tag{33}$$

In terms of the network sensitivity function,* the relative variation in closed-loop gain with small variations in the parameters contained in (32) or (33) is given as

$$\frac{dG}{G} = S_{A_o}^G \frac{dA_o}{A_o} + S_{\beta_o}^G \frac{d\beta_o}{\beta_o} + S_{\epsilon_i}^G \frac{d\epsilon_i}{\epsilon_i} + S_{\epsilon_o}^G \frac{d\epsilon_o}{\epsilon_o} \tag{34}$$

The corresponding sensitivity functions have been calculated and are listed in Table II. If we assume that

$$A\beta \gg 1 \tag{35}$$

and consider the inequalities (23) and (24), the resulting functions can be greatly simplified, as shown in Table II.

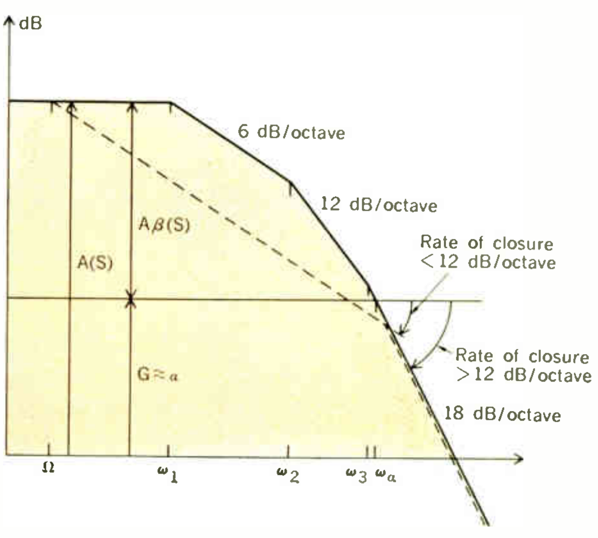
In practice, it is useful to consider finite differentials rather than derivatives when evaluating (34). This technique assumes a linear relationship between the closed-loop gain G and the parameters in (34), which is acceptable as a first approximation as long as the differentials

* The sensitivity[4] of a function ψ with respect to an element x is defined as

$$S_x^\psi = \frac{d\psi}{dx}\frac{x}{\psi} = \frac{d(\ln \psi)}{d(\ln x)}$$

[see Eq. (46)].

FIGURE 5. Bode plot of open-loop gain and closed-loop gain for typical operational amplifier.



remain small. These differentials are obtained as follows:

$\Delta A_o/A_o$ represents the relative tolerance and drift (for example, with temperature) of the open-loop gain. These values are generally specified by the manufacturer.

$\Delta\beta_o/\beta_o$ can be calculated directly from the definition in Eq. (14), from which we obtain

$$\frac{\Delta\beta_o}{\beta_o} = (1-\beta_o)\cdot\frac{\dfrac{\Delta R_G}{R_G}-\dfrac{\Delta R_F}{R_F}}{1+\dfrac{\Delta(R_G+R_F)}{R_G+R_F}} \tag{36}$$

The term

$$\frac{\Delta R_G}{R_G}-\frac{\Delta R_F}{R_F}$$

is a measure of the tracking capability of the external feedback resistors. For integrated circuits, and particularly for tantalum thin-film resistors, it can be made very small—about ±5 parts per million per degree C.

$\Delta\epsilon_i/\epsilon_i$ follows from Eq. (20):

$$\frac{\Delta\epsilon_i}{\epsilon_i} = \frac{\dfrac{\Delta R_i}{R_i}-\dfrac{\Delta R_F}{R_F}}{1+\dfrac{\Delta R_i}{R_i}} \tag{37}$$

The resistor R_F is usually external to the operational amplifier, whereas R_i is a part of it. R_i is thus subject to the relatively wide tolerances and drift common with silicon integrated resistors—typically of the order of ±20 percent. In contrast, the external resistor R_F can be chosen to be arbitrarily stable. Thus, in general,

$$\frac{\Delta R_F}{R_F} \ll \frac{\Delta R_i}{R_i} \tag{38}$$

Therefore,

$$\frac{\Delta\epsilon_i}{\epsilon_i} = \frac{\dfrac{\Delta R_i}{R_i}}{1+\dfrac{\Delta R_i}{R_i}} \tag{39}$$

$\Delta\epsilon_o/\epsilon_o$ follows from Eq. (21):

$$\frac{\Delta\epsilon_o}{\epsilon_o} = \frac{\dfrac{\Delta R_o}{R_o}-\dfrac{\Delta R_F}{R_F}}{1+\dfrac{\Delta R_F}{R_F}} \tag{40}$$

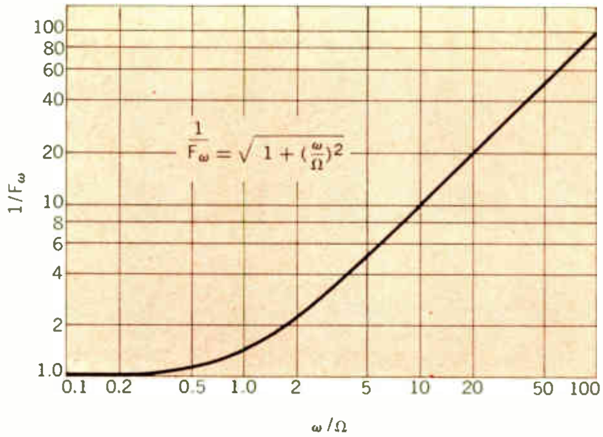


FIGURE 6. Loop-gain correction factor $1/F_\omega$ for the most common method of frequency compensation.

Since R_o, being a part of the operational amplifier, is also a silicon integrated resistor, it can be assumed here that

$$\frac{\Delta R_F}{R_F} \ll \frac{\Delta R_o}{R_o} \tag{41}$$

and thus

$$\frac{\Delta\epsilon_o}{\epsilon_o} = \frac{\Delta R_o}{R_o} \tag{42}$$

By combining Eqs. (36), (39), and (42) with the sensitivity functions in Table II, we can obtain the expressions for gain stability listed in Table III, line 1. For most commercially available operational amplifiers,

$$\frac{\epsilon_o'}{1+\epsilon_o'}\,\frac{\Delta R_o}{R_o} < \frac{\epsilon_o}{1+\epsilon_o}\,\frac{\Delta R_o}{1+R_o} \ll \frac{\Delta A_o}{A_o} \tag{43}$$

Furthermore, if we use temperature-stable external feedback resistors, then

$$\frac{\Delta R_G}{R_G} \approx \frac{\Delta R_F}{R_F} \approx 0 \tag{44}$$

Similarly, if we use integrated resistors that track closely with ambient variations, then

$$\frac{\Delta R_G}{R_G} - \frac{\Delta R_F}{R_F} \approx 0 \tag{45}$$

II. Gain sensitivity functions

	Inverting Mode	Noninverting Mode
1.	$S_{A_o}^{G_I} = \dfrac{1}{1+A_I\beta_I} \approx \dfrac{1}{A_I\beta_I}$	$S_{A_o}^{G_N} = \dfrac{1}{1+A_N\beta_N} \approx \dfrac{1}{A_N\beta_N}$
2.	$S_{\beta_o}^{G_I} = \dfrac{\beta_o}{1-\beta_o}\cdot\dfrac{\dfrac{1}{\beta_o}+\dfrac{1}{A_I\beta_I}\cdot\dfrac{1+2\epsilon_i}{1+2\beta_o\epsilon_i}}{1+\dfrac{1}{A_I\beta_I}} \approx -\dfrac{1}{1-\beta_o}$	$S_{\beta_o}^{G_N} = -\dfrac{1+\dfrac{1}{A_N\beta_N}\cdot\dfrac{2\epsilon_i\beta_o}{1+2\epsilon_i\beta_o}\cdot\dfrac{1+\epsilon_o(1-2\beta_o-0.5\epsilon_i)}{1+\epsilon_o(1-\beta_o)}}{1+\dfrac{1}{A_N\beta_N}} \approx -1$
3.	$S_{\epsilon_i}^{G_I} = -\dfrac{2\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}\cdot\dfrac{1}{1+A_I\beta_I} \approx -\dfrac{2\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}\cdot\dfrac{1}{A_I\beta_I}$	$S_{\epsilon_o}^{G_N} = -\dfrac{2\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}\cdot\dfrac{1}{1+A_N\beta_N} \approx -\dfrac{2\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}\cdot\dfrac{1}{A_N\beta_N}$
4.	$S_{\epsilon_o}^{G_I} = -\dfrac{\epsilon_o}{1+\epsilon_o}\cdot\dfrac{1}{1+A_I\beta_I} \approx -\dfrac{\epsilon_o}{1+\epsilon_o}\cdot\dfrac{1}{A_I\beta_I}$	$S_{\epsilon_o}^{G_N} = -\dfrac{\epsilon_o'}{1+\epsilon_o'}\cdot\dfrac{1}{1+A_N\beta_N} \approx -\dfrac{\epsilon_o'}{1+\epsilon_o'}\cdot\dfrac{1}{A_N\beta_N}$

III. Gain stability

	Inverting Mode	Noninverting Mode
1.*	$\frac{\Delta G_I}{G_I} \approx \frac{1}{A_I\beta_I}\left[\frac{\Delta A_o}{A_o} - \frac{2\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}\cdot\frac{\frac{\Delta R_i}{R_i}}{1+\frac{\Delta R_i}{R_i}} - \frac{\epsilon_o}{1+\epsilon_o}\cdot\frac{\Delta R_o}{R_o}\right] - \frac{\frac{\Delta R_G}{R_G}-\frac{\Delta R_F}{R_F}}{1+\frac{\Delta(R_G+R_F)}{R_G+R_F}}$	$\frac{\Delta G_N}{G_N} \approx \frac{1}{A_N\beta_N}\left[\frac{\Delta A_o}{A_o} - \frac{2\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}\cdot\frac{\frac{\Delta R_i}{R_i}}{1+\frac{\Delta R_i}{R_i}} - \frac{\epsilon_o'}{1+\epsilon_o'}\cdot\frac{\Delta R_o}{R_o}\right] - \frac{\frac{\Delta R_G}{R_G}-\frac{\Delta R_F}{R_F}}{1+\frac{\Delta(R_G+R_F)}{R_G+R_F}}(1-\beta_o)$
2.†	$\frac{\Delta G_I}{G_I} \approx \frac{1}{A_I\beta_I}\left[\frac{\Delta A_o}{A_o} - \frac{2\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}\cdot\frac{\Delta R_i/R_i}{1+\Delta R_i/R_i}\right]$	$\frac{\Delta G_N}{G_N} \approx \frac{1}{A_N\beta_N}\left[\frac{\Delta A_o}{A_o} - \frac{2\beta_o\epsilon_i}{1+2\beta_o\epsilon_i}\cdot\frac{\Delta R_i/R_i}{1+\Delta R_i/R_i}\right]$

* Includes second-order effects.
† First-order approximation.

Under the circumstances specified by (43) and (44) or (45), the expressions for gain stability can be simplified to those given in line 2 of Table III.

Sensitivity considerations in network synthesis using operational amplifiers

Network and pole sensitivity. It can be shown that the sensitivity of a network function $T(s)$ containing simple zeros (z_i) and poles (p_j) in the complex frequency plane can be expressed as follows:

$$S_x^{T(s)} = \frac{dT(s)/T(s)}{dx/x} = S_x^K - \sum_{i=1}^{m}\frac{\mathcal{S}_x^{z_i}}{s-z_i} + \sum_{j=1}^{n}\frac{\mathcal{S}_x^{p_j}}{s-p_j} \tag{46}$$

where
$$\mathcal{S}_x^{z_i} = \frac{dz_i}{d[\ln x]} = \frac{dz_i}{dx/x} \tag{47}$$

$$\mathcal{S}_x^{p_j} = \frac{dp_j}{d[\ln x]} = \frac{dp_j}{dx/x} \tag{48}$$

$$S_x^K = \frac{1}{K}\cdot\mathcal{S}_x^K = \frac{d[\ln K]}{d[\ln x]} = \frac{dK/K}{dx/x} \tag{49}$$

and $n \geq m$

Equations (47) and (48) give a measure of the displacement in the s-plane of a zero (z_i) or pole (p_j) caused by the relative change in a particular network element x. They therefore define the zero and pole sensitivities of a network. Similarly, Eq. (49) represents the sensitivity of the scaling factor K. Thus, the overall sensitivity of a network function to incremental variations of an element x is a partial fraction expansion in which all critical network frequencies become poles and the residues are the pole and zero sensitivities.

It is well known that one of the major problems in active RC network synthesis is that of minimizing the network sensitivity as given by Eq. (46). The two main reasons for this are that (1) the individual network elements, particularly the active ones, may drift appreciably with ambient variations; and (2) in contrast to passive networks, active networks are only conditionally stable—that is, the location of their transmission poles are not limited to the negative s-plane.

One method of minimizing network sensitivity is to select the zeros (z_i) and poles (p_j) of (46) specifically for this purpose. However, this freedom does not usually exist since the zeros and poles are determined by the specified network function $T(s)$. This leaves only the residues—that is, the root sensitivities—to be minimized. These should be considered individually, depending on the position of the respective roots in the s-plane. In particular, the sensitivities related to the network poles should be minimized or at least adequately controlled since zeros in the vicinity of the $j\omega$-axis can be designed to depend entirely on stable passive components, whereas similarly located poles also depend on active elements; moreover, the location of the dominant network poles is highly critical, since drift on their part may cause severe underdamping or even oscillation of the network.

FIGURE 7. Pole displacement in the s-plane for networks containing uniformly varying RC components and one active component.

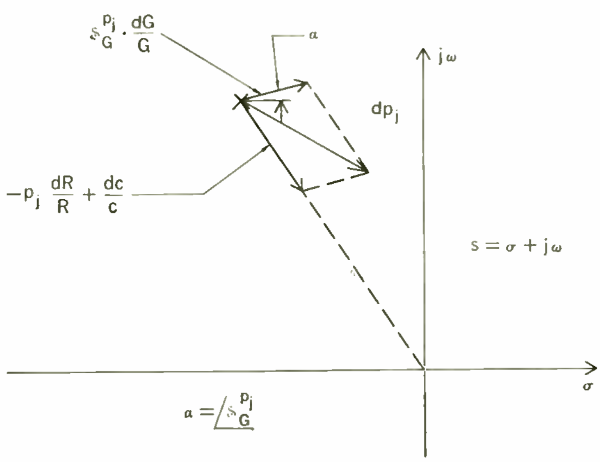
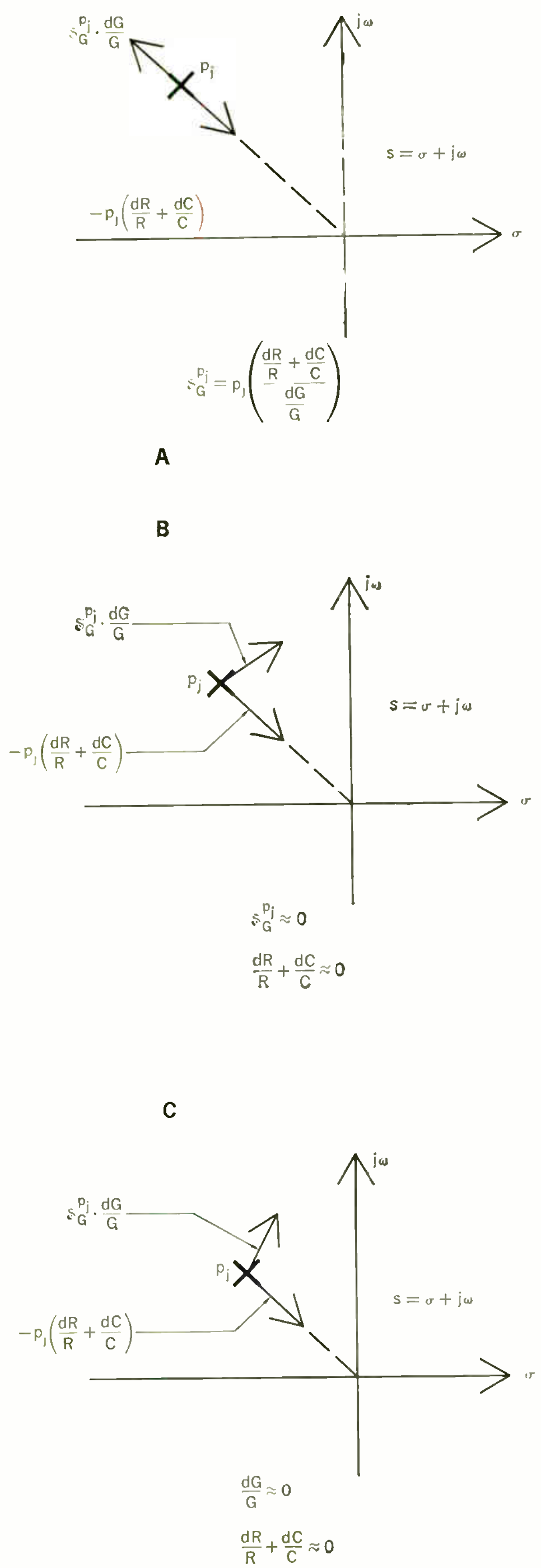


FIGURE 8. Methods of pole desensitization for networks with uniformly varying components and one active component. A—Method 1. B—Method 2. C—Method 3.



Thus, in order to ensure the stability of an active *RC* network, the sensitivity of the dominant poles primarily must be minimized. According to (48), the total drift of the *j*th pole due to drift in all *c* capacitors, *r* resistors, and *g* active elements of an active *RC* network is given by

$$dp_j = \sum_{i=1}^{r} S_{R_i}^{p_j} \frac{dR_i}{R_i} + \sum_{i=1}^{c} S_{C_i}^{p_j} \frac{dC_i}{C_i} + \sum_{i=1}^{g} S_{G_i}^{p_j} \frac{dG_i}{G_i} \tag{50}$$

If we assume that the passive-component variations are uniform,* and that the effects of the active elements can be lumped into those of a single equivalent active device, it can be shown that (50) simplifies to

$$dp_j = S_G^{p_j} \frac{dG}{G} - p_j\left(\frac{dR}{R} + \frac{dC}{C}\right) \tag{51}$$

The pole displacement given by this expression can be respresented by a vector diagram in the *s*-plane, as shown in Fig. 7. It can be easily verified that the displacement due to the passive elements is in a radial direction, whereas that due to the active element is tangential to the root locus with respect to that element. In other words, S_G^{pj} and pole displacement due to G have the same phase angle in the *s*-plane). If the pole sensitivity is to be zero, then from (51) the following condition must be satisfied:

$$S_G^{p_j} \frac{dG}{G} - p_j\left(\frac{dR}{R} + \frac{dC}{C}\right) = 0 \tag{52}$$

This expression provides several means of desensitizing an active *RC* network to variations in network elements. It will be shown next that one of the big advantages of active-network synthesis using operational amplifiers lies in the fact that they allow for a very simple and effective method of pole desensitization.

Pole desensitization. Equations (51) and (52) indicate the following three methods of desensitizing the dominant poles of an active *RC* network.

Method 1. Pole desensitization is achieved here by designing the network in such a way that the passive pole displacement is compensated by the active pole displacement; that is,

$$S_G^{p_j} = p_j\left(\frac{dR}{R} + \frac{dC}{C}\right)\bigg/\frac{dG}{G} \tag{53}$$

This method is illustrated in Fig. 8(A). In this case the pole location is dependent on both passive and active network elements. Pole-drift compensation can be achieved by a network configuration combining several forward transmission paths in single or multiloop feedback structures.[6] This compensation method is somewhat complicated and not very general. The chosen feedback configuration depends both on the physical properties of the active and passive components being used and on the location of the particular pole being desensitized. The requirement imposed by Eq. (53) is restrictive enough that it severely limits the choice of network configurations capable of satisfying this requirement and providing the characteristics of a specified network function at the same time.

This method is not suitable for network synthesis with

* This assumption is quite valid for integrated circuits, which are of primary interest here, since both thin-film and integrated components have very good tracking capabilities.

operational amplifiers since the gain characteristics of the amplifiers must be individually controllable or variable with ambient variations—which they usually are not.

Method 2. In a second desensitization method the network configurations used are those whose critical frequencies (for example, dominant poles) depend only on the passive network elements. The effect of the active elements is negligible because the pole sensitivity to the active element can be made arbitrarily small. The poles are desensitized in the same manner as those of passive *RC* networks—that is, by using resistors and capacitors with uniformly equal but opposite temperature coefficients. Thus, as shown in Fig. 8(B),

$$S_G^{p_i} \approx 0 \tag{54}$$

and

$$\frac{dR}{R} + \frac{dC}{C} \approx 0 \tag{55}$$

Typical networks that permit the realization of (54) are negative-feedback configurations and positive-feedback configurations with unity forward gain.[7] In both cases, the pole sensitivity to G can be made arbitrarily small by the use of amplifiers with sufficiently high loop gain. Obviously, the operational amplifier is ideal for this approach. It is therefore no coincidence that some of the oldest and best-known methods of active *RC* filter synthesis[8,9] utilize the operational amplifier to obtain stable network characteristics in this way.

This method is very effective in decreasing network sensitivity. However, it deprives the designer of the versatility afforded by the active network parameter since he is limited in his choice of network configurations to those that are capable of realizing (54). This limitation is similar to, but by no means as severe as, that specified in Method 1 by (53). Thus, even here the scope of realizable transmission functions is limited. Furthermore, high-Q networks require a wide spread of resistive or capacitive component values—incompatible both with thin-film and semiconductor integrated processing techniques.

Method 3. As in Method 1, the critical transmission frequencies are permitted to depend both on passive and active network elements. However, the active and passive pole displacements are minimized independently, as in Method 2. The passive pole displacement is compensated in the usual way by the use of resistors and capacitors with uniformly equal but opposite drift properties. However, in contrast to the two preceding methods, the active pole displacement is not minimized by placing any constraint on the corresponding pole sensitivity but by minimizing drift in the active element itself. Thus, in this case [see Fig. 8(C)], both (55) and the following expression must apply:

$$dG/G \approx 0 \tag{56}$$

The active element can be stabilized—that is, (56) can be satisfied—by a local negative-feedback network consisting of passive components with tight tracking properties. Since this stabilizing process is much more effective when the available open-loop gain of the active element is as high as possible, it is clear that operational amplifiers are ideal for its application. However, since the available closed-loop gain of an amplifier is greatly reduced by individual feedback, more than one amplifier is often required in high-Q applications.

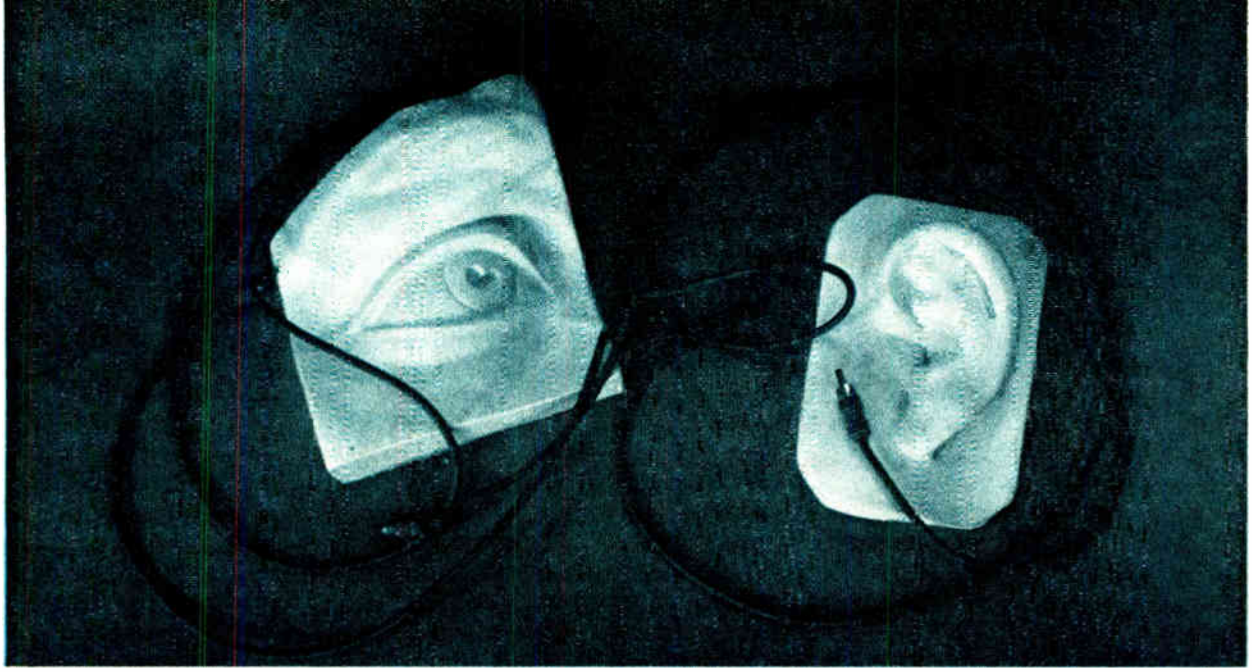
This method of desensitization has been found very effective in the design of high-selectivity networks combining silicon integrated operational amplifiers with high-precision tantalum integrated passive components.[10,11] The temperature coefficients of the tantalum thin-film resistors and capacitors of the frequency-determining networks can be matched closely over a given temperature range, and the semiconductor amplifiers can be stabilized by tantalum thin-film resistors that track closely. Thus, this hybrid integrated technology lends itself particularly well to the constraints required by this method, as specified in (55) and (56).

To summarize, of the three methods of pole desensitization discussed, Method 3 is the most practical and affords the most design flexibility. However, both Method 3 and Method 2 can only be as successful as the quality of the available passive components permits. With both methods, desensitizing becomes a technological as opposed to a circuit-theory problem. It is of interest to note that this puts sensitivity back into the domain it was in previously with passive *LCR* networks.

REFERENCES

1. Giles, J. N., *Fairchild Semiconductor Linear Integrated Circuits Applications Handbook*, Fairchild Semiconductor, 1967.
2. *RCA Linear Integrated Circuit Fundamentals*, RCA Corp., Harrison, N.J., 1966.
3. Stata, R., *Operational Amplifiers*, pts. I–IV, Analog Devices, Cambridge, Mass., 1965.
4. Bode, H. W., *Network Analysis and Feedback Amplifier Design*, Princeton, N.J.: Van Nostrand, 1945.
5. Ur, H., "Root locus properties and sensitivity relations in control systems," *IRE Trans. Automatic Control*, vol. AC-5, pp. 57–65, Jan. 1960.
6. Gaash, A. A., Pepper, R. S., and Pederson, D. O., "Design of integrable desensitized frequency selective amplifiers," *ISSCC Dig. Tech. Papers*, Feb. 1966, p. 34.
7. McVey, P. J., "Sensitivity in some simple RC active networks," *Proc. IEE (London)*, vol. 112, July 1965, pp. 1263–1269, July 1965.
8. Scott, H. H., "A new type of selective circuit and some applications," *Proc. IRE*, vol. 26, pp. 226–235, Feb. 1938.
9. 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.
10. Moschytz, G. S., "Miniaturized filter building blocks using frequency-emphasizing networks," *Proc. NEC*, vol. 23, pp. 364–369, 1967.
11. Moschytz, G. S., "FEN filter design using hybrid integrated building blocks," Paper 4–5, 1969 WESCON.

George S. Moschytz (M) received the M.S. and Ph.D. degrees in electrical engineering in 1958 and 1960, respectively, from the Federal Institute of Technology, Zurich, Switzerland. His Ph.D. dissertation involved the development of a photoelectronic scanner for the automatic sorting of coded envelopes for use by the Swiss Postal Service. From 1960 to 1962 he worked for RCA Ltd. in Zurich on color television signal transmission problems and, in particular, on new high-precision methods of envelope delay measurement in video transmission systems. Since 1963 he has been with Bell Telephone Laboratories, Inc., Holmdel, N.J., where he is now supervisor of the Data Systems and Circuits Group. At BTL he has investigated methods of synthesizing linear and digital circuits for use in data-transmission equipment that can be microminiaturized by combining thin-film and silicon integrated circuit elements. Dr. Moschytz is the author of several papers on active-network synthesis and design and also holds a patent on the design of active-filter building blocks.



What's ahead in communications?

Future advances in communications are expected to become more and more sophisticated. As a result, the increasing expenditures of time and money may cause investors seriously to consider simulating nonexistent technologies in order to evaluate their social effects and values

J. C. R. Punchard Northern Electric Laboratories

Since the publication of the book "The Year 2000" by Kahn and Wiener, it has become popular for prophets to make predictions about what is going to happen in the next 30 years. This article discusses some possibilities and limitations, as they apply to the Canadian telecommunication industry. No attempt is made to forecast accurately future developments in detail.

Since the end of the century is just as far ahead as the beginning of World War II is behind, a thoughtful observation of developments during the last 30 years can help to determine in a general way how we can expect communications in the next 30 years to unfold.

Electronics was developing rapidly at the end of World War II and, by 1948, Bardeen, Brattain, and Shockley had invented the transistor at Bell Telephone Laboratories. This device turned the already steeply rising index of electronic application into an explosive exponential climb. We know now that this invention began a line of development that had, and will have, far greater significance than its illustrious forerunner, the vacuum tube. The original transistors, however, were costly, difficult to make, noisy, and unreliable. Many years of intensive research and development ensued before they were practical for introduction into commercial equipment during the last half of the fifties and early sixties. These devices and their modern derivatives have made possible sophisticated equipment undreamed of even in 1950. A forecaster in 1939, looking ahead to 1969, could not have anticipated the transistor, and therefore could not have accurately predicted the immense strides made in the computer, space, entertainment, telephone switching, and transmission fields. He could have predicted the exploitation of broadcasting and high-frequency radio and television, because these were derived mainly from vacuum-tube technology. No one in 1939 was seriously talking about sending men to the moon. It is quite safe to say that the Apollo 11 moon landing simply could not have been accomplished in 1969 using the available vacuum-tube technology.

From the past history of invention we can learn several facts. The sudden appearance of a radical surprise invention such as the vacuum tube or transistor does not immediately make all previously used communication equipment obsolete, for several basic reasons:

1. Several years are usually required to refine any basic invention before widespread introduction is accomplished. Further time is required to perfect manufacturing and bring costs down to reasonable levels.

2. Most important is the magnitude of the present in-

vestment in plant, which simply cannot be discarded just because a new technology appears.

3. Resistance to change and acceptance of new devices are aggravated by the reluctance to commit large blocks of capital toward provision of new manufacturing processes and plant.

4. Human reaction to new technology and the speed at which men are willing to accept new ideas are entirely unpredictable, even when the economics are sound.

This same reasoning will apply to future surprise inventions and, if past experience is a good teacher, these new inventions will be increasingly sophisticated and require even more time and money before they can be applied in practice. We are now so heavily loaded with advanced products and services that it is difficult to visualize new communication items that would bring about major changes in modes of living. Such items as three-dimensional television, electronic banking and mail service, and home facsimile might be improvements, but hardly cause for drastic changes. However, there may well be surprise inventions that could seriously change our way of life. I refer to further extensions of man's senses—the transmission of touch, taste, and smell. If invented, these systems will likely be very highly sophisticated and therefore will take many years to bring into practical use. You already know how the transmission of sight and sound has altered your life. I leave it to you to imagine the effects that can come about with transmission of touch, taste, and smell!

It is clear that any look into the future cannot be guided solely by predictions concerning technological development. Perhaps more important, and certainly more unpredictable, are human reactions to technological change.

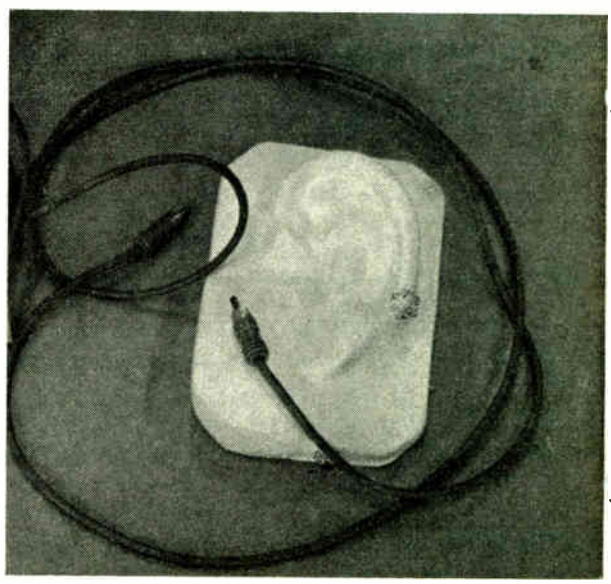
The first major derivative of the transistor is the integrated circuit, which is now being exploited the world over. Such devices are bringing about a revolution in the computer industry and will contribute heavily to the improvement of communication systems in the future. It is to be noted, however, that the monolithic IC is capable of providing right now many services that are not available, and likely will not be available for some time to come. Dick Tracy's wristwatch radio and personalized radio sets for two-way telephone communication with two-way signaling are but two examples. Although both are ruled out at this time for economic reasons, no one really knows what the market acceptance will be, and we are doing little or nothing to increase our ability to predict such acceptance.

Gordon B. Thompson at Northern Electric Labs is advocating some form of human-science research to determine the effect of new devices and systems on the population before commencing programs for such device development. He maintains that the long-run value of a new product or service cannot always be determined by a simple market acceptance survey and economic study. A most important measure of the real worth of a new product is its ability to contribute to society's well-being and/or productivity, resulting eventually in an increase in the nation's gross national product. Thompson also stresses the importance of research work to discover new concepts of communication services, literally, the invention of new businesses. These are radically new thoughts, since they involve doing research in the human and social sciences to determine what technological research programs should be undertaken.

The monolithic IC and its derivatives have already assured themselves a most important place in the future, unless some unlikely "surprise" invention supersedes them. They will find ever-increasing application as design and production processes are improved with corresponding cost reductions, especially for small-quantity runs. Most of their use to date has been in digital systems as found in computers, television receivers, and military equipment. Earth-based analog transmission systems are just now beginning to use integrated circuits extensively as costs are reduced and techniques improved. We can also anticipate a rapid increase in the demand for integrated circuits as digital transmission methods such as pulse-code modulation come into more common use. Because the additional cost of adding a few more transistors or diodes in any one integrated circuit is abnormally small, we can expect to see the development of more and more complex circuitry to accomplish functions entirely uneconomic with discrete devices. Simulation of inductance and capacitance, within limits, are examples.

Medium-scale integration (MSI), in which 20 or 30 integrated circuits are combined on one monolithic chip, is the next derivative of the basic integrated circuit. These are now being introduced into equipment designs for production. On the basis of reasonable cost, reliability, and size, this technique will contribute heavily in both the computer and transmission fields. The next and most exotic technique in an unfolding technology is large-scale integration (LSI), in which hundreds of integrated circuits are combined, with interconnection, on one monolithic chip. This development is now in the laboratory and shows promise for use in memory, logic, and perhaps switching devices. However, it's a long way from the laboratory to the customer's floor, and widespread use of LSI will probably require a long period of evolution. The computer industry will likely be the first fully to exploit its potential, but the telephone industry is already eying the possibilities.

The hybrid microstrip technique, which is an outgrowth of strip line, has proved to be reasonable in cost, lends itself to rapid circuit updating, and is attractive for



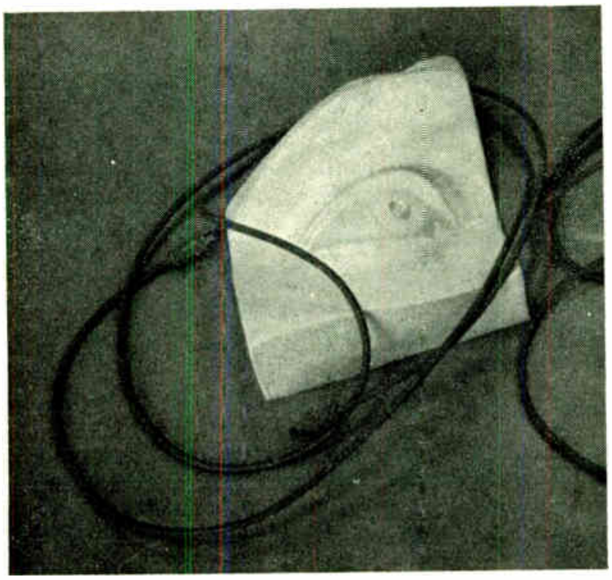
ground-based, airborne, and satellite equipment. It shows considerable promise for bringing about drastic reductions of size and weight of microwave apparatus.

Hybrid thick-film circuitry offers advantages for low-cost designs. One would expect it to be used extensively in domestic-type electronic equipment and other applications where rapid updating is a desirable attribute.

Thin-film circuitry is a more exotic technique, desirable where high reliability and circuit element accuracy are required. High initial capital costs may limit its use to special volume applications such as telephone sets, which are expected to increase in complexity but not in cost as new types of services are offered.

Technology is now available to provide many new services such as push-button supermarket shopping, remote control of household appliances through the telephone network, and home electronic video recording. Again, the measure of the worth of these services and devices depends not only on economics, not only on fast market surveys to determine immediate acceptability, but on long-term answers to such questions as: "Do the ladies have more to gain by pushing buttons at home than by indulging in the pleasure of shopping in a supermarket?" "Will remote control really contribute to society's sense of personal growth, well-being, and productivity?" "Will electronic video recording change personal living habits, television programming, etc., and if so, will such changes in some direct or indirect way increase the gross national product?" Without research in social or human science, we really don't know the answers. We have very little, if any, factual information for trying to predict how humans will react to the possible intrusion of privacy through the introduction of, say, the "checkless" society or the development of computer data banks of readily accessible personnel information. If the reaction is negative, sophisticated technology alone cannot guarantee a successful service.

Without completely designing and building such systems, it is possible to simulate their performance, whether or not the technology exists. Under the proper research conditions, the resulting social effects and social values can be studied and evaluated to determine whether or not major research programs will be worthwhile and should be undertaken. The cost of this kind of research will be significant, but small compared with the long-term commitment for capital and expense dollars required for development and implementation. Today's managers generally do not readily accept such radical thinking but, when the future stakes are so high, can we afford not to attempt to measure the results of technical innovations that have a direct affect on the individual?



One major development we can be certain is coming is the broadband switched network—the so-called "wired city." It is evident that markets are growing for services requiring wider and wider bandwidths. Although dedicated circuits for wide-band transmission are available now, the ultimate in service requirement for flexibility cannot be achieved without automatic broadband switched networks. Although it is easy to visualize a coaxial cable system as comprehensive as the present voice-band telephone network, its implementation will be several orders of magnitude more expensive. Since Canadian telephone companies and other common carriers already have a plant investment of around $5.5 billion, it is evident they cannot quickly retire such an investment and replace it with something more exotic and expensive just because the technology for wide-band transmission *can be developed*. The justification for large expenditures such as required to build the "wired city" concept of a broadband switched network, must be based on the potential increase in productivity and ability to innovate in the society at large. This concept will not be viable from a business standpoint unless it meets this criterion.

It is inevitable that such systems will come into being, but only as a result of years of further research and development and very careful physical, economic, and social planning. Let us also remember that whatever is done must be compatible with existing networks. Although we hear a lot about digital transmission systems using advanced designs of present pulse-code modulation equipment, there are still a lot of unanswered questions concerning the technical and economic feasibility of space- or time-switching methods. At the moment, no one really knows exactly what is involved in terms of total network requirements. In reality, we do not yet have the necessary technology to engineer viable broadband switched systems, but this is just a matter of time. Dedicated radio systems using digital methods are of course firmly established, especially for deep-space communication. We can expect many new developments in digital techniques in the next few decades. But when we ponder on switched broadband networks, let's not fool ourselves—the road will be long and expensive.

The possibility of satellite communication in many forms has excited a great deal of current interest. Engineers are talking of direct television broadcasting from nuclear-powered synchronous satellites as if they will be here tomorrow. The production of kilowatts of radio-frequency power in the sky is only a matter of time, but its achievement will not guarantee that efficient and profitable systems can be built. There still will be many administrative and human problems to be solved, involving investment in existing ground-based equipment, local advertising, time zones, and somebody's new dollars.

The future of communication satellites beyond the first domestic systems now being designed is somewhat hazy because the use of repeaters in the sky will be limited by economics rather than by technology. The program for the Canadian domestic satellite is a matter of government policy based on staking a claim for parking space and for improved communications to Canada's north. However, as technology improves and costs are reduced, this type of repeater will become increasingly more desirable as we move into the seventies, and we would expect an expanding market for communication satellites.

Educational television spattered from above is perhaps the most inefficient way of handling the learning process from a technical standpoint. Ground-based information-retrieval systems on coaxial cable networks that can be more heavily loaded show much more promise, although even these systems have yet to be evaluated by the educators.

Even educational television by way of terrestrial microwave systems seems wasteful of frequency spectrum space while providing the minimum of channel capacity. It is estimated that 250–500 television channels will be required to service the schools in the city of Ottawa alone. On the subject of education, a number of organizations are hard at work developing prototype input and output devices for computer-aided learning together with experimental software. Good progress is being made, but in the long run the educators must rule on whether such systems will come into being on a nationwide basis.

Conjecture about future developments should not fail to mention the use of laser pipes for long- or short-distance transmission of thousands upon thousands of television channels, and voice and data channels in the hundreds of thousands. The demand for channel capacity in Canada is growing rapidly as the country expands, but it is most unlikely we will need the kind of capacity provided by laser pipes before the year 2000. In addition, the technology is a long way from being developed, since completely satisfactory modulators, isolators, amplifiers, and demodulators simply do not exist today. The problems associated with provision of optically perfect pipes, kilometers in length, with suitable reflectors are simply horrendous in Canada's climate. The possibilities of multiple coaxial cables and waveguides are a long, long way from being exhausted at this time. Only when these become uneconomical will the laser pipe seem attractive.

It is hard to see how anyone can be anything but optimistic about the future of the computer since it is becoming almost as essential as bread in our society. We can surely look forward to great expansion in this field until eventually most homes will be equipped with computers or computer service, provided the homeowner will accept these machines as readily as he accepts telephones and television. After all, the acceptance of the telephone, which is one of the worst invaders of privacy, is purely an accident of human nature caused by curiosity.

The availability of computers and the improvements in this field have caused a great deal of thinking about information-retrieval systems over and beyond the present on-line computer facilities now springing up like dandelions. These information systems, sometimes referred to as data banks, are complex and costly and will need impetus by the government before they can be implemented on a national scale. Their advent is inevitable because of the increasing volume of literature necessary to carry out technical and managerial jobs. The Science Secretariat of Canada is now proposing a national system but so far industry has not jumped at the opportunity to contribute to its cost. Perhaps it is just too far ahead of its time and our pocketbooks!

Only a few of the many developments in the communication art have been mentioned here, but let us reiterate. New discoveries generally take 10 to 20 years to come to full commercial fruition and, therefore, the expansion of communications in the next 30 years will be determined by what happens in the next decade. The blossoming of new types of communication services will depend more on economics, government policies, or human reactions than on technical feasibility. The telecommission studies recently announced by the Minister of Communications, Mr. Kierans, will play a most important role in shaping Canada's future communications.

We can make one solid prediction. The future of communications shows great promise but, because it involves the extension of man's senses, the kinds of communication services provided eventually will affect in one way or another the life of every citizen. The phonetic alphabet, the printing press, the telegraph, the telephone, radio, and now television have each changed man's mode of living in giant leaps. No planning of the effect on human beings attended the introduction of these advances in the coupling of the minds of men, mainly because research in the human sciences lagged so far behind current scientific knowledge.

Today, we are witnessing the beginning of research studies to determine the long-range effect of technology on society. Time alone will tell whether or not this approach is practical compared with the obvious supply–demand economic approach. In any case, when we design and introduce communication systems today, we may well be designing the men of the future. It behooves us to plan these systems intelligently and with the best insight that can be mustered, without emotion and with our feet firmly planted on the ground.

Essentially full text of a paper presented at the International Electronics Conference and Exposition, Toronto, Ont., Canada, October 6–8, 1969.

J. C. R. Punchard (F) received the bachelor of applied science degree in electrical engineering from the University of Toronto in 1933. After "demonstrating" at the university, he joined Northern Electric Company, Ltd., Montreal, Que., Canada, where he was engaged in transmitter design and point-to-point radio system development. In 1944, he obtained a leave of absence to develop radar equipment at Research Enterprises, Ltd., Toronto, Ont.; upon returning to Northern Electric, he was appointed chief engineer of the Belleville, Ont., Electronic Division, responsible for Pinetree radar system development and production. He became manager of the Belleville plant in 1954. Transferred to the Research and Development Laboratories of the company in 1960, Mr. Punchard has held the position of director of transmission development in addition to the position of director of research. He is now assistant vice president responsible for liaison with government, universities, technical societies, and other organizations. A past Director of IEEE Region 7, he is a member of the Association of Professional Engineers of Ontario and the Engineering Institute of Canada, and is president of the Canadian Radio Technical Planning Board.

Transistorized ground-fault interrupter reduces shock hazard

Even the home outlet can deliver a lethal electric shock. Yet for a decade there have been simple sensitive devices for ridding us of such dangers

Charles F. Dalziel University of California

The ground-fault interrupter is the most successful device for eliminating the hazard from low-voltage electric shocks in the home, on the farm, and in industry. It is the newest of four recognized means—isolation, insulation, grounding, and shock interruption—for reducing the danger from electric shock and it is, by far, the most radical.

The ground-fault interrupter, or GFI, is a device that interrupts an electric circuit when the fault current exceeds a predetermined value less than that required to operate the overcurrent devices of the circuit. (In Europe this device is called an earth-leakage circuit breaker.) Such apparatus have been used to protect high-voltage power lines since the 1920s, and they were set to operate at 10 to 20 percent of the minimum operating current, or trip value, of the associated overcurrent devices. Thus, a power circuit breaker having an overload trip value of 200 amperes was set to trip on ground faults of only 20 to 40 amperes, which was considered a great achievement of the day. Some ten years later the importance of protecting against low-voltage burndowns in industrial equipment was recognized in Germany and devices were developed having a line-to-ground-trip value of about 500 mA.

More recently, about ten years ago, both the French and the Austrians developed two-wire earth-leakage circuit breakers having a trip value of 25 to 30 mA.

The French–Austrian developments were followed in the U.S.A. in 1962 by the development of two- and three-wire transistorized ground-fault interrupters having a continuous rating up to 100 or 200 amperes and a ground-current trip value of 5 mA. This performance parameter means that the circuit breaker will trip in the event of a 5-mA line-to-ground fault current. In the case of an accident, this would be the current through the victim. The 5-mA level is prescribed both by Underwriters' Laboratories, Inc. (U.S.A.), and by the Canadian Standards Association.

The operating time of these devices is so fast, and the corresponding shock energy so small, that the modern GFIs virtually eliminate electrocutions, burndowns, and fires due to currents flowing to ground. However, it must be recognized at the outset that the sensitive ground-fault indicating mechanism does not respond to line-to-line or three-phase faults unless zero phase-sequence currents are involved.

To verify that the GFI really will prevent electrocution, Dr. Archer S. Gordon, of Statham Instruments, Inc., Oxnard, Calif., administered 2400 shocks to experimental dogs under anesthesia. A commercial 5-mA GFI was used, and dogs were connected in series with the "hot" wire of the 120-volt laboratory circuit and ground. The dogs were given 800 shocks using a current pathway between right forepaw and left hind paw to simulate the frequently experienced arm-to-leg pathway

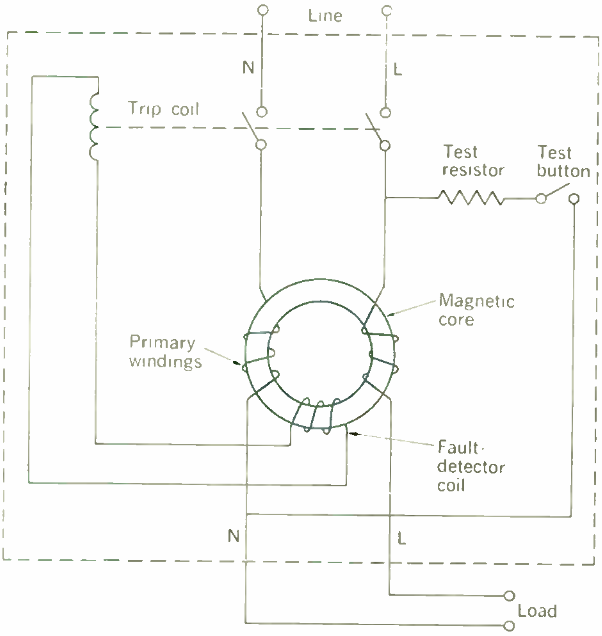


FIGURE 1. Schematic diagram, French device.

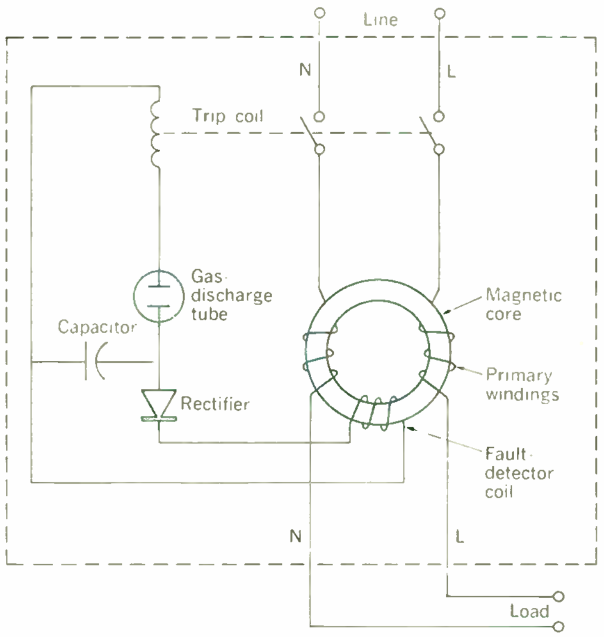
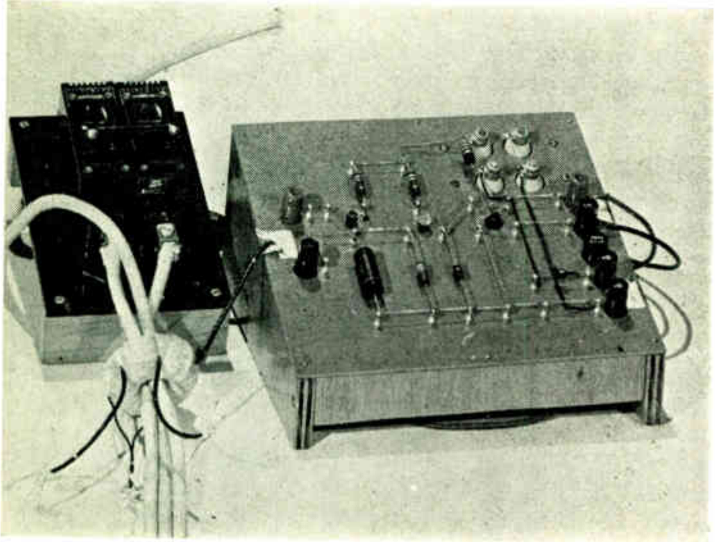


FIGURE 2. Schematic diagram, Austrian device.

in many human electrocutions. No incidence of ventricular fibrillation was observed. Eight hundred shocks were then given to the dogs after electrodes were placed on the right forepaw and left forepaw. None of these 800 shocks produced ventricular fibrillation. However, 36 fibrillations were produced during the course of 800 shocks applied with electrodes placed on opposite sides of the chest. This result is not important from a safety viewpoint, since such a pathway is unlikely in human accidents. Moreover, since the minimum current for producing ventricular fibrillation in mammals is approximately proportional to body weight, it is evident that the GFI will protect human beings, including the very young.[1]

FIGURE 3. First United States GFI.



It is pertinent that during the past four years several technicians and salesmen have received 120-volt shocks incidental to installing, selling, or demonstrating GFIs, and no injuries occurred. European experience during the past nine years has been equally good, and many lives have been judged saved by the operation of 30-mA, 220-volt, two-wire devices.

Types of sensitive ground-fault interrupters

Among the several types of interrupters are the following:

Direct trip. The direct-trip type is used in Europe; it obtains its operating power from the shock or fault current alone, by means of a differential transformer, which is the leakage-current or ground-fault-current sensing mechanism. The device has the advantage of being entirely self-contained, and it can be used where the distribution system is grounded at only a few places—such as a system grounded only at the substation. It is also low cost. However, it has a limited 30-A ampacity. Also the tripping reliability has been questioned because of the lack of robust tripping power and the necessarily delicate mechanisms. This genre comprises:

The (predominantly) mechanical type, which consists of a differential transformer and a polarized relay-trip coil. Electricité de France has purchased some 30 000 units, manufactured by l'Industrie Electrique de la Seine, for use as entrance switches for residential customers. The unit is essentially instantaneous in operation, has a minimum trip value of 30 mA, a full-load rating of 30 amperes at 50 Hz, 220 volts, and is two-wire, single phase.

Figure 1 shows the schematic diagram of this French device. A sensitivity of 30 mA is achieved by the use of a polarized, spring-loaded relay trip coil. The device operates on the principle of magnetic balance. As long

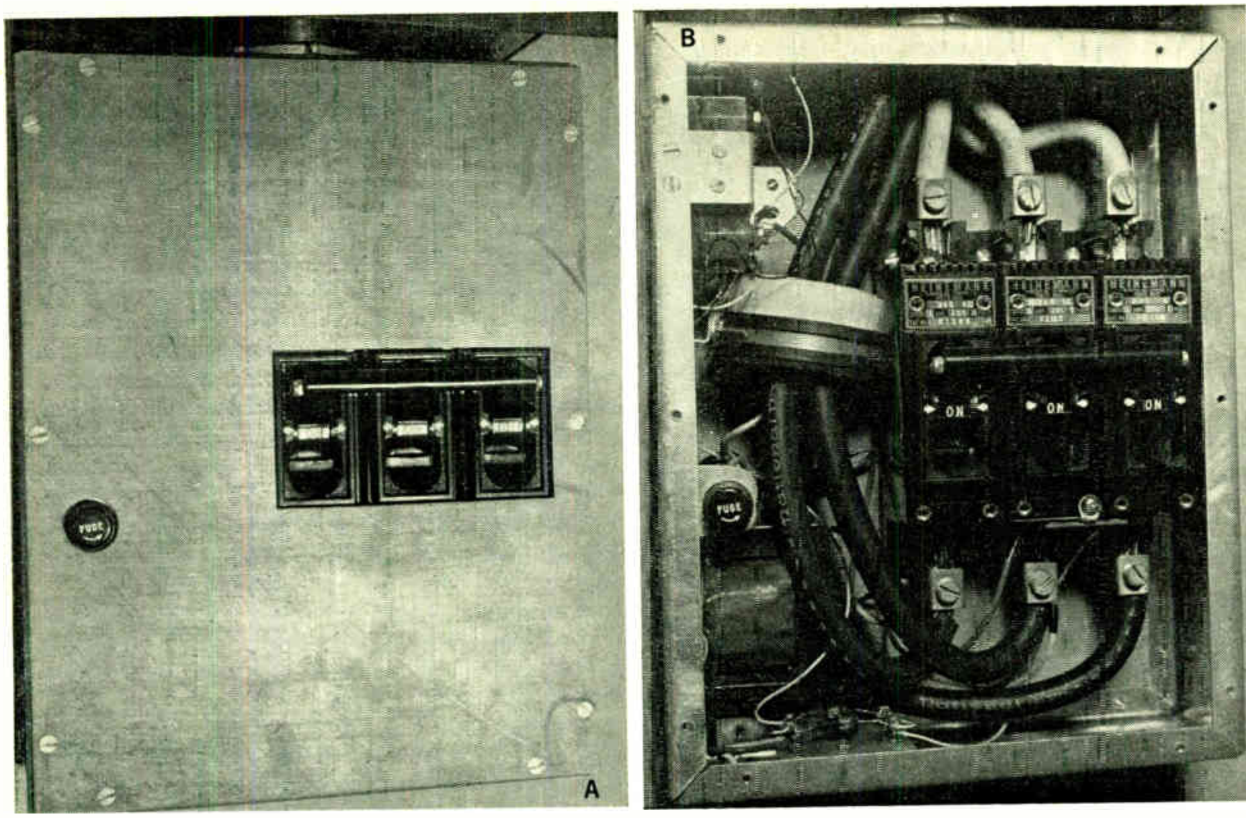


FIGURE 4. GFI protecting author's home since 1962; (A) cover in place, (B) cover removed.

as all of the current from the incoming lines flows in at *L* and exactly the same current flows out at *N*, there will be no flux in the transformer core and no voltage will be induced in the fault-detector coil. However, should some current be diverted to ground beyond the load terminals, the magnetic balance will be upset and a flux will be created in the core, which will induce a voltage in the secondary winding. If this induced voltage is sufficient, the device will operate and trip.

The electrical type, which consists of a differential transformer, and a single or multistage solid-state rectifier–capacitor voltage multiplier that fires a gas tube, which energizes the trip coil of the circuit breaker. Although not quite so delicate as the "mechanical" type, it has considerable inherent time delay. Manufactured by the Felten and Guilleaume Company of Vienna, it is rated 16 amperes, 500 volts, and has a nominal trip value of 30 mA.

Figure 2 shows the schematic of the Austrian earth-leakage circuit breaker. The device functions similarly to the French apparatus. However, sensitivity is achieved by the charge stored in the capacitor.

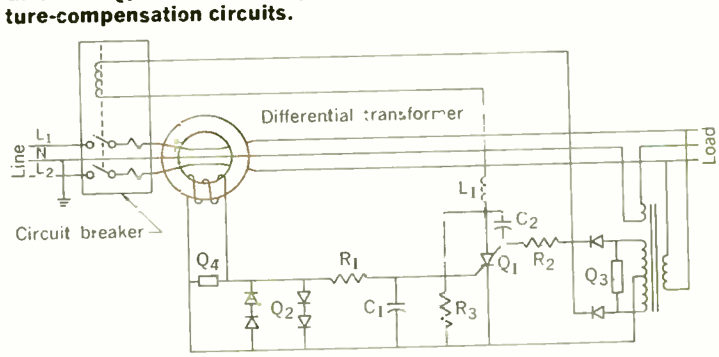
Power-operated. In the power-operated design, the differential transformer is used only to detect leakage current or ground-fault current, and tripping force is obtained from the electric power line. The devices are not limited in size, and can be readily designed for 15 to 200 amperes, 120 to 480 volts, two- to four-wire, one- or three-phase, and with very low trip values. Demonstration models are available having a trip value of only ½ mA. Trip values of commercial units start at 5 mA and some units have adjustable trip values from 10 to 100 mA. Their chief limitation is that they must be used only on multigrounded systems (as required in the U.S.A. by the National Electrical Code) so that a break in the neutral wire in the distribution system will not impair protection to customers located downstream from the severed neutral.

Figure 3 is believed to represent the first successful GFI constructed in the U.S.A. (1961).

Figure 4(A) shows the three-wire, 100-ampere, 120/240-volt, ground-fault interrupter that has been protecting this author's home since November 1962; it employs a single transistor.

Figure 4(B) shows the interrupter with the cover removed. Note the differential transformer, the epoxy-

FIGURE 5. Schematic diagram, three-wire transistorized GFI. Path Q_4 includes the trip-value control and temperature-compensation circuits.



encased sensing circuit, and the three companion trip circuit breakers. The series trip in the center or neutral circuit breaker has been replaced by a shunt trip coil but the series trips in the line circuit breakers have not been altered. Because of this design, the sensitive trip provides a line-to-ground feature in addition to that normally provided for overcurrent or overload protection.

There are at least four different types of power-operated GFIs:

The magnetic-amplifier type, manufactured by F. W. Industries, Ltd., is used in South Africa. The device is inherently large and heavy because of its magnetic cores. It is reported to have excellent operating characteristics, but has the reputation of being the most expensive of all GFIs.

The transistorized type is manufactured by The Rucker Company in the U.S.A. and in Canada, and by the Federal Pacific Electric Company in Canada. This device comprises a differential transformer, solid-state amplifier, and a circuit breaker.

A *reed-switch* actuated circuit breaker, manufactured by Harvey Hubbell, employs two magnetic cores. The unit is competitively priced.

A trickle-charged battery type assures power for tripping a circuit breaker for an extended period even after the main power supply has failed. A European firm is marketing a transistorized device of this design. This is the least expensive of all GFIs, but has the disadvantage that when the battery wears out, the device becomes inoperative without warning or indication; i.e., when it fails, it fails stone dead.

The transistorized GFI

Figure 5 illustrates a schematic diagram of the three-wire, 120/240-volt, modern, transistorized, highly sensitive GFI. As long as all of the current flowing in at L_1 returns through N or L_2, the core of the differential transformer remains unmagnetized and the voltage of the secondary winding is zero. However, if any current flows to ground beyond the load terminals, the magnetic balance is upset and a voltage appears across the secondary-winding terminals. When this voltage reaches about ½ volt, the solid-state switching device Q_1 becomes conducting and the circuit breaker is tripped by energy supplied from the control transformer. Such a device is capable of high sensitivity and large ampacity.

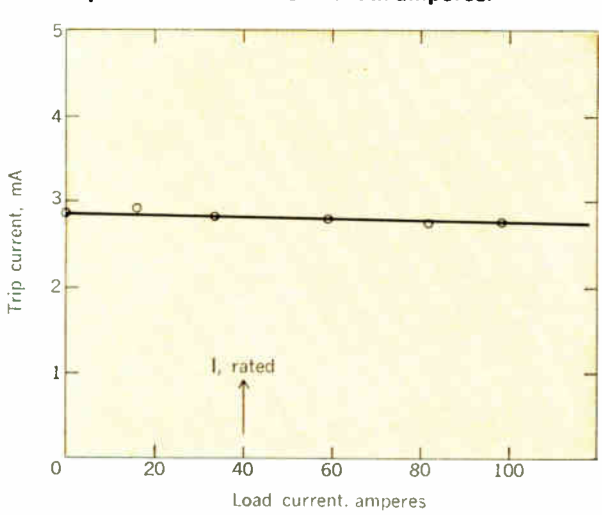
If, in a protected circuit, a person were to make contact between an energized defective appliance and a grounded object, causing a ground-current flow in excess of the trip value, the circuit breaker will be actuated and the electrocution hazard eliminated. However, the device is actuated only by the current flowing to ground, and it cannot differentiate between the body–appliance circuit and any other current pathway to ground having the same resistance. Thus, it is imperative that the normal or standing leakage current of the protected system be as small as possible. Obviously, standing leakage must be less than the trip value or the device will not remain closed. The insulation resistance of modern wiring is very high, and possibly the chief source of abnormal leakage is damp, poorly taped joints, and old wiring immersed in water-filled conduits, dirty terminal blocks, or defective appliances.

Although the National Electrical Code prohibits grounding the neutral wire on the load side of the service equipment, such grounds frequently exist, and such grounds will cause the GFI to trip out, much to the consternation of the electrician making the installation. Because of the required grounds on the utility system, such a ground would shunt some of the load current returning to the differential transformer, upset the magnetic balance, and cause variations in the trip value. To prevent such unwanted tripping, the neutral wire is fed approximately 100 mV from a third winding on the GFI control transformer. If extraneous grounds are present on the load circuits, it will be impossible for the GFI to remain closed, and the installer will be immediately aware that something is wrong.

Another requirement is that the device withstand high-current line-to-ground faults (presently prescribed at 5000 amperes rms) for at least the time it takes for the circuit breaker to clear the fault. The secondary or sensing circuit is protected from proportionately large currents: by saturation of the core of the differential transformer, by the secondary winding, which appears as a very high-impedance source (> 50 000 ohms), and by the symmetrical pair of double diodes, Q_2 that clip any transients to a voltage level within the voltage tolerance of the solid-state device, Q_1. Any steep transient—such as that due to lightning or switching surges—is further attenuated by the R_1 C_1 filter (that does not impair tripping time, but does shunt high-frequency, short-duration transient voltages from Q_1). These means have been found sufficient to eliminate false tripping due to transients coming in the "front door."

Eliminating erroneous tripping due to transients coming in the "back door"—i.e., through the control transformer—is a more difficult problem. Capacitor C_2 and resistor R_2 reduce the susceptibility of the anode gate to respond to unwanted signals. Response to transients is also reduced by inductance L_1 and resistor R_3 and voltage clipper Q_3, which becomes conducting and prevents the control transformer's secondary voltage from rising to unacceptable values.

FIGURE 6. The characteristics of a new differential transformer, illustrating the accuracy with which the device discriminates between leakage current to ground in milliamperes and the load current in amperes.



False or nuisance tripping

False tripping can be defined as an unwanted power outage as viewed from either the customer's, the installer's, or the engineer's viewpoint:

From the customer's standpoint, any power outage will usually be classified as a nuisance. The device may be performing its intended purpose in every respect, but if the housewife plugs in a defective appliance and the circuit goes dead, from her viewpoint the power outage is a nuisance, notwithstanding that the tripping action may have saved her life.

When the installer is equipping a home with a ground-fault interrupter, the GFI will not stay in the "on" position unless the system is free of inadvertent grounding of the neutral conductor on the load side of the device. Likewise, the GFI will not stay in the closed position if the system has an excessive amount of standing leakage. If the device does not stay "on," and trips repeatedly either because of grounding of the neutral conductor or excessive leakage, it is only natural if the installer considers this tripping a nuisance.

From the engineer's standpoint, if the device has been installed and responds to a transient—due to switching, lightning, or whatever—and the device interprets these signals incorrectly and responds in the same manner as it responds to a ground fault constituting a shock hazard, then truly the device has falsely tripped.

To summarize, tripping may be classified into at least three categories:

1. Intended operation. Tripping in the intended manner when a shock or fire hazard exists is the normal operating mode for the device.

2. Warranted nuisance tripping. In this category, consider tripping of the device in its intended manner, but where the very act of tripping constitutes a nuisance to the customer. Included in this category are operation of the device under grounded neutral conditions and excessive standing leakage in the system. Here the unit is functioning in its intended manner, however, when interpreted from the standpoint of the customer or the installer, the operation of the device does in fact constitute a nuisance.

3. Unwarranted or false tripping. This category comprises tripping due to a response of the device to signals that are not line-to-ground faults constituting a shock hazard.

It is readily apparent that the first two cases cannot be circumvented by any design refinement. In fact, the power outages resulting from either are usually consistent with the design parameters required to produce a device that provides shock- and fire-hazard protection to the user. Any compromise to reduce the so-called nuisance tripping in these cases would compromise the ability of the device to respond in a shock situation.

Truly false tripping (3) is a result of inadequate engineering. Earlier in the development of these devices, this was a considerable problem. However, continuing effort has produced a very reliable and accurate device; see Fig. 6.

Trip value

It is evident that the trip value of the GFI is a compromise between providing maximum protection and avoiding unwanted tripping due to abnormal leakage currents—the sum of wiring and appliance leakage. In January 1969, the American National Standards Institute* issued a new Standard for Leakage Current for Appliances (C101.1), limiting such current (for portable, cord-connected, 120-volt appliances) to 0.5 mA.[2] Although the leakage current of appliances has been gradually reduced over the years, previous regulations set the leakage-current limit at 5.0 mA, and it may take some time for these usable, old, but leaky appliances to be discarded. Compounding the problem is the fact that any mass-produced product must be allowed a manufacturing tolerance, and the transistorized GFIs operate between 3.0 and 4.5 mA from zero to full load, over a temperature range of $-35°$ to 66°C, and with an applied 102–132 volts. Thus, there may not be much margin between a 3.0-mA-trip-value device, and the sum of normal wiring leakage and likely leakage current of appliances—especially some of the older appliances. Fortunately, as far as is known, this effect has not been a source of complaint, and field experience indicates that the 5-mA trip value is satisfactory for protecting individual appliances and 15- or 20-ampere branch circuits.

It is time to dispel two rather widespread misconceptions concerning GFIs. There is an erroneous impression that the GFI reduces the magnitude of the ground-fault current to a safe value. Actually the GFI does not limit current at all; if it did it would seriously limit the load-carrying capability of the protected circuits. The GFI achieves safety by limiting the exposure time and hence the energy in the shock. Danger of electrocution depends upon energy, not current. The shorter the duration of the current flow, the less energy that has coursed through the victim and the greater the current that can be tolerated by the human body.

It is necessary to stress that the trip value of the GFI has no relationship whatever to let-go currents. Although the GFI has to be set to operate at a prescribed current, the device actually monitors the impedance to ground of the protected circuit. Thus a GFI set to trip at 5 mA will cause circuit interruption when the circuit impedance to ground drops to $Z = E/I = 24\,000$ ohms; for a trip value of 10 mA, the critical impedance is 12 000 ohms; and for a trip value of 15 mA, impedance = 8000 ohms.

Why GFI?

In 1930, a series of tests was conducted at Underwriter's Laboratories, Inc.[3], in an effort to determine the maximum current that an individual could withstand for a short time and still have voluntary control of his muscles. In these tests, members of the laboratories' staff were used as subjects. The electrodes consisted of pliers held in each hand. (Although the tests were recorded for alternating current, tests with direct current indicated slightly higher values could be withstood for a short time until a hot spot occurred at the point of contact.) The results obtained for 13 subjects were as follows:

Subject	V, ac	Current, mA	Resistance, Ω
Maximum	40.0	10.0	6670
Average	27.8	7.8	3560
Minimum	20.0	6.0	2330

* Formerly USA Standards Institute.

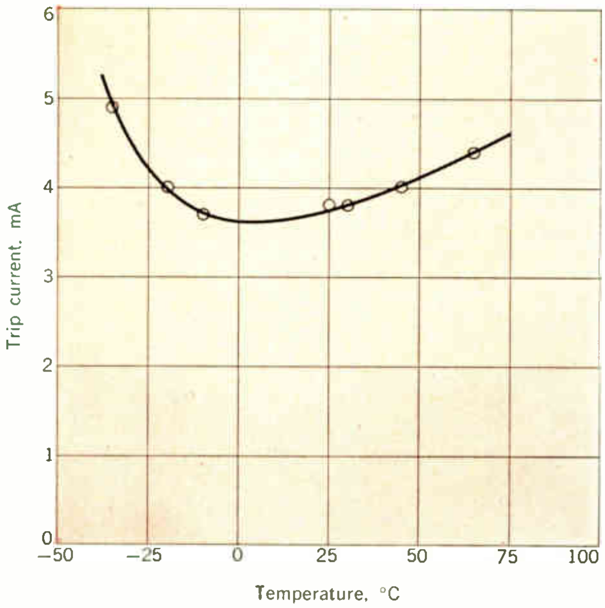


FIGURE 12. Characteristics of a temperature-compensating circuit. Note the satisfactory response over the considerable range in temperature.

conventional overcurrent protective devices, which will not operate at all on the currents that flow through the human body during accidental contact with the 15- and 20-ampere circuits used in the U.S. home or on the farm.

Application of these sensitive GFIs should result in little inconvenience if properly wired, except for possible interruption of illumination. Certainly it is undesirable to be without lights in the dead of night when attempting to locate trouble. Here, the answer is to provide protection only where protection is a must—that is, for the circuits supplying wall and floor receptacles. (It is widely recognized that the electric shock hazard is largely associated with portable appliances in contrast to ceiling or wall-bracket illuminating fixtures.)

A requirement that receptacle outlets be supplied from a subpanel protected by two- or three-wire GFIs should result in a great improvement in the electrical safety of the U.S. home. This should protect the barefooted hedge clipper, the college student studying near a metal desk lamp with his bare feet on the hot-air register, the family sitting or playing in the living room between the metal floor lamp and steam radiator, the daughter accidentally dropping a radio in the bathtub, the father working on the toaster on the stainless-steel sink work area, the mother vacuum-cleaning the house with a 25-year-old machine, the son cutting the lawn with an electric lawn mower, the sister playing the record player at the edge of the swimming pool, the baby chewing the frayed insulation of the sewing-machine cord with wet diapers or bare knees making contact with a grounded metallic object, and the professor sitting on an aluminum chair in the back yard dictating his next article on electric shock into the tape recorder.

Such protective means would also protect the mechanic using the electric hand tool in the wet garage or basement and the fix-it man crawling under the floor with a brass-shell socket extension lamp. It would also mitigate many potentially dangerous situations that occur in and around the all-electric home. Some typical installation and performance characteristics of available commercial GFIs are shown in Figs. 8–12.

Operation on reduced voltage is also of considerable importance. Tests on commercial units indicate that reliable tripping is obtained from rated voltage down to 75 volts. The present emphasis is to study and develop impulse-testing techniques suitable to application of these devices.

Caution

The ground-fault interrupter is a safety device, and not a toy. The public, especially children, must be patiently educated in the proper and safe use of electric appliances and tools. The installation of a GFI is no license for improper, careless, or unsafe practices either in the home or in industry.

The author wishes to acknowledge and thank the several engineers of The Rucker Company, Oakland, Calif., for their splendid cooperation in furnishing technical data and the photographs.

REFERENCES

1. Dalziel, C. F., and Lee, W. R., "Reevaluation of lethal electric currents," *IEEE Trans. Industry and General Applications*, vol. IGA-4, pp. 467–476, Sept/Oct. 1968, and vol. IGA-4, pp. 676–677, Nov./Dec. 1968.

2. "Proposed USA standard for leakage current for appliances," USA Standard C101.1, Jan. 1969.

3. Whitaker, H. B., "Electric shock as it pertains to the electric fence," Bulletin of Research No. 14, Underwriters' Laboratories, Inc., Dec. 1939.

4. Dalziel, C. F., Lagen, J. B., and Thurston, J. L., "Electric shock," *Elec. Eng.*, vol. 60, pp. 1073–1079, Dec. 1941. (Discussion, same vol., pp. 1295–1297.)

5. Dalziel, C. F., Ogden, E., and Abbott, C. E., "Effect of frequency on let-go currents," *Elec. Eng.*, vol. 62, pp. 745–750, Dec. 1943. (Discussion, same vol., pp. 998–1000.)

6. Dalziel, C. F., "Transistorized, residual-current trip device for low-voltage circuit breakers," *AIEE Trans. Power Apparatus and Systems*, vol. 81, pp. 978–983 (Feb. 1963 section).

Charles Francis Dalziel (F) showed an early flair for electricity. As a teenager, he became the 381st licensed amateur radio operator after World War I. During high school and college, he served as a substation operator with a utility company. After receiving the B.S. degree from the College of Mechanics, University of California (there was no Electrical Engineering Department) in 1927, he took a brief excursion through Europe and then joined the General Electric Company as a testman. In 1929, he left G.E. for the San Diego Gas and Electric Company where he was soon given charge of system protection. In 1932, he joined the then new Department of Electrical Engineering, University of California, Berkeley, subsequently received the M.S. degree (1934) and the E.E. degree (1935), and became a professor emeritus in 1967. As this article suggests, his work at the university has centered about electrical safety. Professor Dalziel has volunteered his services to many professional, advisory, and public service groups, including duty as chairman of the San Francisco Section of the AIEE, and vice chairman of the San Francisco Section of the IEEE Power Group. He has lectured throughout the world and has authored 120 papers—some having earned commendations (recently, first prize, IGA Group Transactions, 1969). Professor Dalziel is a registered professional engineer.

A perspective on integrated electronics

Although integrated-circuit technology has enabled dramatic advances to be realized in the electronics field in a very short period of time, it is also placing severe demands on electronic equipment designers

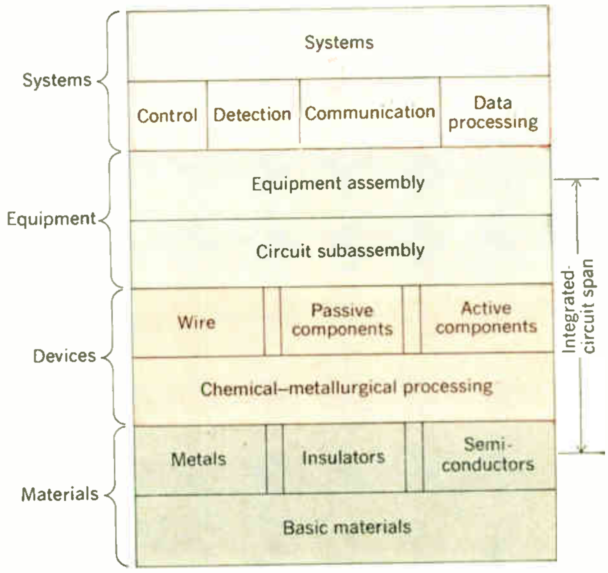
J. J. Suran **General Electric Company**

Integrated-circuit technology may be expected to bring the price of complex electronic equipment within the sphere of industrial and consumer markets and, at the same time, make possible the development and use of significantly more powerful equipment, such as large computers, for military, government, and industrial applications. However, it is important to recognize that the basic ground rules of design are being drastically changed as is the internal structure of the "classical" businesses.

Discovery of the transistor, and its subsequent development in the early 1950s, was an event of major significance in the electronics field. The transistor opened the door to new levels of electronic equipment complexity and miniaturization, and was largely responsible for the development of such major new business areas as digital computers, space applications, and pocket radios. As processing techniques developed, concurrent with quality-control procedures that resulted in high yields and high reliability, transistor fabrication evolved into integrated-circuit manufacturing. The full significance of the integrated circuit is that for the first time it is possible to build complex electronic assemblies in a batch production process that effectively increases equipment reliability while reducing equipment cost.

The evolution of transistors from single components to complete circuits, and the continuing evolution of integrated circuits to complex integrated assemblies (large-scale integration, or LSI), has blurred the lines between component, circuit, and equipment manufacturers. This trend is illustrated in Fig. 1. Traditionally, the electronics-based industry was divided into three basic categories: materials suppliers, component manufacturers, and equipment producers. As the complexity of electronic equipment increased, and as electronics spread into more application areas, a fourth category evolved: systems. Systems design ties together the four

FIGURE 1. Electronics industry structure.



basic uses of electronics—control, detection, communications, and data processing. Hence it may be thought of as the "software" content of electronic equipment applications. As a result of integrated-circuit technology and systems technology, it is apparent that modern electronics is in the process of creating a new industrial configuration—namely, materials suppliers, equipment fabricators, and system designers. Therefore, although electronics continues to create new business opportunities, the advancing technology is also changing the internal structure of the "classical" businesses.

In addition to stressing traditional organizational relationships within industry, integrated-circuit technology has placed severe demands on electronic equipment designers. A transition from discrete-semiconductor to integrated-circuit design is considerably more demanding than was the transition from vacuum tubes to transistors. In the latter case, certain principles of "duality" could be invoked to ease the transition. In general, the duality principles preserved the old ground rules. However, in the move from the discrete world to the integrated world, the basic ground rules of design have changed. Discrete component circuit designs not only are nonoptimum but, in general, may be technically unworkable in integrated form. Thus, large bodies of design knowledge, particularly of the "handbook" type, have been made obsolete.

Effects of ground-rule changes

Cost and component availability. Figure 2 summarizes the ground-rule changes in terms of component cost for the worlds of discrete technology, silicon-integrated-circuit (SIC) technology, and hybrid-integrated-circuit (HIC) technology. Of paramount significance is the complete reversal of relative cost of the active element (transistors) between discrete and SIC assemblies. Also significant is the high cost or lack of availability of inductors in integrated circuits. The low cost of active elements has led to an "explosion" in the numbers of transistors used in circuit design. For example, Fig. 3 compares a SIC television chroma demodulator, marketed by Fairchild, with a discrete vacuum-tube design still widely used in color television receivers. Not only has the SIC-to-discrete active element count in the demodulator increased by a ratio of 21 to 1, but the SIC circuit employs no energy-storage elements. Apparently, energy management in the integrated world follows completely different principles. Furthermore, circuit complexity can no longer be measured by active-element count.

Component tolerances. In the days of discrete components, circuit designers were generally free to select components within relatively narrow specifications. Thus, resistors with tolerances of ±10 percent were considered "sloppy" (albeit inexpensive) and most equipment had at least a few components that were specified to within 1 percent. Tolerance rules have changed drastically, as illustrated in Table I. Circuit designers in SIC technology have become resigned to tolerances of ±25 percent, although HIC techniques still allow an out for precision-component requirements. Nevertheless, with SIC technology dominant, today's designers must accept broad tolerances.

One effect of the new tolerance rules has been a substantial increase in the use of feedback. Since gain is now

I. Effects of design-rule changes on tolerances

	Tolerance, percent		
	Discrete	SIC	HIC
Manufacturing	1.0	25	10
Adjustable	0.1	—	0.1

FIGURE 2. Design-rule changes relative to cost.

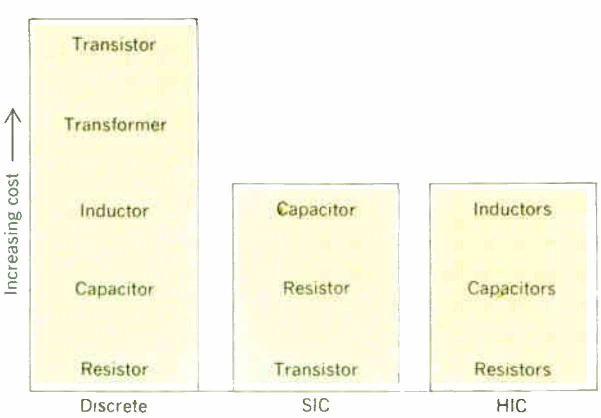
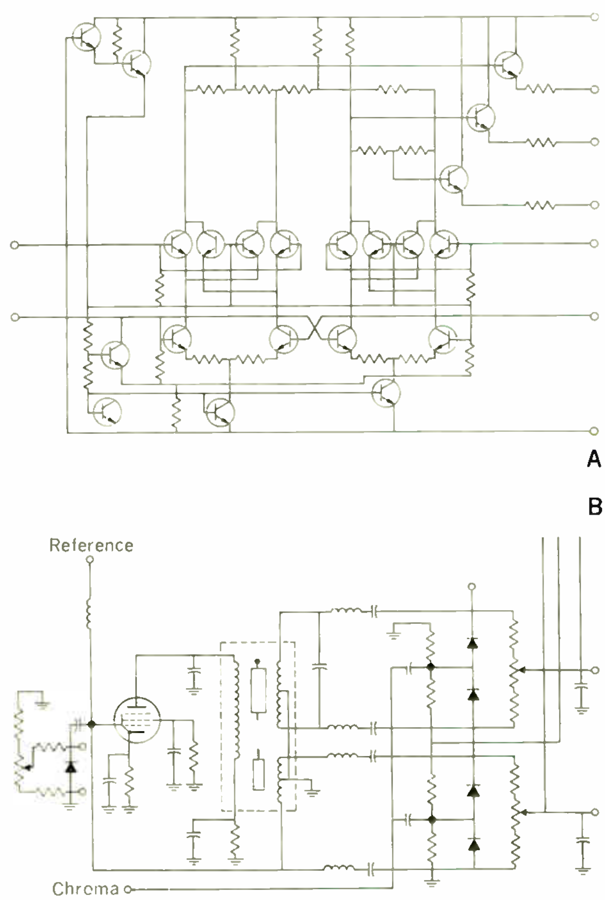


FIGURE 3. Comparison of (A) integrated and (B) discrete chroma demodulator circuits.



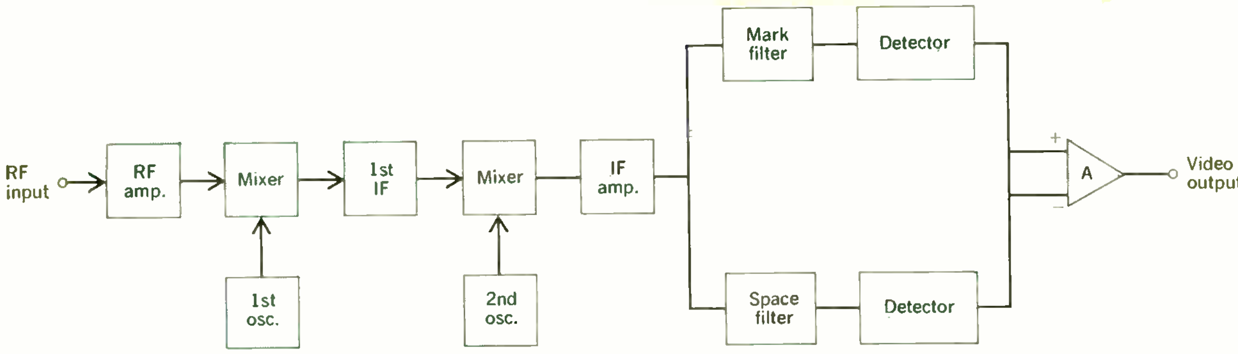


FIGURE 4. FSK receiver designed for discrete-circuit technology.

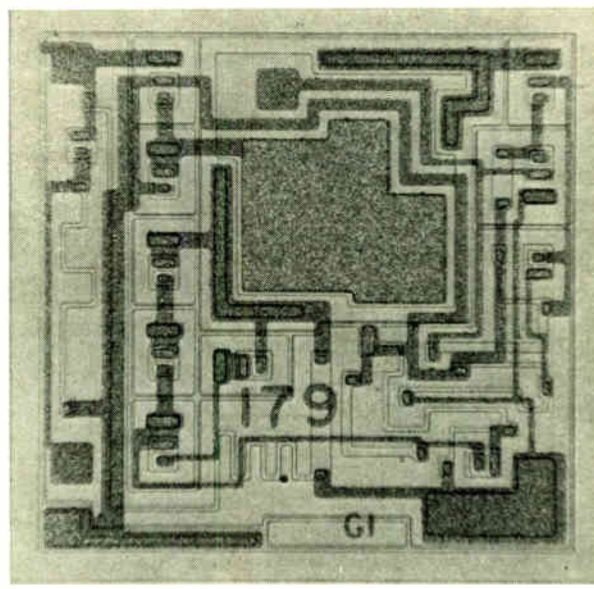


FIGURE 5. FM discriminator chip for silicon integrated circuit.

inexpensive, circuit stabilization by degenerative feedback methods is economical. And since semiconductor junctions tend to be well-matched if they are fabricated simultaneously in proximity to each other, self-tracking circuits employing differential-pair feedback may be used to compensate for temperature changes. Self-adjusting feedback methods have also greatly increased in SIC designs whereby circuits "lock on" to signals—in phase-locked loops, for example. An important consequence of the increased economy of feedback, at the equipment level, is the ability to make electronics relatively free of human adjustments. This can be expected to reduce maintenance costs as well as manufacturing "tune-up" costs.

Another effect of tolerance considerations has been the strong trend to digital circuits. At first glance it might appear that a "switch" is relatively simple and tolerant compared wtth a linear amplifier. Hence, the digital trend makes "intuitive sense." However, to experienced digital-circuit designers who have wrestled with tradeoffs between such interrelated variables as fan-in, fanout, speed, thresholding, directionality, isolation, reshaping, tolerances, stability, power dissipation, logic capability, etc., the trend must be based on sterner stuff than switch simplicity—and it is. An important implication of digital circuits is the broadband and nonlinear nature of the signals. Since energy can usually be managed by low-pass filtering, the digital approach relieves the selectivity problem in integrated circuits, thereby avoiding the high cost of capacitors and inductors. And, since the signal amplitudes are usually constrained to binary values, clocking techniques can be used for significantly limiting noise and distortion effects. Furthermore, electrical storage can often be replaced by logical storage (for example, in the use of counters as integrators), and this too unburdens integrated technology from the fabrication of reactive elements. Finally, digital circuits tend to be highly iterative. Computers, for example, can be built by iterating a single gate design. Consequently, although *apparently* much more complicated than linear systems, digitally designed systems may provide important economies in the synthesis and layout stages of SIC fabrication. Thus, the old measures for cost comparison no longer apply.

Design methodology. Designing an integrated circuit requires considerable knowledge of the fabrication processes. This knowledge may be in the form of design rules or may be based on first-hand familiarity with the processes involved; in any event, however, an integrated-circuit design must start with process modeling. At the system end, because of the major changes in ground rules, the design must be based on broader considerations than individual circuit blocks. A circuit-block by circuit-block transition to integrated design may appear comfortable and may be relatively easy to digest; however, this approach, if it succeeds at all, will succeed only accidentally. Systems and equipment designed for discrete-technology implementation are not partitioned optimally or practically for integrated technology. Hence, the design must start at the equipment or system level, with the blocks partitioned in accordance with the new ground rules, and once the parts are functionally compatible with each other as well as with the system requirements, a block-by-block design becomes feasible.

Consider, for example, the frequency-shift keyed (FSK) receiver shown in partial-block-diagram form in Fig. 4. The receiver converts a binary frequency code to a binary amplitude representation of the code. Starting at

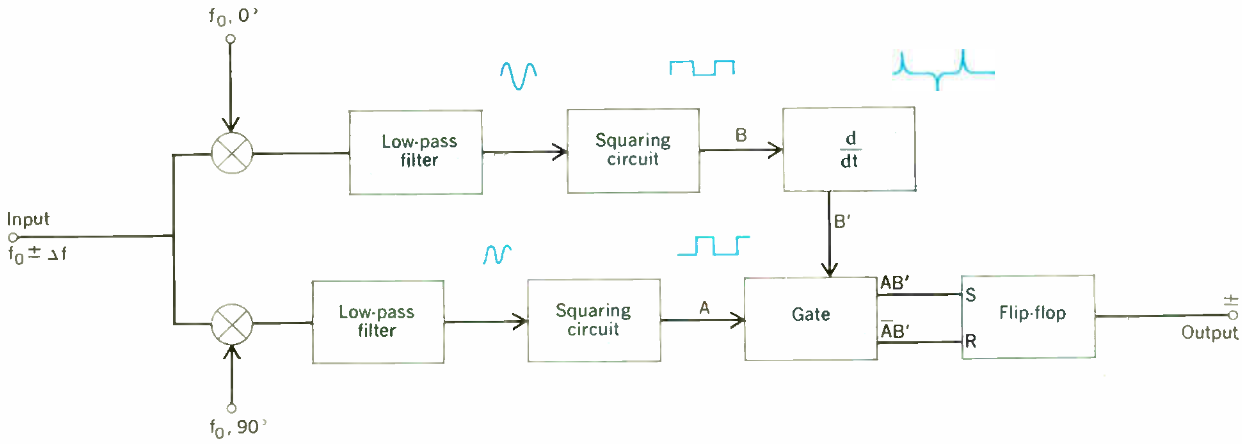
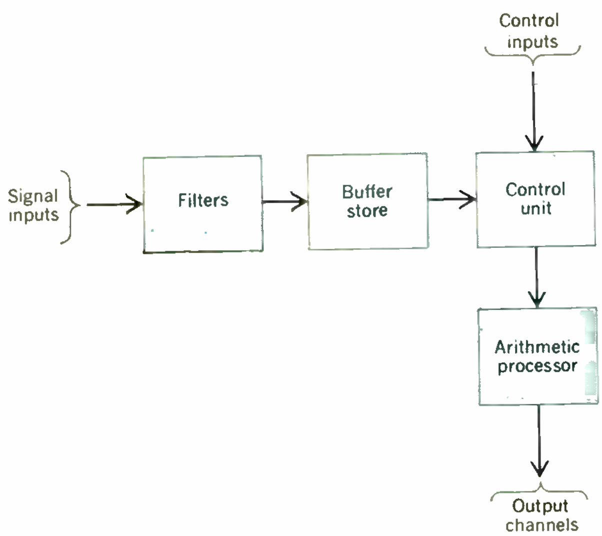


FIGURE 6. FSK receiver designed for IC technology.

FIGURE 7. Detection-system data processor.



the back end, the discriminator can be implemented in the form of a pulse-counting circuit so that no tuning or expensive reactive elements are required. Such discriminators have been built in SIC form and one, having dimensions of 50 × 50 × 10 mils (about 1.3 × 1.3 × 0.25 mm), is illustrated in Fig. 5.[1] The SIC version does the job functionally but, because of its broader band compared with conventional discriminators, it has a lower conversion gain and it also introduces pulse noise into the system. Thus, other parts of the system must be redesigned to compensate for the new characteristics of this block; for example, additional gain must be added in the video section and the filtering networks must be reevaluated and probably redesigned to block the added energy components. By the time the interfaces are adjusted to compensate for the new discriminator design the advantages of integration could be lost.

A more practical approach is to redesign the FSK receiver from scratch in accordance with the ground rules of integrated-circuit processing. Such a redesign is illustrated in Fig. 6.[1,2] Instead of using a conventional superheterodyne front end, the design employs an arrangement whereby the signal is converted in quadrature to baseband and then detected by logic circuits operating on the phase difference between the quadrature channels. With the exception of the oscillator, all blocks are broadbanded and can be realized with a minimum number of reactive elements. Although the quadrature detection configuration uses many more active elements than the conventional superheterodyne receiver, it can be fully integrated and realized with one SIC chip and an HIC low-pass filter assembly.

Integrated-circuit design implies integration not only with fabrication processes but also with system requirements. The intersection of integrated-circuit technology with systems technology in the chart of Fig. 1 is real. This intersection point becomes dominant in the world of large-scale integration.

Large-scale integration

Large-scale integration is defined as the interconnection of 100 or more circuits of logic-gate complexity by silicon processing. Since several transistors are used in the design of a single gate, LSI implies the interconnection of at least several hundred transistors in a volume of silicon approximating 0.0001 in^3 (about 1.7 mm^3). One typical LSI chip, based on silicon MOSFET technology, houses about 750 transistors interconnected as an equipment subassembly. This chip barely meets the LSI definition. Assemblies containing 10 000 transistors per chip have been fabricated.[3]

Large-scale integration portends a new age for electronics, the outlines of which are not yet clearly discernible; LSI will change the ground rules of system design just as surely as SSI (small-scale integration) changed the ground rules of circuit design. Electronics is now entering the world of large numbers, where millions of interconnected active devices become economically and technically realistic. Problems of design, test, and evaluation of such large systems are formidable, but at most these problems will delay, not prevent, the onset of LSI.[4] Designers have already begun to think in terms of large numbers, particularly in the data-processing field, where problem-solving capability is size-related.

An example. An example of a system that has been planned for LSI realization is illustrated in block-diagram form in Fig. 7.[5] The configuration represents a special-purpose data processor for use in detection applications. A summary of the numbers of components used is given in Table II. One of the first entries to catch the eye is the number of transistors used: 1 300 000. However, the first identifiable part is at the chip level. Each of the 4000 chips has approximately 100 interconnected logic gates built into it. The total volume of silicon used for the 1 300 000 transistors is less than 8 cm^3.

The partitioning plan assumes that about eight LSI chips will be interconnected by hybrid integrated processes (thick or thin films) into functional assemblies referred to as "modules," of which there are 500 in the system. About ten modules will then be interconnected by printed-board wiring (multilayer) into "backboards," of which there are a total of 50 in the system. Thus, each module, measuring about 5 × 2.5 × 0.6 cm, contains some 2800 transistors and each printed board, measuring about 23 × 15 × 0.6 cm, contains some 28 000 transistors.

Reliability and availability requirements dictate that self-diagnostic techniques be built in as part of the system. Diagnostic and repair considerations to a large extent determine how the system is partitioned, and these considerations are carried right down to the chip-design level. For the system under discussion, repair is effected only on backboards, not within modules; consequently, the throw-away level is in units of 2800 transistors. The entire system utilizes only eight types of modules and five types of backboards. Minimizing the number of types of assemblies reduces inventory and production costs.

Table III summarizes the interconnection count at each component level. Approximately 600 connections must be made on each chip by semiconductor processes. But it is interesting to note that the interconnection complexity at the module and backboard level is about the same as at the chip level, though at lower densities because of the increasing area of subsequent interconnect carriers. In order to insure that needlessly severe demands are not made at any one level of interconnections, it is desirable to design the interconnect system as a single entity, making tradeoffs at the module level to simplify the chip design or for other reasons.

As might be expected when packing densities approach the levels of LSI, thermal dissipation becomes a major problem. Table IV summarizes the power and power density dissipated at each component level for the data-processor example. Although the average power dissipation of each transistor is very small, the power density is quite large. For example, the power density at the chip level is approximately equivalent to that of an average-size 500-watt lamp. Thus, the interface between chip and module must be carefully designed to insure good thermal conduction out of the chip to avoid excessive chip temperatures. It might appear from Table III that once heat has been successfully removed from the chip the thermal problem has been solved; but here again, unless thermal design is carried through the entire system, serious consequences could ensue. The maximum temperature rise in the printed board can become very severe despite the relatively modest temperature difference between module and chip. Hence it is not sufficient to concentrate thermal design on only those parts of the system where power densities are high.

Some implications of LSI. Despite the novelty of the new electronic world of large numbers, several provisional generalizations might be drawn from specific examples of the kind just discussed. One important implication is that partitioning, including chip design, will be determined by *total* system considerations. Tradeoffs must be made between process capability, device performance, circuit design, logic design, interconnect techniques, thermal design, maintenance requirements, environmental specifications, and other factors, in order to optimize system figures of merit, such as cost or performance. Optimization at the chip or module level is not necessarily synonymous with optimization at the equipment or system level.

II. Component count for special-purpose data processor

Component Level	Numbers	Types
Transistors	1 300 000	6
Gates	400 000	6
Chips	4 000	8
Modules	500	8
Backboards	50	5

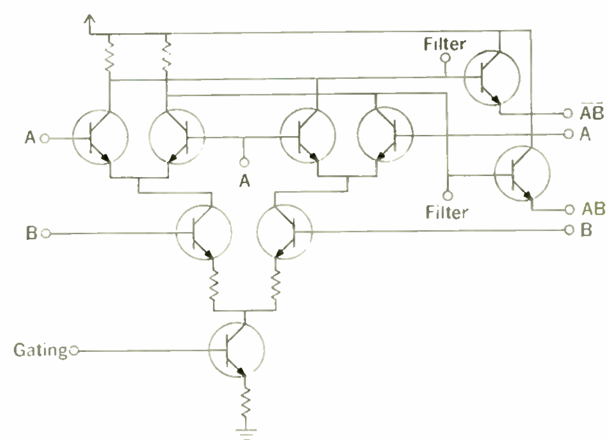
III. Interconnection count for special-purpose data processor

Component Level	Connections	Pins	Gates per Pin
Transistors	3	—	—
Gates	6	—	—
Chips	600	30	3.3
Modules	500	40	20
Backboards	500	100	80

IV. Power dissipation for special-purpose data processor

Component Level	Power Dissipation, watts	Power Density, W/cm² (W/in²)
Transistor	0.0005	15.5 (100)
Gate	0.0015	7.8 (50)
Chip	0.170	3.1 (20)
Module	1.36	0.15 (1)
Backboard	13.6	0.015 (0.1)

FIGURE 8. General-purpose multiplier.



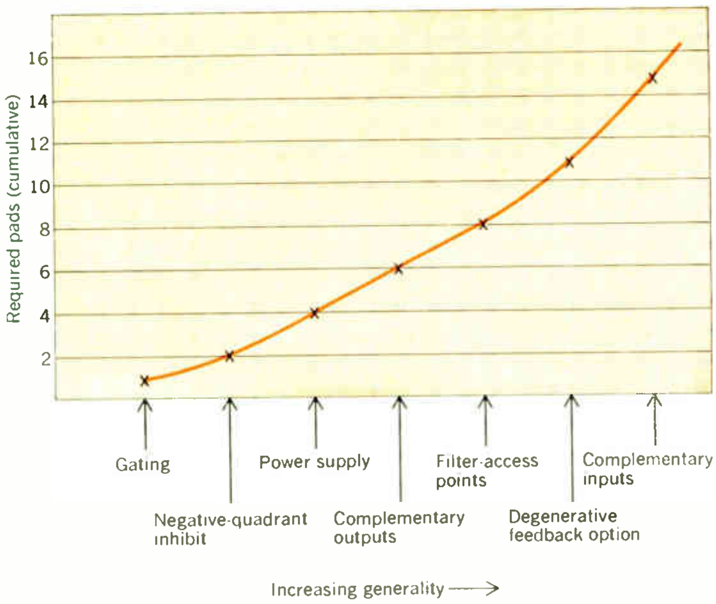
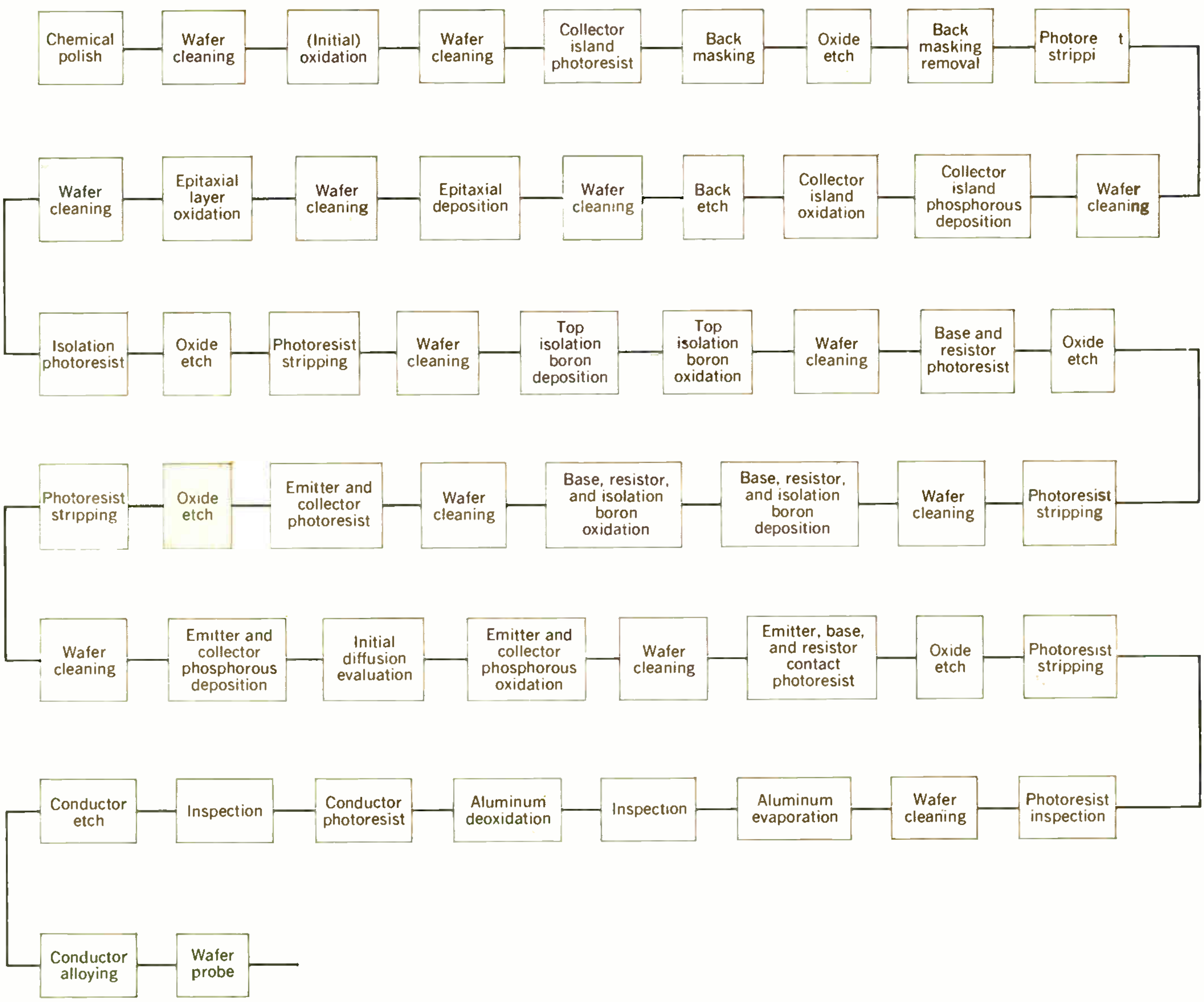


FIGURE 9. Pad count vs. "universal" multiplier circuit.

A second implication is that standardization may move from the circuit level to the device-structure and processing levels. Chips containing hundreds of interconnected transistors are highly committed configurations and will probably be custom-designed to achieve system- or equipment-level optimization. For example, the special-purpose data processor outlined in Fig. 7, and discussed previously, requires 4000 LSI chips of eight different types, each custom-designed; if the same system were implemented using standard circuits of the complexity of a *J-K* flip-flop, approximately 40 000 chips would be required.[5] It is not difficult to see how high a price would be paid at the system level in the sheer handling of 36 000 additional parts for the "convenience" of standardized circuits.

At times, circuit standardization imposes severe penalties at the equipment level in considerably less complex equipments than data processors. Shown in Fig. 8 is a multiplier configuration designed for appropriate generality to serve as a broad-based standard SIC in many kinds of applications. Figure 9 illustrates the num-

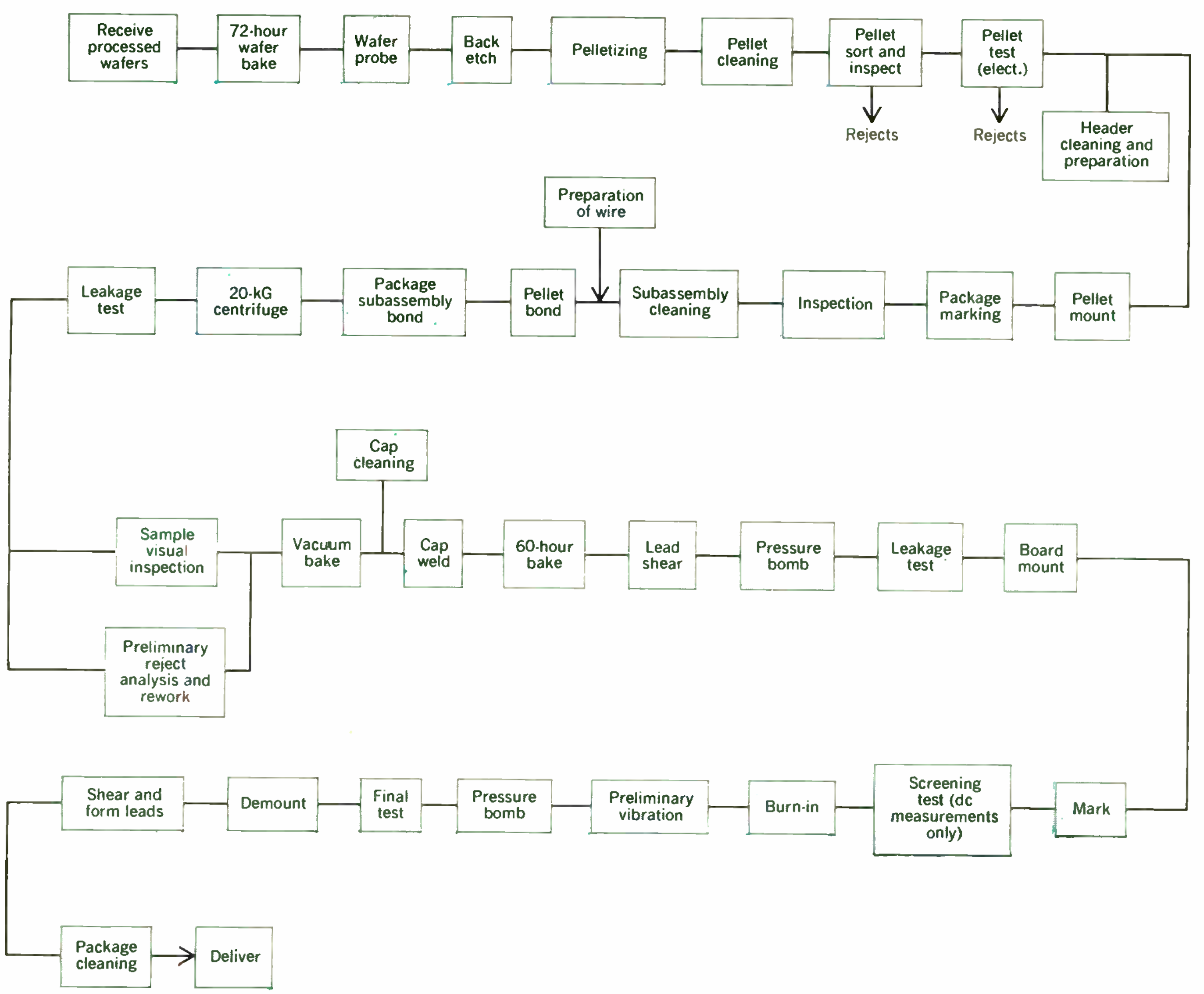
FIGURE 10. Wafer-processing flow diagram.



ber of pads required on the chip in order to accommodate the many options under which the circuit may be employed. Four such multipliers might be used to build a TV chroma demodulator of the type illustrated in Fig. 3, but 48 additional pads, out of a total of 56 on the chip, would be nonfunctional.[6] This would result in unnecessary added cost in semiconductor processing due to the excess silicon used to provide for the unneeded pads. Even worse, from a cost point of view, the excess pins would have to be accommodated in interconnections at the equipment level. It is considerably more economical to design the chroma demodulator as a custom chip. Thus, where high-volume production of relatively simple equipment such as television receivers is expected, custom design of chips is the sensible route because of the liquidation of design costs over a high production volume; but interestingly enough, where low-volume production of complex equipment is expected, such as in special-purpose data processors, the same route of custom-designed chips is indicated because of the economies achieved in using many less parts or less costly diagnostic procedures. Economies expected from circuit standardization may thus apply to only exceptional cases in the era of integrated circuits.

The use of customized integrated circuits where production volume is low raises doubts concerning quality control. How is reliability to be assured if large-scale testing of circuits becomes too expensive for the volumes of circuits to be sold? One answer to the question is to move reliability assurance from the product to the process. If processes become standardized to the extent that process changes are under tight control and if, further, the same processes are used for both high-volume and low-volume production, it would seem that reliability assurance at the process level becomes feasible. Indeed, LSI chips will never be fully tested, at functional levels, prior to use because of the complexity of the configuration. (For example, a 100-stage flip-flop assembly could have as many as approximately 10^{30} logic states. If it were desired to cycle through all possible logic combinations in order to test the assembly fully, and if each cycle check could be achieved in 10^{-9} second,

FIGURE 11. Chip-packaging flow diagram.



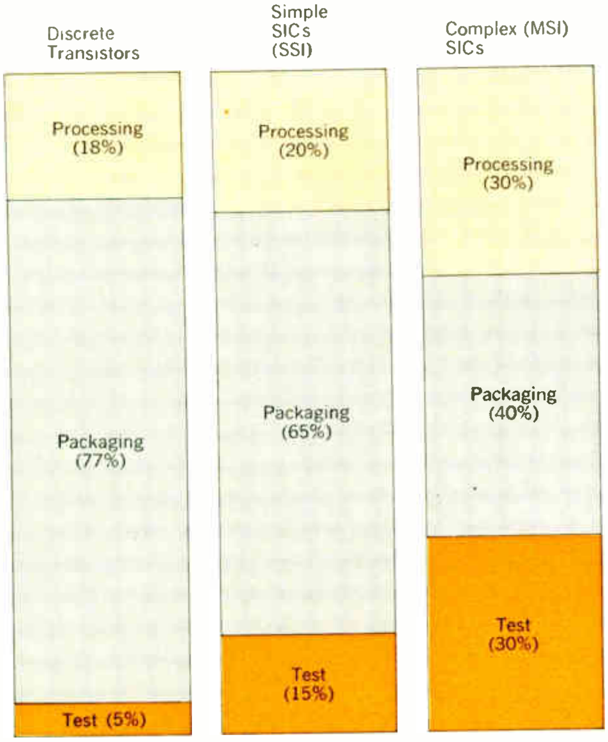


FIGURE 12. Semiconductor manufacturing cost breakdowns.

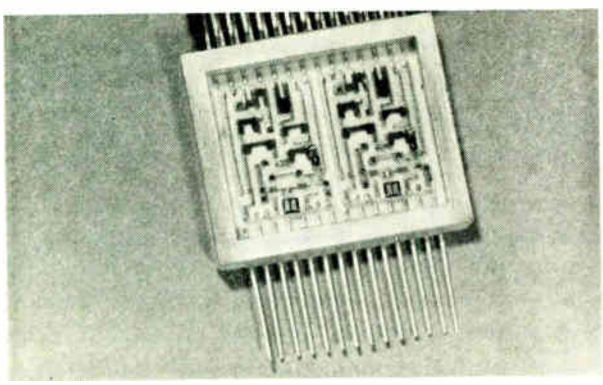
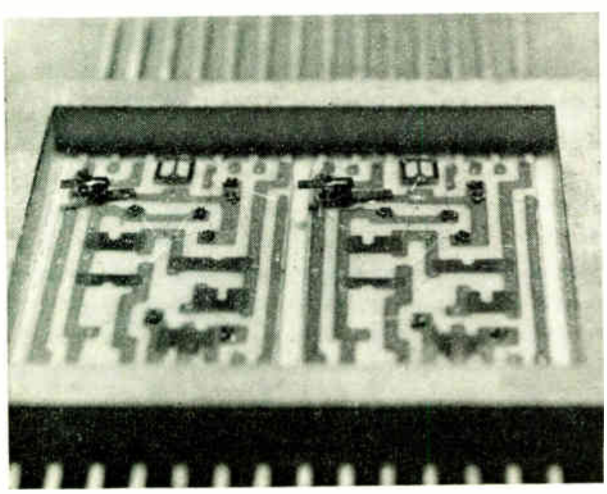


FIGURE 13. Hybrid integrated circuit with several silicon chips on a thick-film substrate.

FIGURE 14. Close-up of assembly shown in Fig. 13.



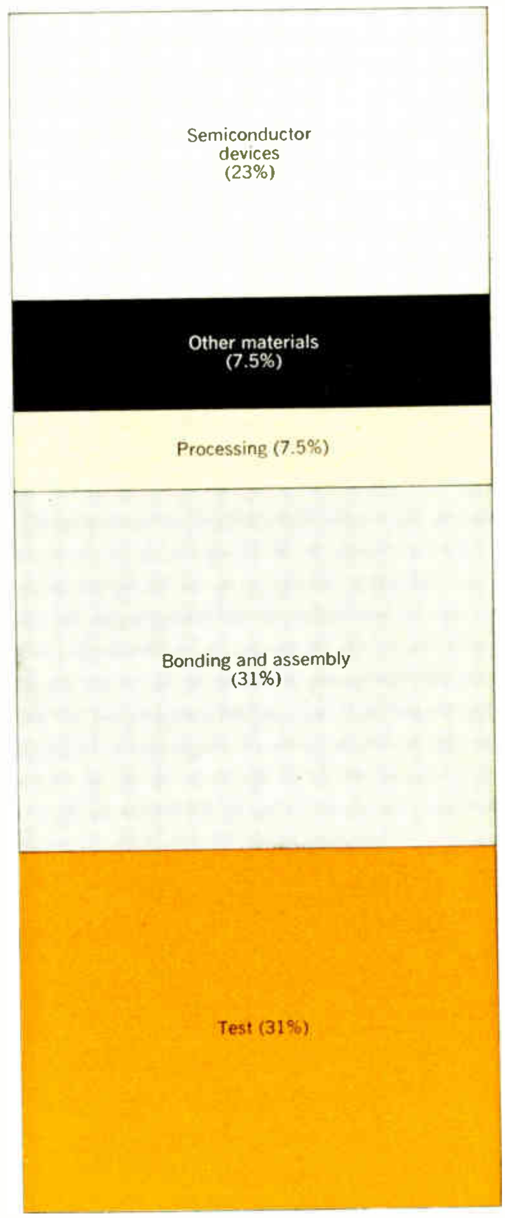
it would take more than 10^{10} years to make a complete test.)

Confidence of the reliability of such arrays, in the absence of decades of experience, is based upon the fact that the same semiconductor processes used to produce millions of reliable transistors per year are used to fabricate LSI chips. Process control, coupled with gross input–output tests at the functional level, may be the key to reliability assurance.

The shape of integrated electronics

Silicon fabrication processes provide the means for forming active and passive components as well as for interconnecting them. The processes are essentially chemical and photolithographic in nature and involve over 100 tightly controlled steps.[7] Figure 10 is a block diagram of the basic semiconductor process up to the stage of finished wafers. Silicon integrated circuits would

FIGURE 15. Hybrid IC manufacturing cost diagram.



be prohibitively expensive if each one were fabricated separately by such a complex process. The fact that circuits and circuit connections are batch-processed—that is, many are produced simultaneously—makes the unit cost economic.

Once wafers have been processed, the circuits must be individually tested and marked to identify those that meet specifications. The wafer is then scribed or chemically etched to separate the circuits into chips, and finally the chips may be packaged into suitable containers for shipment or use. The process used in converting a wafer to packaged chips is also complex. A block diagram of this process is illustrated in Fig. 11. Since many of the steps associated with packaging are sequential, as opposed to the batch technique, it may be expected that the cost of packaging individual chips is high. This is confirmed by the estimated cost breakdown shown in Fig. 12. (Hermetic packaging is assumed.) It is apparent from this chart that even for complex SIC manufacturing, the operations that are associated with assembling, connecting, and packaging chips constitute approximately half the cost of the final product.

It is not surprising, therefore, that the industry trend is toward packaging several chips in a single container in order to try to distribute the high costs of package "real estate" over several chips. "Super packs," as illustrated in Fig. 13, are becoming increasingly common. Chips are bonded to a suitable substrate, such as alumina, and the chips are interconnected by runs formed on the substrate. Thin- or thick-film processes are employed to fabricate this second level of interconnections (for example, see Ref. 8).

The interface between conductor pads on the chip and conductor runs on the substrate may be bridged by flying leads, as revealed in the super-pack close-up of Fig. 14. Besides being esthetically incompatible with the batch-connection processes used on the chips and on the substrate, flying leads are a source of trouble. Their fragility sometimes leads to failure. And since they are placed and bonded one at a time, the operation of lead attachment is time-consuming and costly. A relative cost breakdown for the manufacture of super-pack modules containing five to ten SIC chips, using either thick- or thin-film interconnect patterns on an alumina substrate, is shown in Fig. 15. Chip bonding and flying-lead attachment accounts for about one third of the cost of the finished product.

Several alternative methods of crossing the SIC–HIC interface have evolved over the past few years, each having the common feature of extending the concept of batch processing to batch interconnecting. Beam-lead technology extends semiconductor processes to include overhanging leads on the chip; the leads are rugged enough to withstand thermocompression bonding to conductor patterns on dielectric substrates.[9] All of the beam leads on a chip are bonded simultaneously, thus achieving a form of batch interconnection. Flip-chip technology provides "bumps," made of a solder material, either on the chips or the substrates; when the bumps

FIGURE 16. HIC interconnection system.

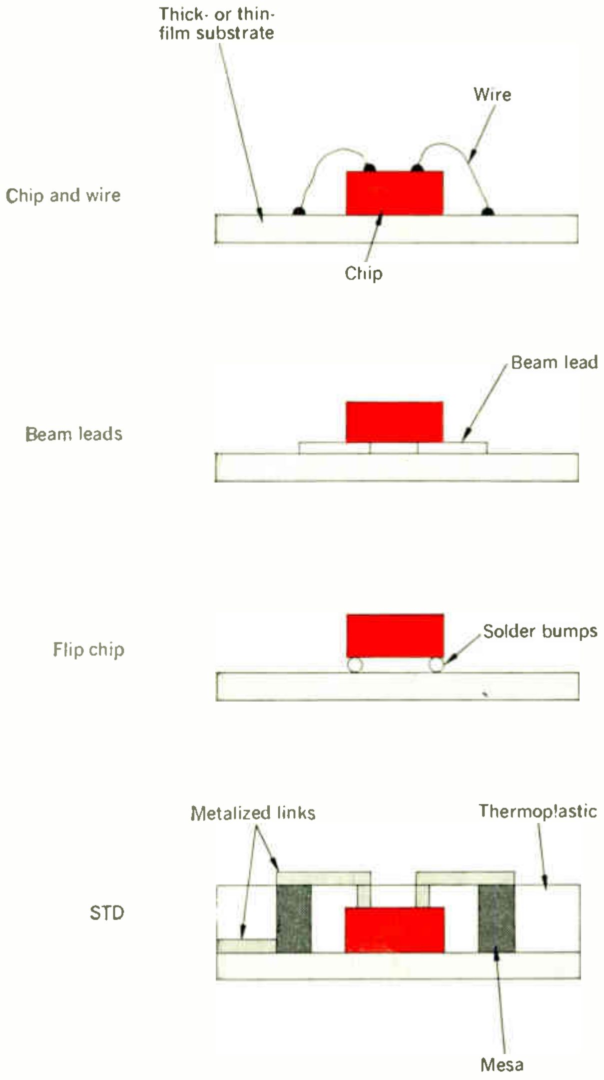
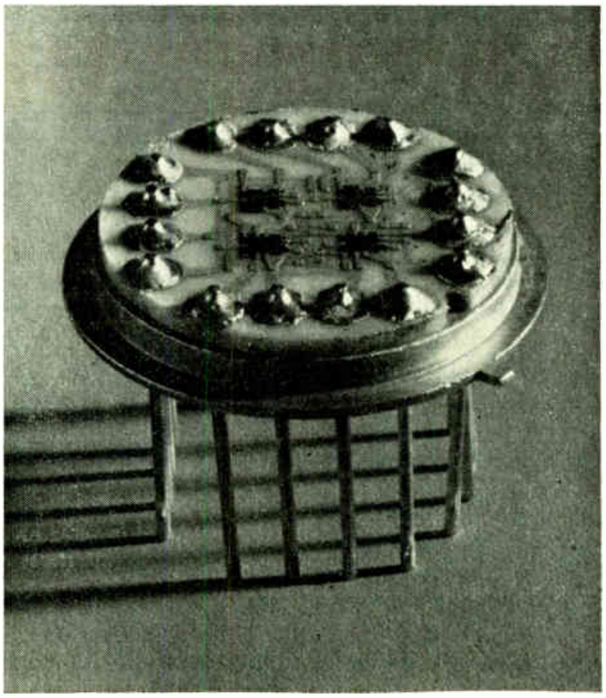
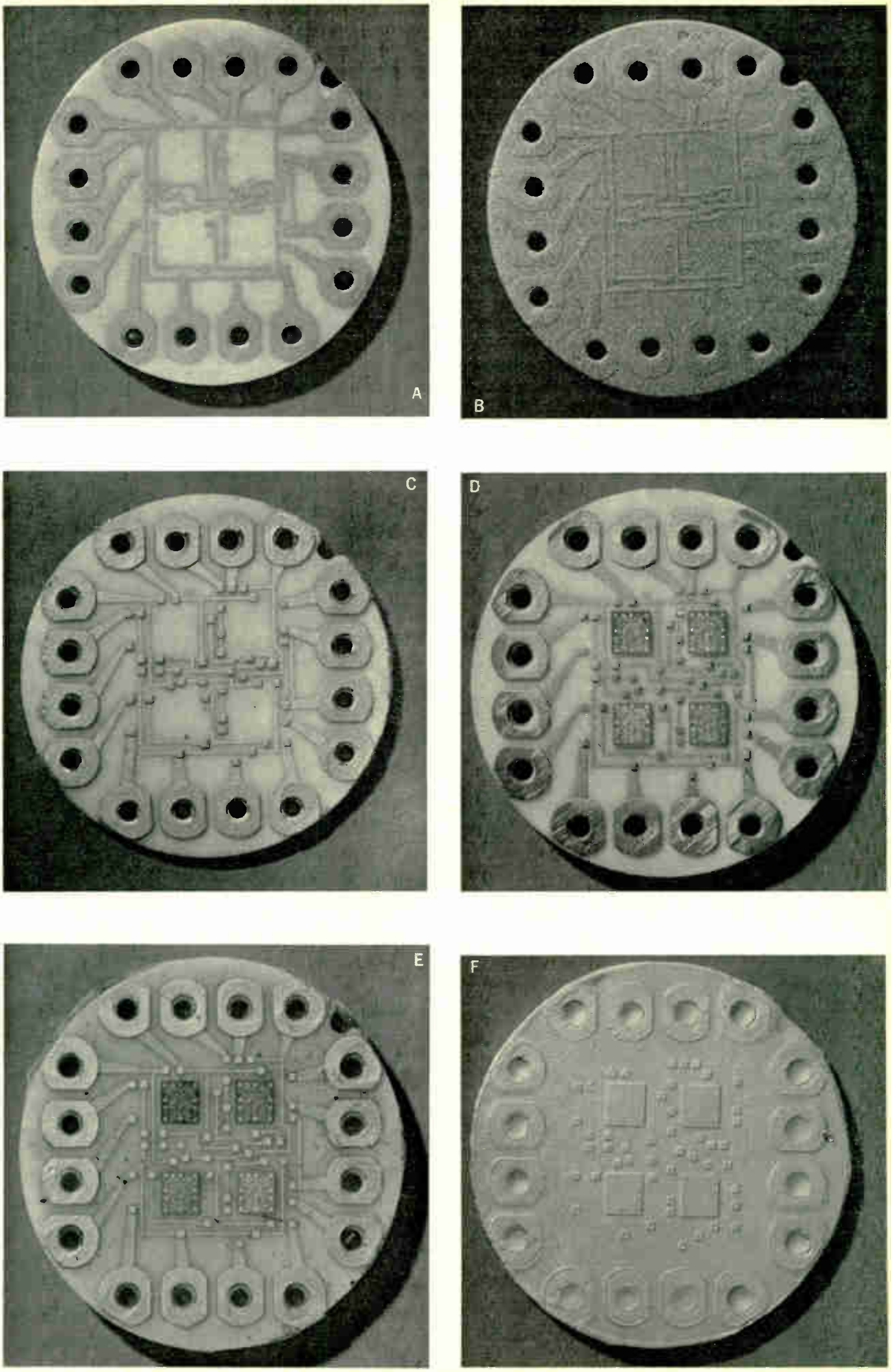


FIGURE 17. STD four-chip circuit.




A
B
C
D
E
F

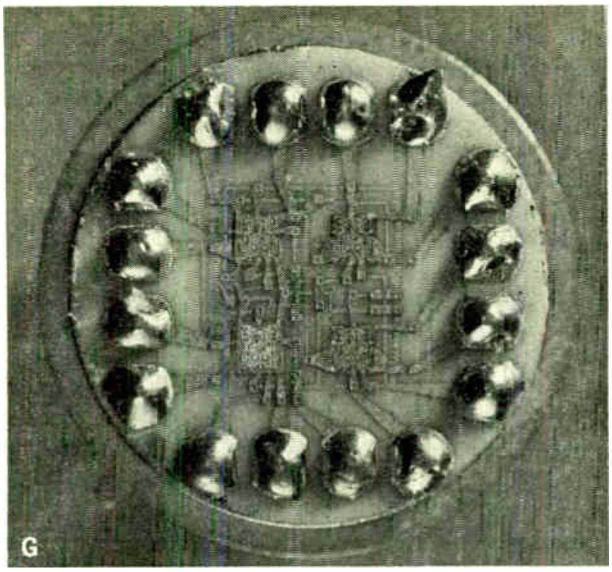


FIGURE 18. Major process steps in four-chip circuit fabrication. A—First-layer conductors on substrate. B—Mesa metalization over first-layer conductors. C—Etched mesas on first-layer conductors. D—Chips bonded to substrate. E—Thermoplastic (FEP) pressed on substrate and holes etched through FEP to chip pads and mesas. F—Second-layer conductor metalization on thermoplastic. G—Etched second-layer conductors.

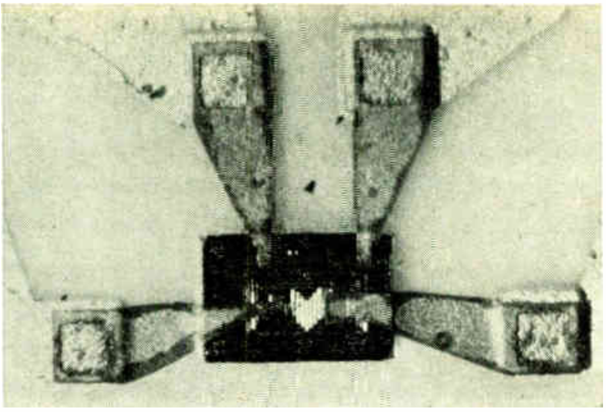


FIGURE 19. STD interconnect to a MOSFET tetrode showing mesas and connections to internal chip pads.

are melted and then resolidified, rigid connections are made between the SIC chip pad and the conductor patterns on the substrate.[10] The bumps on a chip can also be simultaneously bonded, and thus batch interconnections are achieved. STD (semiconductor–thermoplastic–dielectric) technology bonds the chips onto the substrate in a face-up fashion, covers the assembly with an inert thermoplastic material, etches through the thermoplastic to provide access to the chip pads and conductor pattern on the substrate, and then metalizes a second conducting layer over the thermoplastic while simultaneously interconnecting the chips.[11] Flying-lead, beam-lead, flip-chip, and STD techniques are compared schematically in Fig. 16.

The second level of interconnections. Since interconnecting LSI chips in a subassembly involves about as many interconnections as may be on each chip (see Table III), it is not surprising that this second level of interconnections is attracting a great deal of research and development interest. SIC chips are relatively delicate and so the processes used to batch-interconnect them must be tightly controlled and carefully engineered. It is of interest to note that some advanced batch-interconnect methods are beginning to resemble the chemical and photolithographic techniques commonly used in SIC fabrication.

Beam-lead, flip-chip, STD, and other techniques for achieving batch interconnection of semiconductor chips have as their common objective the crossing of the semiconductor–film interface in an economic and reliable manner. All of these techniques have their proponents. One that General Electric has been developing in recent months is STD; in view of the fact that little information has been published on this process it may be instructive to consider STD in more detail.

A circuit of four chips, interconnected by the STD process and mounted in a TO-8 header, is shown in Fig. 17. An example of the major process steps used in fabricating the four-chip circuit is illustrated in Figs. 18(A) through (G). The first layer of interconnects is provided by standard thin- or thick-film processes; see Fig. 18(A). A thick layer of copper and gold is then plated over the substrate, as shown in Fig. 18(B). By the use of photolithographic methods, a "mesa" pattern, Fig. 18(C), is then etched out of the copper–gold layer. The mesa pattern may be thought of as the inverse of plated-through holes in a multilayer printed board. SIC chips are then registered to the substrate; see Fig. 18(D). After chip bonding, a cover sheet of thermoplastic (in this case, FEP-fluorinated ethylene propylene) is bonded to the substrate assembly by a combination of elevated temperature (about 300°C) and moderate pressure. After cooling, the thermoplastic serves as a rigid cover plate over chips and substrate. A second photolithographic process is then employed to delineate a hole pattern on the cover sheet so that holes subsequently can be etched out of the thermoplastic over the chip-pad areas and the mesas; see Fig. 18 (E). After the holes are etched, a metal layer [chrome–copper in circuit shown in Fig. 18(F)] is deposited over the thermoplastic in such a way that all the etched holes are filled. A final photolithographic step forms the conductor pattern on the second layer, simultaneously interconnecting all chips and providing crossovers between the first and second layers. Some 100 connections, batch-processed, have been made in the circuit shown in Fig. 18(G).

Because of its tie to photolithography, STD is a flexible process. A chip of any complexity can be accommodated, thus easing second-level interconnect processing for custom-designed chips. Gold- or aluminum-chip metalizations can be used in conjunction with thin- or thick-film substrates, and therefore different SIC or HIC processes can be accommodated. Since the thermoplastic material completely covers the chip, conductor runs can be made over the chip; hence, flexibility is provided to the SIC designer, who is no longer constrained to bring all pads out to the periphery of the chip. Figure 19 illustrates an STD interconnect to a MOSFET tetrode; here, for conve-

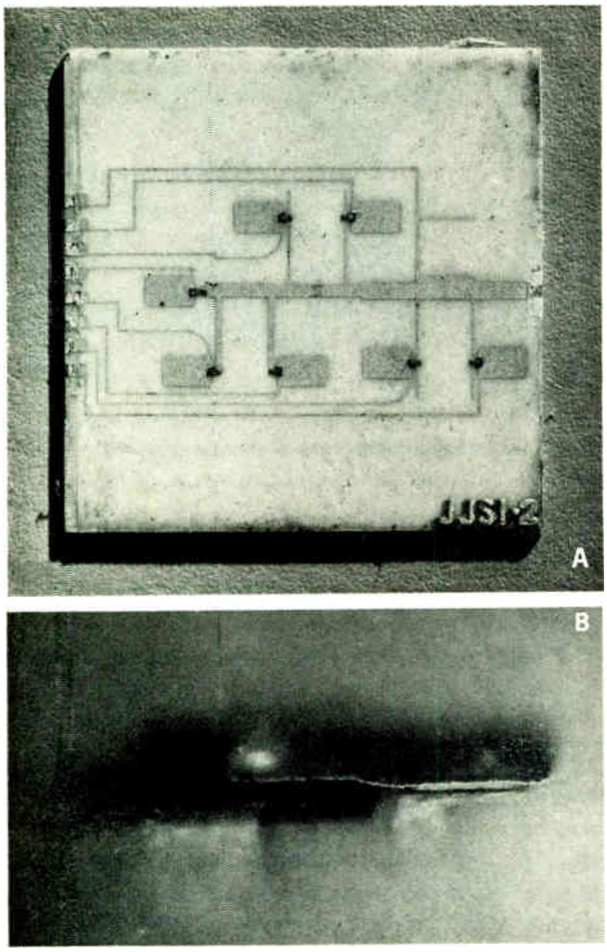


FIGURE 20. A—Microwave X-band phase-shifter circuit. B—Close-up view of an STD connection to one of the p-i-n diodes shown in (A).

nience of design and for performance advantages, the gate and drain pads are completely surrounded by the meandering gate −2 pattern. The inert thermoplastic material provides a means for reaching the internal pads without requiring an additional metalization layer on the semiconductor and also provides protection for the delicate MOSFET surface.

Semiconductor chips that can be eutectically bonded to the substrate have much lower thermal resistances at the chip substrate interfaces than chips that depend on heat conduction through lead connections. There is an order-of-magnitude advantage in thermal conductivity of eutectic bonding, which becomes increasingly important as LSI chip densities increase and as higher power chips are deployed in integrated electronic structures.

Microwave integrated circuits, unlike LSI, generally consist of few active elements. Because of the distributed nature of the circuit, packing density is low. However, the interconnections between chips in a microwave configuration are critical to the circuit's electrical performance. Microwave microelectronic interconnections must be designed to carefully specified transmission-line characteristics. Uncontrolled parasitic inductances of the order of one nanohenry can seriously degrade the performance of X-band circuits. The photolithographic processes used in the STD interconnection system can delineate transmission-line patterns of high resolution and quality. These lines are fed directly into the metalization pattern of the semiconductor devices. Thus, parasitic reactances are minimized. Furthermore, because photolithographic processes result in highly repeatable patterns, whatever small parasitic impedance is encountered in the semiconductor–transmission-line interface can be controlled to within extremely tight tolerances from circuit to circuit.

A microwave microelectronic X-band phase shifter fabricated by STD processing is shown in Fig. 20(A).* A magnified view of the transmission-line connections to the p-i-n diodes is shown in Fig. 20(B). The diodes are bonded to transmission lines on the substrate for both electrical and thermal conduction; since the diodes are nonplanar—that is, the anode and cathode are on opposite faces of the chip—the multilayer capability of STD is used to bring transmission lines to opposite sides of the chip. A multilayer capability, coupled with the high resolution of photolithography, can also be used to fabricate interesting two-layer interactive microelectronic transmission-line structures, such as hybrid couplers, for example.

The ability of second-level interconnect processes to span the range of applications from low frequency through microwave fulfills an important need in integrated electronic technology—namely, the desirability to standardize basic processes while maintaining sufficient flexibility to allow optimization at the equipment-system levels of design. Further, the use of photolithographic processes at all levels of interconnections, from semiconductor through printed boards, makes possible the full utilization of the heavy investment in computer-aided design for layout and mask fabrication. As electronics enters the high-complexity-per-circuit era, software costs associated with design and test may increasingly overshadow the costs of hardware fabrication.

Some challenges

The complexity of LSI chips, taken together with the fact that most elements on the chip are inaccessible to testing, imposes severe strains on design synthesis and analysis. Large computers have become indispensable tools in the design of SIC layout patterns, in the generation of diagnostic methods for input–output testing, and in modeling or simulation.[12-14] However, the largest computers are still incapable of performing complete analyses in a large SIC array and, despite the important gains which have been made in computer-aided design (CAD),[15] automated SIC synthesis, even for simple configurations, is a long way from realization. Hence, a major challenge in the area of design theory is to develop methods for solving electronic design problems that involve the use of thousands of active elements in complex circuit configurations.

Referring to the industrial structure chart of Fig. 1, much of the emphasis in this discussion has centered on the intersection of integrated electronic and systems technologies. However, future advances in integrated electronics are most apt to come from the other end of the chart—the intersection with materials technologies. Just as semiconductor process development fueled the

* Circuit design and fabrication by C. Lee, J. Lunden, and J. Dietz of the General Electric Company's Electronics Laboratory.

SIC trend, other materials discoveries may substantially change the current emphasis on silicon. The optoelectronic properties of compound semiconductor materials, like gallium arsenide, make them attractive for use not only in display applications but also for the fabrication of complex LSI arrays that utilize optical–electrical interactions for interconnection or diagnostic purposes.[16, 17]

Magnetic materials, which have served as keystones for computer memories, have recently been discovered to have controllable domain states at micro levels.[18] Emphasis in the future could well shift back to magnetic integrated circuits for complex logic as well as memory functions. It may be important, therefore, to develop methods for integrating circuits made of differing material substrates in future systems.

A major challenge in the area of applications outside military and industrial uses concerns the potential of LSI in consumer products. It should be possible to achieve assemblies of 10 000 to 50 000 transistors at costs compatible with consumer budgets, but how to utilize this level of electronic complexity in order to fulfill consumer wants is not at all clear. Perhaps, as James Hillier of RCA put it, "I believe we should think more of entrepreneuring services rather than selling boxes."[19] If so, the entrepreneur should be as challenged by LSI as the electronic designer.

Industrial and academic organizations are also confronted with a major challenge—that of tying together the various traditional departments in order to effect the interdisciplinary requirements of integrated electronics. Vertical integration of industrial organizations—whereby system organizations attempt to master integrated-circuit processing and semiconductor organizations attempt to move into system applications—may in the end result only in verifying that the Second Law of Thermodynamics applies to men as well as to machines. It seems that only the very large companies may have the resources to be expert across the board and even such large companies may find their broad-based expertness limited to narrow market applications. The most important organizational challenge, it would seem, is for chemists, metallurgists, physicists, circuit engineers, systems experts, economists, marketing specialists, government and business leaders, and others to learn to communicate with one another. Interdisciplinary teams, having relatively large supporting organizations in each discipline behind them, will have to learn how to make the tradeoffs that ultimately define the mask sets for chips, hybrid circuits, and printed boards. And in order for this to happen, the universities will have to start turning out specialists who are able to talk to each other.

In 20 years, we have come from the vacuum tube to the transistor to the integrated circuit. Even the brilliant John von Neumann could not predict the pace of progress. Consider his statement from a classic technical paper: "We saw earlier that a fast memory of several thousand words is not at all unreasonable for an all-purpose instrument. Hence about 10^5 flip-flops or analogous elements would be required! This would, of course, be entirely impractical."[20] An attempt to predict where integrated electronics will take us 20 years from now is probably futile. But as electronics enters the age of large numbers, it is possible to predict that unless man can keep up with the change, and benefit thereby, the tremendous potential of integrated electronics will be dissipated.

The author is indebted to many colleagues in General Electric Company's Electronics Laboratory for providing the information used in this article. In particular, the assistance of Richard Moeller is gratefully acknowledged for help in obtaining the illustrative material.

REFERENCES

1. Danielson, G. H., Schindler, H. R., and Zampini, J. J., "Microelectronic frequency discrimination techniques," *IEEE (London) Conf. Pub. No. 30*, pp. 215–226, 1967.

2. Danielson, G. H., "Frequency selective circuits for a microelectronic FSK receiver," *Proc. IEEE*, vol. 54, pp. 435–436, Mar. 1966.

3. Thornton, C. G., "LSI subsystems assembled by the silicon wafer-chip technique," *Dig. Tech Papers 1968 Internat'l Solid State Circuits Conf.*

4. Hochman, H. T., "The art of building LSIs," *IEEE Spectrum*, vol. 6, pp. 29–36, Sept. 1969.

5. Greer, D. L., Private communication.

6. Peil, W., Private communication.

7. For an excellent description of integrated-circuit manufacturing techniques see *The Engineer* (Western Electric Co.), Dec. 1967.

8. Thun, R. E., "Thick films or thin?" *IEEE Spectrum*, vol. 6, pp. 73–77, Oct. 1969.

9. Morton, J. A., "Strategy and tactics for integrated electronics," *IEEE Spectrum*, vol. 6, pp. 26–33, June 1969.

10. Miller, L. F., "Controlled collapse reflow chip joining," *IBM J. Res. Develop.*, May 1969.

11. Suran, J. J., Kim, C. S., and Palmer, G. G., "Plastic structured microelectronics," *Dig. Tech. Papers 1968 Internat'l Solid State Circuits Conf.*

12. Spitalny, A., and Goldberg, M. J., "On-line graphics applied to layout design of integrated circuits," *Proc. IEEE*, vol. 55, pp. 1982–1988, Nov. 1967.

13. Weinberger, A., "Large-scale integration of MOS complex logic: a layout method," *IEEE J. Solid-State Circuits*, vol. SC-2, pp. 182–190, Dec. 1967.

14. Jones, E. R., and Mays, C. H., "Automatic test generation methods for large-scale integrated logic," *IEEE J. Solid-State Circuits*, vol. SC-2, pp. 221–226, Dec. 1967.

15. Calahan, D. A., *Computer Aided Network Design*. New York: McGraw-Hill, 1968.

16. Loebner, E. E., "Opto-electronic devices and networks," *Proc. IRE*, vol. 43, pp. 1897–1906, Dec. 1955.

17. Stewart, R. D., "Optoelectronic memories: light to read out by," *Electronics*, vol. 42, no. 5, Mar. 3, 1969.

18. Bobeck, A. H., "Properties of cylindrical magnetic domains in orthoferrites," *Abstr. 1968 Intermag. Conf.*

19. O'Donnell, J. D., *et al.*, "Electronically expanding the citizen's world," *IEEE Spectrum*, vol. 6, pp. 30–40, July 1969.

20. Goldstine, H. H., and von Neumann, J., "On the principles of large-scale computing machines," May 15, 1946, in *John von Neumann Collected Works*, Vol. 5. New York: Macmillan, 1963.

J. J. Suran (F) is manager of the Electronic Applications and Devices Laboratory at the General Electric Company's Electronics Laboratory facility in Syracuse, N.Y. He is responsible for technical and financial program planning, major project organization, integration of laboratory efforts with company product departments, and evaluation of laboratory effectiveness. Prior to joining General Electric, Mr. Suran held engineering positions at the J. W. Meaker Company, where he was engaged in control system development, and at Motorola Inc., where his work was in research and development of mobile communication equipment. From 1959 to 1963, in addition to his work at General Electric, he served as a nonresident member of the M.I.T. faculty. Mr. Suran holds the B.S.E.E. degree from Columbia University and has done graduate work there and at the Illinois Institute of Technology. He is the coauthor of two books, has published more than 35 papers, and holds 18 patents.

A philatelic history of electrical science

There is a wealth of history to be garnered from the postage stamps of the world about electrical and electronic discovery over the past centuries

Sidney V. Soanes Ferranti-Packard Ltd.

Because countries are proud of their traditions, heritage, and accomplishments and want the rest of the world to know about them, they find it useful and convenient to disseminate such information on tiny testaments called stamps. The philatelic approach to recording history is studded with acknowledgments of the electrical and electronic accomplishments and includes events dating back to the early 17th century. Although the number of stamps commemorating such events is far too numerous to describe adequately in this article, some of the more noteworthy events and inventors are described with reference to their appearance on postal engravings.

Of all the advertising and propaganda mediums ever devised, the postage stamp is certainly one of the most effective. These little pieces of paper carry their message to all parts of the world. It is no surprise, therefore, to find nearly every country using them to full advantage to publicize the country's culture, its industry, and its political views and ambitions. Millions of stamp collectors eagerly seek out these labels; indeed, some countries print many, many more than they really need, just for sale to philatelists!

Nearly every imaginable subject and a few unimaginable ones have been represented on postage stamps. A great many (several thousand) are of special interest to members of the IEEE—documenting people and events in the history of electrical and related sciences. It would be quite impossible to describe more than a small percentage of these stamps without using up most of this issue. The choice has been made quite arbitrarily, and the author offers profuse apologies if the reader's particular favorite has been omitted. An effort will be made, however, to show the very wide scope of subject matter covered, and a few of the more interesting stamps are illustrated. It should also be kept in mind that some of the more important people and events have been commemorated on several stamps (often from many different countries) and then only one or two examples are generally given.

The pioneers

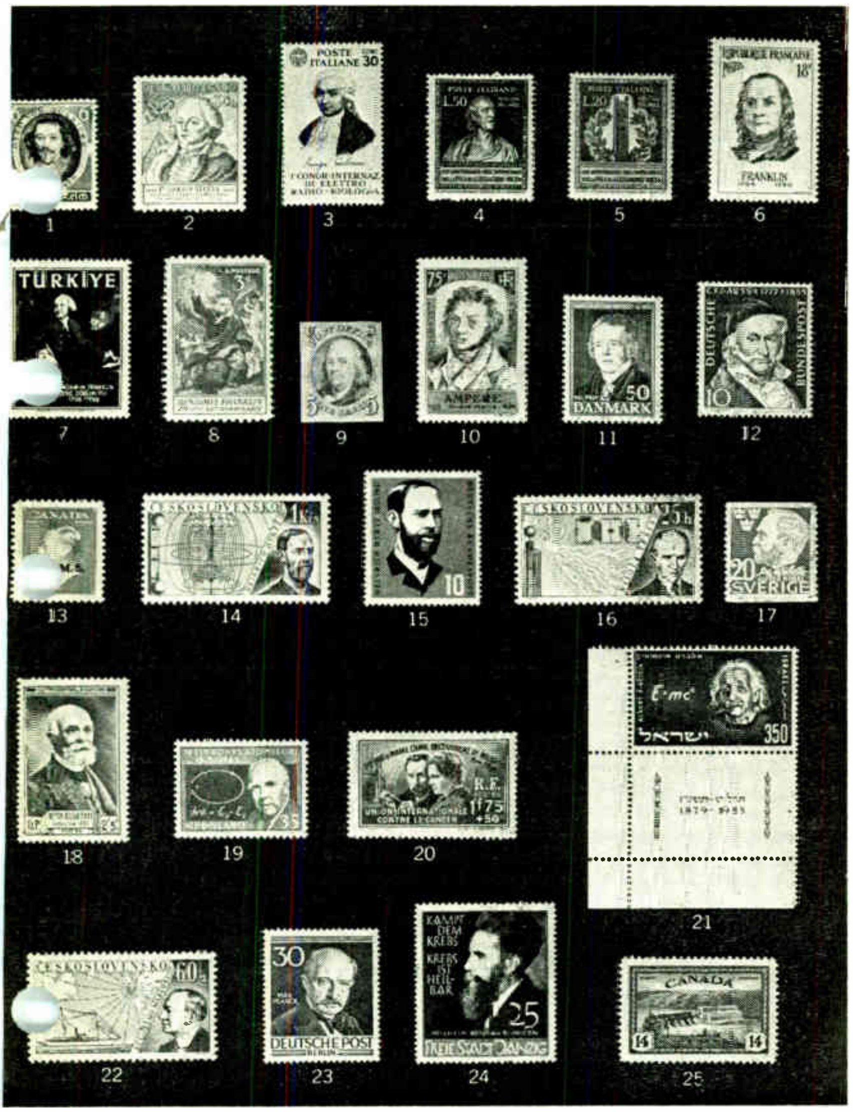
Our philatelic history of electrical science begins with events of the 17th century. It is true that the ancient Greeks recognized, but did not understand, the phenomenons of electric and magnetic attraction. Also, there is strong evidence that the Chinese used some form of magnetic compass about 1000 B.C.—possibly even as early as 2600 B.C. But progress and understanding developed very slowly until, at the end of the 16th century, William Gilbert, physician to Queen Elizabeth, published a treatise on electricity and magnetism. (Gilbert has rightly been called the "father of modern electricity" and the "Galileo of magnetism.") Then things started to happen.

Otto von Guericke (1602–1680) was a German natural philosopher, having studied law and mathematics. He was a combination politician, physicist, and astronomer, and was notably successful in all three pursuits. He probably deserves credit for inventing the first electric generator, a crude hand-operated electrostatic machine. He is commemorated on a German stamp issued in 1936, the 250th anniversary of his death (**1**, *472**).

Another major 17th century contribution came from Holland. Christiaan Huygens (1629–1695) (his father, Sir Constantijn Huygens, was probably the most brilliant figure in Dutch literary history) became an eminent mathematician–astronomer–physicist. He invented the pendulum clock when he was 27 years old. His studies of centrifugal force enabled him to predict the equatorial bulge of the earth to about half of its actual amount. Huygens is best known, however, for his work in optics. Not only did he make many important advances in the technology of lens grinding, but he also was an exacting theorist. Undoubtedly, his greatest contribution to posterity, and the one of most importance to electrical science today, was his development of the wave theory of light. The Netherlands issued a stamp in his honor in 1928 (*B36*).

The next century saw the first real glimmers of understanding of electricity and the possibilities of its practical use. Some of the experimenters' names have found their way into our modern scientific dictionaries: Prokop Diviš (1698–1765)—a Czech naturalist who discovered that lightning is atmospheric electricity (Czechoslovakia, **2**,

* All stamp numbers (*italics*) in this paper have been taken from *Scott's Standard Postage Stamp Catalog* and are used with the permission of the publishers, Scott Publications, New York. Numbers in boldface, when they appear, correspond to those stamps pictured on these pages and relate to the italics that follow.



(1)	Germany	472	Otto von Guericke
(2)	Czecho-slovakia	661	Prokop Diviš
(3)	Italy	329	Luigi Galvani
(4)	Italy	527	Alessandro Volta
(5)	Italy	526	Voltaic pile
(6)	France	814	Benjamin Franklin
(7)	Turkey	1259	Benjamin Franklin
(8)	U.S.A.	1073	Benjamin Franklin
(9)	U.S.A.	1	Benjamin Franklin
(10)	France	306	André Marie Ampère
(11)	Denmark	329	Hans Christian Oersted
(12)	Germany	725	Karl Friedrich Gauss
(13)	Canada	015	OHMS
(14)	Czecho-slovakia	953	Heinrich Rudolf Hertz
(15)	Germany	762	Heinrich Rudolf Hertz
(16)	Czecho-slovakia	949	Nikola Tesla
(17)	Sweden	380	Alfred Nobel
(18)	France	B202	Henri Becquerel
(19)	Greenland	57	Niels Bohr
(20)	France	B76	Pierre and Marie Curie
(21)	Israel	117	Albert Einstein
(22)	Czecho-slovakia	952	Guglielmo Marconi
(23)	Germany	9N92	Max Planck
(24)	Danzig	240	Wilhelm Roentgen
(25)	Canada	270	Hydroelectric station

661, 662); Luigi Galvani (1737–1798)—an Italian physiologist, best known for his work on muscle twitching due to electric current in frogs' legs (Italy, **3**, *329*, *330*); Alessandro Volta (1745–1827)—an Italian physicist (later made a senator of the Kingdom of Lombardy), who invented the first electric battery and is also credited with the first electric capacitor and the hydrogen lamp. Volta's picture appears on several Italian and Italian-colony stamps (Italy, *188–191*, **4**, *527*, *et al.*). One Italian stamp, issued in 1949, shows a picture of his voltaic pile (**5**, *526*).

Benjamin Franklin (1706–1790) really needs no introduction. Quite aside from his scientific achievements, he was an important figure in the history of the American colonies, holding several responsible positions including that of postmaster general from 1753 to 1774. We can be thankful that he was not electrocuted doing his famous kite experiment. He has been honored philatelically by several countries, including France (**6**, *814*—one of a set of stamps issued in 1956 in honor of famous men who lived in France), Turkey (**7**, *1259*, *1260*—to commemorate the 250th anniversary of his birth), and the U.S.A. (**8**, *1073*—also for the 250th anniversary of his birth. The United States has also issued dozens of other noncommemorative stamps showing his picture, including **9**, *1*).

Continuing into the golden age of the 19th century, we have André Marie Ampère (1775–1836), who established the relationship between magnetism and electricity (France, **10**, *306*, *326*) by expanding upon the work of Danish pharmacist and physicist Hans Christian Oersted (1775–1851) (Denmark, **11**, *329*).

Karl Friedrich Gauss (1777–1855), German mathematician, physicist, and astronomer, proposed a system of absolute units based on length, mass, and time (Germany, **12**, *725*). Georg Simon Ohm (1787–1854), a German physicist, determined the relationship between voltage and current in a resistance and developed the law that bears his name—a law very unsympathetically received at first. There appears to be no stamp showing Ohm's portrait, but (if the reader will excuse a small jest) his name appears on many stamps for official government use with the overprint OHMS (On His Majesty's Service). (See, for example, Canada, **13**, *O15*.) Hermann Ludwig Ferdinand von Helmholtz (1821–1894) was a German philosopher and scientist whose specialty, interestingly enough, was physiology. He wrote a book, *Physiological Optics*, of exceedingly great importance. As a physicist he can be considered one of the founders of the law of conservation of energy. His other significant investigations included sound, hydrodynamics, meteorology, and elec-

(26)	Russia	1794	Atomic electric station
(27)	Greece	734	Agra River hydro-electric station
(28)	Switzerland	328	High-voltage distribution line
(29)	Italy	851	Pacinotti's dynamo
(30)	Germany	966	Siemens' dynamo
(31)	Belgium	217	Zenobe Gramme and his dynamo
(32)	Germany	965	Three-phase power transmission
(33)	U.S.A.	654	Edison's first electric lamp
(34)	Japan	577	Japan's first electric lamp
(35)	Japan	827	Tokaido Railroad
(36)	Switzerland	412	Trans-Europe express
(37)	Peru	408	Samuel Morse
(38)	U.S.A.	924	Telegraph centenary
(39)	Central African Rep.	46	Claude Chappe's telegraph
(40)	Gabon	180	Morse telegraph
(41)	Mongolia	63	Telegraph operator
(42)	Switzerland	340	Morse code
(43)	Romania	1020	Teletypewriter
(44)	Bulgaria	B19	Lineman
(45)	U.S.A.	1112	Atlantic cable centenary
(46)	Canada	274	Alexander Graham Bell
(47)	Germany	846	Reis telephone
(48)	Germany	B59	Telephone
(49)	Ryukyu	39	Telephone and dial
(50)	Netherlands	391	Dial system

trodynamics. In 1871 he showed that the velocity of propagation of electrodynamic induction was greater than 314 000 m/s (East Germany, *62*). One of his pupils was Heinrich Rudolf Hertz (1857–1894), whose work on electromagnetic waves is well known (Czechoslovakia, **14**, *953;* Germany, **15**, *762*).

Germany did not quite have a monopoly on electrical discovery in the 19th century. Nikola Tesla (1856–1943) was a Serb, although most of his work—largely related to the use of alternating current—was done in the United States. He is credited with many useful inventions and in 1888 received a patent for the induction motor (Yugoslavia, *136*, *137*, *373*, and *374*). He also appears on the Czechoslovakian stamp (**16**, *949*) because he studied in Prague for a time.

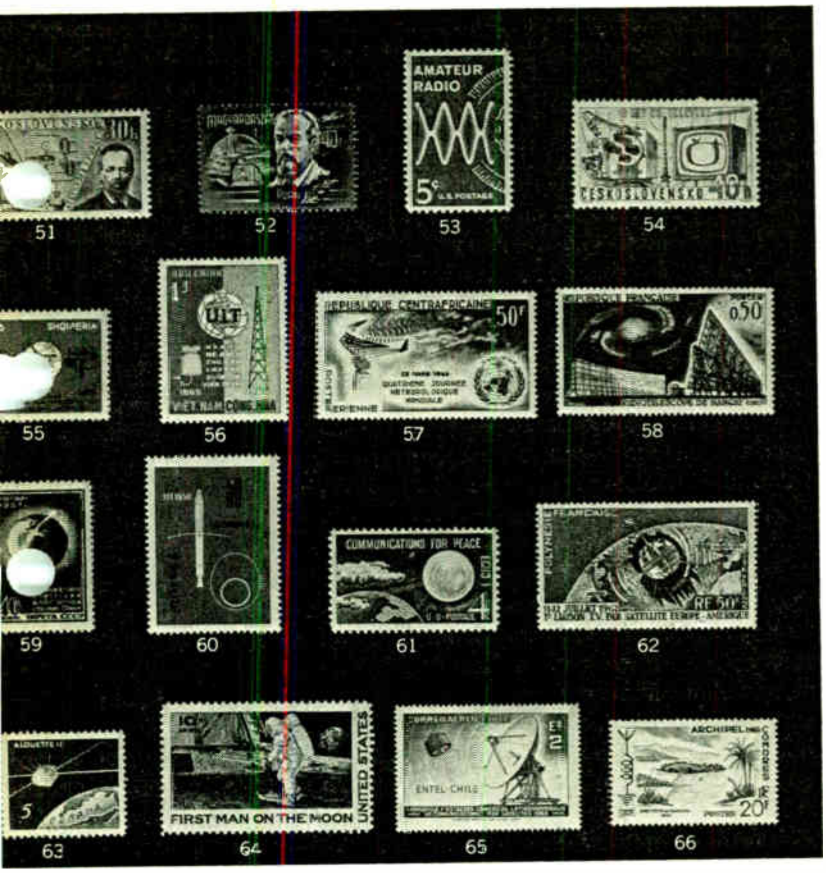
The Nobel Prize winners

Special mention is due those who have won the Nobel Prize. Alfred Nobel, himself, appears on a set of Swedish stamps (**17**, *380–382*). Several of the Nobel Prize winners in physics made important contributions to our knowledge of electrical science, particularly in extending our understanding of the atom and how it behaves. Some of those who have been honored on postage stamps include: Henri Becquerel (France, **18**, *B202;* Sweden, *638*); Niels Bohr (Denmark, *409*, *410;* Greenland, **19**, *57*, *58*); Pierre and Marie Curie (Afghanistan, *RA2;* Monaco, *B24;* Panama, *RA1–4*, *et al.;* Sweden, *638*, France, **20**, *B76*, plus about 21 other countries in the French community.) Pierre appears alone on Bulgaria *957;* Rumania *1126*, Russia *1883;* Marie, who was born in Warsaw, appears alone on Poland *401* plus about five others; Surinam *B49*, *B52;* Turkey *B67*.

Other honored Nobel scientists include: Albert Einstein (Ghana, *190;* Israel, **21**, *117;* Poland, *882;* U.S.A., *1285*); Guglielmo Marconi (Czechoslovakia, **22**, *952*, plus others including those listed on page 84 under "Radio and television"); Max Planck (Germany, **23**, *9N92*, plus others); Wilhelm Roentgen (Danzig, **24**, *240*, plus others); J. J. Thomson (Sweden, *710*, *712*).

Electric power

Basic to all industry is the generation and distribution of electric power. Many stamps show generating stations and high-voltage distribution lines. Hydroelectric stations are shown on Canada **25**, *270;* China *1410;* Czechoslovakia



(51)	Czechoslovakia	950	Aleksander Popov
(52)	Hungary	C62	Aleksander Popov
(53)	U.S.A.	1260	Amateur radio
(54)	Czechoslovakia	1167	Television
(55)	Albania	814	ITU centenary
(56)	Vietnam	253	ITU centenary
(57)	Central African Rep.	C18	Radar
(58)	France	1067	Radio telescope
(59)	Russia	1992	Sputnik I
(60)	Poland	1180	Explorer I
(61)	U.S.A.	1173	Echo I
(62)	French Polynesia	C29	Telstar I
(63)	Canada	445	Alouette II
(64)	U.S.A.	C76	Apollo 11
(65)	Chile	C290	Satellite communications station
(66)	Comoro Islands	46	Comoro radio station

731; Russia *1277–1279, 1598.* An atomic electric station appears on Russia **26**, *1794, 1795,* and *1796.* The Rance River tidal power station appears on France *1170;* Lebanon *AP60* shows an electric generating station of an unstated kind. In April 1962, Greece issued a set of seven stamps to publicize Greece's electrification project (*728–734*). The highest valued (**27**, *734*) shows the generators inside the Agra River hydroelectric station. A high-voltage distribution line is shown on Switzerland **28**, *328.*

The development of the dynamo is exemplified by the following selection of stamps: Italy **29**, *851, 852*—commemorating the 50th anniversary of the death of Antonio Pacinotti who invented the continuous-current dynamo in 1861; (Pacinotti, himself, is shown on Italy, *322, 323*); Germany **30**, *966*—commemorating the centenary of the discovery of self-excitation by Werner von Siemens; Belgium **31**, *217*—for the Liège Exhibition of 1930, 60 years after Zenobe Gramme combined Pacinotti's and Siemens' principles into the first commercially practicable dynamo. The 75th anniversary of three-phase power transmission is depicted on a German stamp (**32**, *965*).

The electric light. In 1929 the U.S.A. issued three stamps (**33**, *654, 655,* and *656*) to commemorate the golden anniversary of the invention of the incandescent lamp by Thomas Edison on October 21, 1879. Edison, himself, appears on U.S.A. *945* issued February 11, 1947, upon the centenary of his birth. Japan issued a stamp (**34**, *577*) in 1953 to commemorate the 75th anniversary of electric lighting in that country.

Electric railroads. A great many countries have electric railroads and, of course, these are often publicized on their stamps. Most of the Belgian parcel-post stamps issued during the past 30 years have shown parts of their railway system. Many of these are excellent examples of electric traction (for example, *Q327, 337, 341, 359* and others). In 1956, Japan issued a stamp (*632*) to commemorate the electrification of the Tokaido line, and eight years later issued another (**35**, *827*) for the opening of the new Tokaido line. Switzerland is another country largely dependent on the electric railroad for its transportation (shown on *310*, **36**, *412*). The former commemorates the centenary of the opening of the first Swiss railroad, between Zurich and Baden, and shows the Gotthard Express. The other stamp was issued to publicize the introduction of the Swiss electric TEE trains and shows the Trans-Europe Express.

Closely related to these are Belgium *265* showing an electric railroad signal and Russia *1411* showing an electric trolley coach (1949).

Telecommunications

A separate book could be written about the influence of telecommunications on the postage stamps of the world.

The telegraph. We think of Samuel F. B. Morse as the inventor of the telegraph, and several countries have honored him on their stamps, including Argentina (*B1*), Peru (**37**, *407, 408*), and the U.S.A. (*890*). In 1944, the United States commemorated the centenary of the Morse telegraph with a special stamp (**38**, *924*).

Actually a visual-telegraph system was invented in France about 1790 by the Chappe brothers, and it became widely used. Claude Chappe appears on France *474* and

625. Chappe's apparatus is shown on a stamp issued in 1965 by the Central African Republic as one of a set commemorating the centenary of the International Telecommunications Union (**39**, *46*). The Morse telegraph is shown on a stamp issued by Gabon for the same purpose (**40**, *180*). Many other countries have issued stamps showing telegraph apparatus and wire lines, including Austria (*495*), Australia (*270*), Chad (*112*, *114*), Denmark (*352*), Hungary (*C184*), Japan (*604*), Mongolia (**41**, *63*), Nicaragua (*20*), Paraguay (*436*, *C135*), Portugal (*813–815*), and others. The Morse code appears on stamps from Russia (*347*, *348*) and Switzerland (**42**, *340*). Teletypewriters are shown by East Germany (*161*) and Romania (**43**, *1020*). Some of the early workers in this field who have been honored on postage stamps include Baudot (France, *627*) and Ferrié (Austria, *B150*).

The telegraph repairmen have not been forgotten; see, for example, Austria *B276*, Bulgaria **44**, *B19;* Croatia *B78;* Hungary *C91;* Yugoslavia *B106;* and Romania *592*.

Another stamp of interest in the telegraph-communications category is U.S.A. **45**, *1112* issued in 1958 for the centenary of the first Atlantic cable.

The telephone. In March 1876, Scottish-born Alexander Graham Bell (1847–1922) became the first person to transmit speech over an electric wire, and he is, therefore, generally recognized as inventor of the telephone. He has been honored philatelically by at least three countries: Argentina (*B3*), Canada (**46**, *274*), and the U.S.A. (*893*). Five years before Bell's success a German, Johann Philipp Reis (Germany, *693*), developed an instrument that transmitted single tones not speech (Germany, **47**, *846*).

A number of stamps feature telephones and other aspects of telephone plants; only a few can be listed.

- *Telephones:* Argentina (*597*); Chad (*113*); Germany (**48**, *B59*—interesting because it shows a man talking into a handset with no connecting wire!); Saar (*210*); Syria (*C175–C177*); Ryukyu (**49**, *39*); Switzerland (*341*).
- *Switchboards:* Bulgaria (*B20*), Croatia (*B79*).
- *Radiotelephone exchange:* French West Africa (*C19*).
- *A few others of interest:* Australia (*157*, *158*)—commemorating the telephone link between Australia and Tasmania; Japan (*859*)—the 75th anniversary of telephone service (1890–1965); and the Netherlands (**50**, *391–393*)—completion of their dial system.

Radio and television. We think of Marchese Guglielmo Marconi (1874–1937) as being the "father of radio." In 1895 he succeeded in transmitting signals without a connecting wire. His home country, Italy, honored him with a set of three stamps one year after his death (*397–399*). A Newfoundland stamp (*155*) shows Cabot Tower on Signal Hill in St. John's, where Marconi received the first transatlantic radio signal on December 12, 1901. He received the Nobel Prize for physics in 1909. Other stamps showing Marconi are Italy *909* and Monaco *615*. (See also page 82 under Nobel Prize winners.)

Aleksander Popov (1859–1905), a Russian physicist, and a contemporary of Marconi also demonstrated the possibilities of radiotelegraphy. Several stamps show Popov—mostly Russian ones—(*328*, *329*, *989–991*, *1352–1354*). But Bulgaria (*722–723*), Czechoslovakia (**51**, *950*), and Hungary (**52**, *C62*) have also featured his likeness.

Other stamps featuring radio technology include: Brazil *855*—Sarapui Central Radio Station; China *1168–1170*—30th anniversary of China broadcasting in 1957; Czechoslovakia *954*—honoring E. H. Armstrong; Guatemala *C216*—radio tower; Hungary *1119*—girl assembling a radio set; Japan *499*—25th anniversary of Japanese broadcasting in 1950; Russia *2063*—radio tower; and the U.S.A. **53**, *1260*—50th anniversary of the ARRL, and *1329*—25th anniversary of the Voice of America.

Stamps particularly related to television include: Czechoslovakia **54**, *1167*, *1168*—tenth anniversary of Czech television; France *766;* Germany *770*; Hungary *1181*, *1182*, *1390;* Italy *649*, *650;* and Switzerland *343*.

Special telecommunications events. In 1965 the International Telecommunications Union celebrated its centenary, and this event was honored by about 200 stamps from more than 130 countries. Only a very few of these countries can be mentioned here: Albania (**55**, *814*, *815*), Andorra (*167*), Antigua (*153*, *154*), Barbados (*265*, *266*), Ethiopia (*439–441*), Libya (*279–281*), Mauritius (*291*, *292*),

I. Some stamps commemorating space flight

Subject	Event	Country Issuing*	Catalog Numbers
Sputnik I	U.S.S.R.—Oct. 4, 1957—first artificial satellite in orbit	Russia	1992 (**59**)† plus about 30 others
Explorer I	U.S.A.—Jan. 31, 1958—first U.S. satellite in orbit	Poland	1180 (**60**) plus about 5 others
Lunik II	U.S.S.R.—Sept. 12, 1959—first moon landing	Albania	C69 plus about 15 others
Echo I	U.S.A.—Aug. 12, 1960—first communications satellite	U.S.A.	1173 (**61**) plus about 3 others
Vostok I	U.S.S.R.—Apr. 12, 1961—first man in space, almost one orbit	Surinam	C28 plus about 68 others
Friendship 7	U.S.A.—Feb. 20, 1962—first U.S. astronaut to orbit earth (three orbits)	U.S.A.	1193 plus about 46 others
Telstar I	U.S.A.—July 10, 1962—telecommunications satellite	French Polynesia	C29 (**62**) plus about 50 others
Syncom I	U.S.A.—Feb. 4, 1963—telecommunications satellite	Liberia	415 plus about 7 others
Alouette II	Canada—Nov. 28, 1965—"Topside Sounder," part of the Canada-U.S. program of space research	Canada Poland	445 (**63**) 1471
Surveyor I	U.S.A.—May 30, 1966—first U.S. spaceship to soft-land on the moon	Togo	563 plus about 11 others
Lani Bird II	Japan-U.S.A.—Jan. 11, 1967—communications satellite	Japan	904
Apollo 8	U.S.A.—Dec. 21, 1968—first satellite to put men into orbit around the moon	U.S.A.	1371 plus about 20 others
Apollo 11	U.S.A.—July 20, 1969—man lands on the moon!!	U.S.A.	C76 an (**64**) plus an uncounted number of others

* Refers only to stamps specified by numbers. "Others" may originate in other countries.
† Boldface numbers in parentheses refer to numbers in photograph.

Stamps have often been used to advance the claims of various nations to priority of invention. Italy, Germany, and Yugoslavia separately commemorate the beginnings of rotating machinery through the contributions of their respective nationals, Pacinotti (67), Siemens (68), and Tesla (69); Italy and the Soviet Union vie for the honor of having "invented radio" in the work of Marconi (70) and Popov (71).

Ruanda (*109–113*), Seychelles (*218, 219*), Thailand (*430*), Vietnam (**56**, *253, 254*).

Many of the smaller countries are especially proud of their membership in the ITU. For example, in 1957 Korea issued two stamps (*243, 244*) commemorating the fifth anniversary of their admission to the ITU. Five years later it issued a tenth anniversary stamp (*348*).

The opening of a new communications link is always good cause for rejoicing and celebrating and a new stamp or two! Two philatelic examples are COMPAC (Commonwealth Pacific Telephone Cable December 3, 1963)—issued by Australia (*381*), Fiji (*192*), Great Britain (*401*), New Zealand (*364*)—and SEACOM (South-East Asia Commonwealth Cable, March 30, 1967)—issued by Hong Kong (*236*) and Malaysia (*42, 43*).

Radar and radio astronomy. These two categories do not quite fall into the previous classification of telecommunications. They are lumped together here only because there are very few stamps relating to these topics. Most of them are listed in the following:

Barbados *303*, Central African Republic **57**, *C18*, and the United Nations *188, 189*—showing a radar antenna (issued for World Meteorological Day).

Iran *1393*—radar dish pictured to commemorate the inauguration of the radio telecommunication system of the Central Treaty Organization of the Middle East.

Czechoslovakia *836*—radio telescope featured for the International Geophysical year.

France **58**, *1067*—radio telescope at Nançay.

The space age

Since Sputnik I was launched on October 4, 1957, nearly 600 satellites have been sent into space, of which some 100 have been commemorated on stamps. The total number of these stamps is close to 1500! It is especially interesting to note when looking at the sampling in Table I that many of the North American satellites appear on stamps from "Iron Curtain" countries. Undoubtedly, there will be many more such stamps.

A recent pair of stamps by Chile (*375*, **65**, *C290*) was issued to publicize the inauguration of ENTEL–Chile, the first commercial satellite communications ground station at Longovilo.

The emerging nations

We have witnessed more new countries come into existence in the past ten years than possibly in any previous decade in man's history. They are all very proud of their new technological progress, and their achievements are well documented on their postage stamps. Again, a complete listing is quite impossible.

Upper Volta *190*—issued to commemorate the opening of the automatic telephone office in Bobo–Dioulasso (September 1968).

Guyana *64–67*—issued to publicize the Guyana–Trinidad tropospheric-scatter radio link.

Comoro Islands **66**, *46, 47*—issued to publicize the Comoro radio station.

Tanzania *5*—showing the Hale hydroelectric plant.

Lesotho *68*—Radio Lesotho.

It has been quite impossible, in this relatively short article, to do more than introduce the reader to a few of the more obvious and, it is hoped, interesting aspects of the hobby of philately as it relates to our profession. If the reader has learned some detail of the history of electricity he did not know before, then this is an added bonus. Unlike most other SPECTRUM articles, this one has been intended primarily to entertain rather than to educate.

Ideas on which this article has been based have come from many sources, including *Encyclopaedia Britannica*, *Minkus Stamp Journal*, *Radio Philatelia* (by Herbert Rosen), *Topical Time* (*Journal of the American Topical Association*), and *Scott's Standard Postage Stamp Catalog*. All photographs are by J. V. Scott of the Ryerson Polytechnical Institute.

Sidney Vincent Soanes (SM) is a mathematician-physicist by vocation and a naturalist (among other pursuits) by avocation. This blend of interests is reflected by his membership in an impressive total of 13 scientific and professional societies—a list that also attests to his dedicated work for national and international technical standards. And as this article suggests, Dr. Soanes is an amateur philatelist. Although Dr. Soanes' interests are varied, his loyalty is not. For the past 20 years—since receiving the Ph.D. from the University of Toronto, Toronto, Ont.—he has worked for Ferranti-Packard Ltd., Toronto. His present position there is as staff engineering supervisor for the Electronics Division—a position that he has held for the past two years. Dr. Soanes counts more than 18 articles and conference presentations among his varied technical contributions to the electrical and the electronics fields.

New product applications

New lock-in amplifier recovers repetitive signals in the presence of noise without tuning

The System 82 Lock-In Amplifier was designed to give an uncomplicated means for the recovery of repetitive signals in the presence of noise. It consists of the Model 822 Phase Sensitive Detector, the 823 Nanovolt Amplifier, and 821 Phase Shifter.

Typical applications include the measurement of nuclear magnetic resonance, Hall effect, contact potentials, photodetector signals, molecular beams, noise bandwidth characteristics of solid-state devices, and complex permittivity. It can also be used in ac bridge null detection, and for electrospin paramagnetic studies.

The $1895 instrument uses a wideband approach that eliminates the need for tuning. The instrument has a range of 5 nV to 0.3 volt, a frequency range of 1 Hz to 1 MHz, a gain of 30 dB to 120 dB, a linearity of ±0.05 percent, and the capability of noise rejection to 70 dB.

A typical application is the automatic measurement of semiconductor junction capacitance. Continuous measurement of the capacitance permits its variations to be recorded against changes of temperature or applied voltage without the need for point-by-point plotting. It is generally possible to make the measurements at any frequency between 1 kHz and 1 MHz with only minor changes in setting up of the apparatus for each frequency.

The simplest measurement principle is that using the circuit shown in Fig. 1. An alternating voltage V_1 is applied to the test capacitance C_x in series with a standard capacitor C_o. The voltage measured across C_o is

$$V_o = V_1 \cdot C_x/(C_o + C_x)$$

If $C_o > C_x$

$$V_o = V_1 \cdot C_x/C_o$$

and, therefore, if C_o has a fixed value

$$|V_o/V_1| = AC_x$$

where A is a constant.

This basic circuit is capable of three-terminal measurement because stray admittances appear only across the voltage generator, which is by definition low impedance, and across C_o, which can be low impedance also.

It is necessary to consider the effect of conductance in parallel with C_x. If the conductance $G_x < C_o$, then

$$V_o = V_1 \cdot \frac{-jG_x/\omega + C_x}{C_o}$$

which can be represented by two voltages; one in phase with V_1 or V_1C_x/C_o and one in quadrature with V_1 or $V_1G_x/\omega C_o$. Both voltages may be measured with a phase-sensitive detector, by giving the detector a reference voltage in phase with V_1, in the first case, and a reference voltage in quadrature with V_1 in the second case. In this respect, the use of a phase-sensitive detector is mandatory. However, it also brings further advantages.

It is desirable to keep the applied voltage V_1 sufficiently low so that any point on the capacitance voltage curve can be measured with the required precision. The final voltage V_o is always considerably less than V_1 and the measuring equipment usually has to recover this signal from noise. The narrow-band properties of the detector enable it to enhance the SNR while giving negligible dc offset due to noise voltages. The ideal measurement system is shown in Fig. 2.

A second method uses the circuit shown in Fig. 3. This circuit is useful when it is necessary to apply direct voltages to the semiconductor junction. The assumptions that need to be made are: $(R_1 = R) < G_x$ so that the entire direct voltage appears across the device, stray admittances are low compared with $1/R$ and $1/R_1$. $1/\omega C_1 < R_1$, and the frequency stability is consistent with the required accuracy of measurement.

The disadvantages of the circuit of Fig. 3 are sensitivity to stray admittances and the dependence on frequency of the capacitance reading. These disadvantages can be eliminated by using an operational amplifier as shown in Fig .4.

More information is available from Keithley Instruments, Inc., 28775 Aurora Rd., Cleveland, Ohio 44139.

Circle No. 85 on Reader Service Card

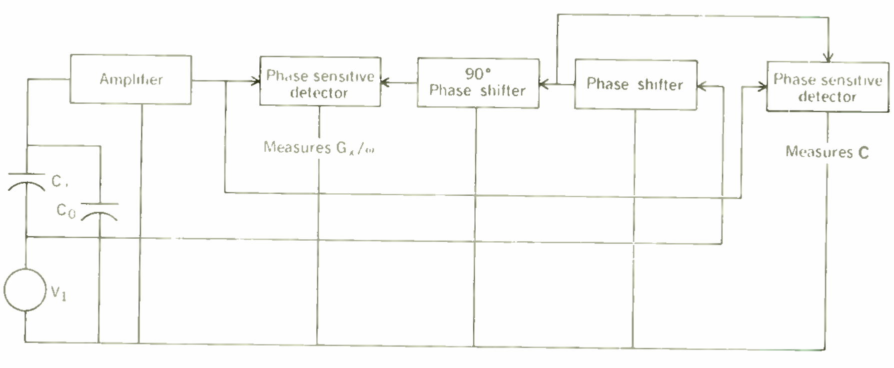


FIGURE 2. Ideal measurement system for junction capacitance.

FIGURE 3. Direct-current test circuit.

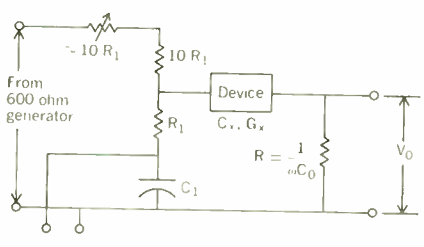


FIGURE 4. Operational amplifier circuit.

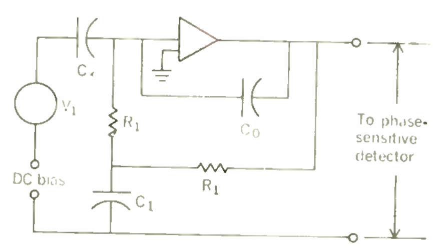


FIGURE 1. Capacitance test circuit.