VOLUME 3 □ NUMBER 6 JUNE 1966

IEEE spectrum

features

Departments: *Please turn to the next page*

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

departments

the cover

The contourograph, which makes use of the three-dimensional aspects of the cathode-ray tube, is a device for organizing and displaying certain types of data, as described in an article beginning on page 77. This month's cover shows an EKG contourogram, made with a stationary film, linear vertical sweep, and vertical signal deflection.

Spectral lines

Increased costs. Like most individuals and organizations (other than governments), professional societies must remain financially solvent. As the recent letter from President Shepherd to the IEEE membership indicated, and as the article on page 42 of this issue of SPECTRUM substantiates, the Institute must increase its dues in order to maintain a balance of income and expenditures. The recent financial history of both the predecessor societies and the merged Institute are described in detail in President Shepherd's article.

In this "Spectral Lines" article, I should like to discuss some of the implications of this dues increase with regard to the operation of the Institute. It has become increasingly apparent that the Institute cannot continue appropriate growth and development without a wider base of stable income than it has at present. There are many possible improvements in the services rendered to the members by the Institute; however, many of these improvements require greater funding than is presently available. A few examples chosen primarily in the publications area will serve to illustrate the point.

Recently, a very substantial effort has been made to minimize the delay between the submission of an article for publication and its appearance in print in the IEEE periodicals. To accomplish this it has been necessary to develop a better organized and more capable staff at Headquarters and also to obtain better schedules from the printers who serve the Institute on a contractual basis. Such improvements usually involve increased costs.

Another example is the improved method of reviewing "PROCEEDINGS Letters" that is being started. We are very anxious to keep the delay between the submission of a Letter to the Editor of PROCEEDINGS and its publication to less than eight weeks on the average. In order to meet this objective and still subject the Letter to review by an individual competent to make a good technical evaluation of the contents, it is necessary to reduce the delay introduced by the reviewing process as much as possible. The method planned is to call the reviewer by telephone after he has had a few days in which to consider the Letter. In this manner, the reviewing delay can be reduced considerably. Again, costs are increased in this process.

Another large area in which the Institute must develop know-how is in the automatic processing of technical information. It is our plan to utilize computer-controlled composition on one of our periodicals in the near future. It will be more costly at first to try this experiment but we feel that is it very desirable to gain in experience. It is also apparent that there are many other innovations possible in the not-too-distant future with the increasing availability of remote-access time-shared computers. The possibilities of developing techniques for automatic indexing of technical papers, for storing and retrieving documents by machine, and for the dissemination of technical information on a selective basis so that an individual's attention is directed only to the information in which he is interested are all in need of consideration and investigation. The IEEE cannot afford to be unaware of these possibilities and, indeed, it has the responsibility of moving aggressively ahead so that its membership has available to it as adequate an information system as is economically and technically feasible. Again, additional costs are involved.

The citing of these examples from the publications areas does not imply that there are not important and significant new ideas available in other parts of the Institute's activities. It is clear that there is a need for experimentation in areas such as the provision of visiting lecturers for local Section meetings and in efforts to do a more effective job in the continuing education of engineers and scientists, a subject which has been subject to much study recently.

The important point is that the Institute needs to have sufficient resources available so that it can exploit those ideas that prove to be useful. The Institute and its members do not appear to be idea-limited at this time; the major limitation appears to be in the resources needed to try out and develop the available ideas. It was in recognition of this fact that the Board of Directors unanimously approved the increase in dues.

The impact of the dues increase on IEEE members who do not live in United States or Canada was of great concern to the Board of Directors. To develop means for increasing the service the Institute renders to the members residing outside of the United States and Canada, the Board set up a special *ad hoc* committee under the chairmanship of Dr. George Sinclair of the University of Toronto to study this question and to report to the Board at its next quarterly meeting. This committee's function is to develop recommendations as to how to minimize the adverse effect of the dues increase in those countries where foreign exchange problems are most severe and where the new dues structure may make it difficult to attract new members, or to retain present ones.

The Board of Directors welcomes comments by the members on suggested areas for improved membership services. It is planned to devote a special section of the next few issues of SPECTRUM to the discussion of various possibilities. The particular recommendations that relate to members outside the United States and Canada will be directed to Dr. Sinclair's *ad hoc* committee.

Your contributions are solicited.

F. Karl Willenbrock

An important message from your President

IEEE dues to be increased in 1967

As announced in President Shepherd's letter to the membership, the Institute dues will be raised to $25, effective next January. Herewith is the factual background

William G. Shepherd *President IEEE*

My letter of April 8 to IEEE members announced that the IEEE dues would be raised next year and that the detailed background would be published in this issue of IEEE SPECTRUM. The letter contained summary statements of the factors that underlay the decision to raise the dues. These are repeated here, followed by explanatory comments and supporting data.

Introduction

• *The Institute exists primarily as a mutual educational society to provide the most important generally available vehicle for continuing education for the practicing electrical engineer.*

• *To accomplish this purpose, adequate support for the Institute publications and technical activities is essential.*

• *Membership support of the Institute should be sufficient to avoid major fluctuations in program arising from changes in income beyond our control.*

These statements reflect the basic purposes of IEEE and emphasize that the services rendered to its members, to the profession, and to the public are dependent on the resources provided not only by its members but by such nonmember sources as advertising. Since advertising revenue has been subject to unforeseen major fluctuations, the IEEE Board of Directors has been able to maintain the basic IEEE activities only by drawing on reserve funds. The point has been reached where it is no longer prudent to follow this course.

Dependence on non-dues income

• *The present overdependence on nonmembership income has forced curtailment of previous programs and services at a time when these services need to be maintained and expanded.*

The charts in Fig. 1 indicate the distribution of income and expense for 1966. These show that the 1966 income from sources other than dues is 63 percent of the total. Figure 2 shows the trend of income and expense since 1962, compared with the income from dues and entrance fees. The average dues income has remained steady at about $12.50 per member per year (this is the average of $15 for members other than students and $5 for students). The expenses have averaged about $37 per member, and these have exceeded the income from all sources by an amount exceeding $1 per member per year, averaged over five years.

Figure 3 shows the major categories of income and expense since 1962. The reduction in advertising revenue from the premerger level of 1962 is evident. Table I shows the detailed figures. As stated in my letter, this reduction in advertising revenue is likely to be permanent. We expect growth in this area, but from a reduced base.

Of major significance is the fact shown in Table I that if advertising income had remained steady at the 1962

I. Income from advertising

Year	Advertising Income	Reduction from Premerger (1962) Income
1962 (IRE and AIEE)	$1 811 020	$........
1963 (IEEE)	1 397 263	413 757
1964 (IEEE)	1 392 744	418 276
1965 (IEEE)	997 188	813 832
1966 (est.) (IEEE)	1 100 000	711 020
	Total reduction accumulated	$2 356 885

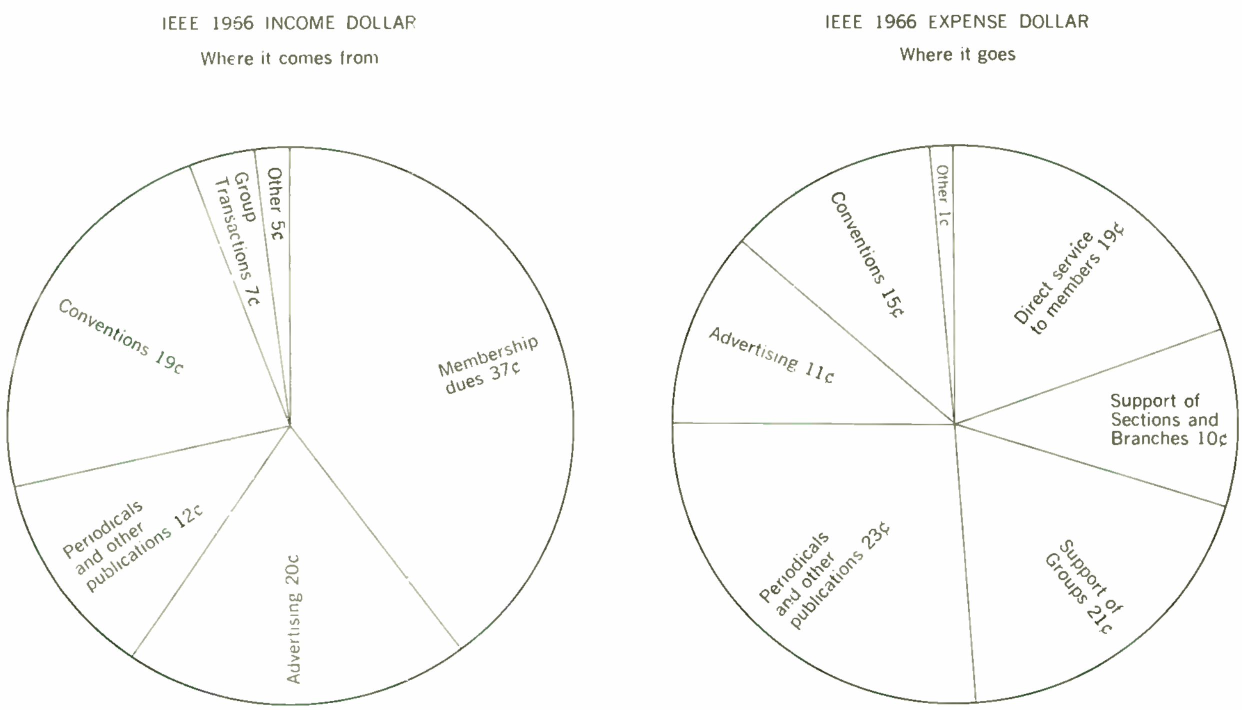


Fig. 1. The IEEE 1966 dollar.

Fig. 2. Overall income and expense.

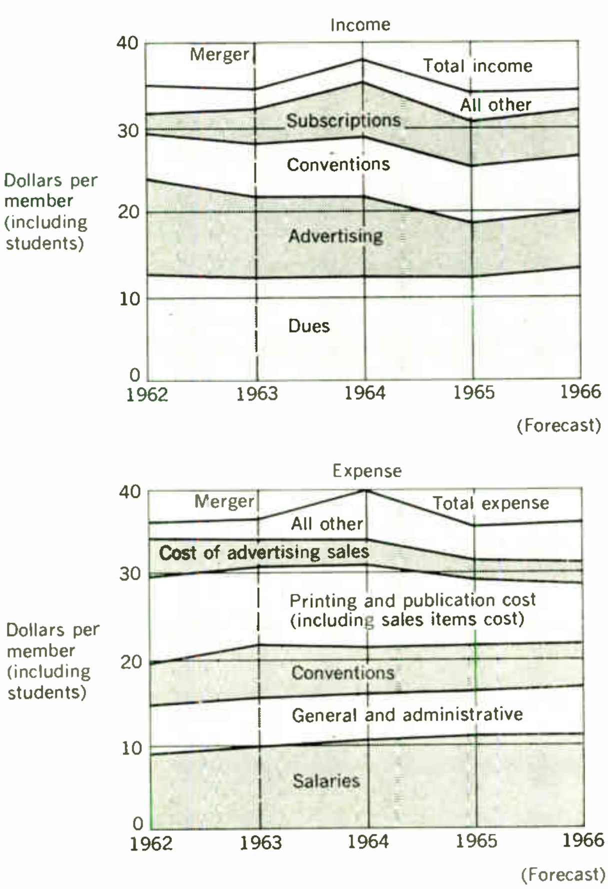
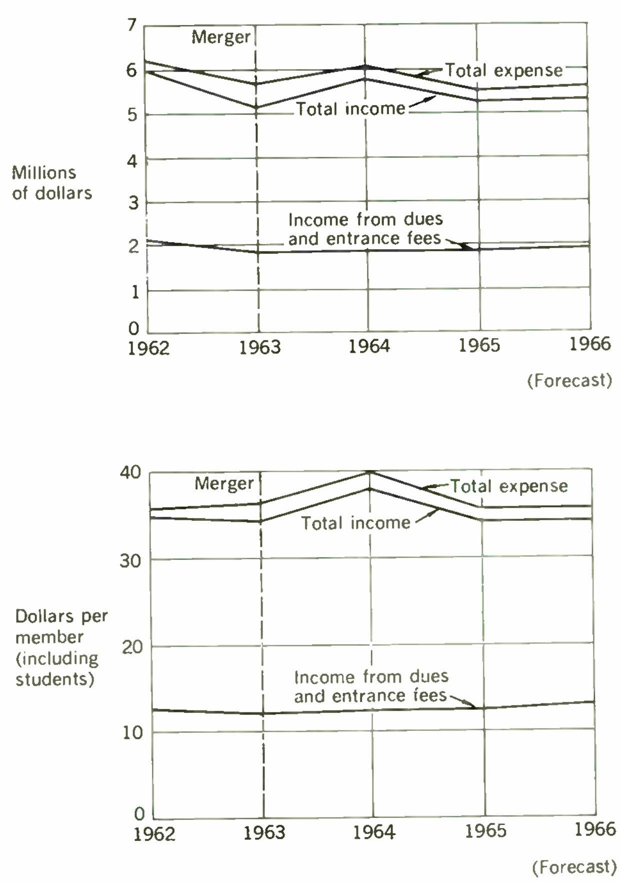


Fig. 3. Distribution of income and expense.



II. Dues of professional societies

Society	Dues	Effective Date
ASME (mechanical engineering)	$30	1963
SAE (automotive engineering)	30	1960
ASCE (civil engineering)	25	1954
AIChE (chemical engineering)	25	1963
AIME (mining, metallurgical, and petroleum)	20	1950
AIAA (aeronautics and astronautics)	20	1963
APS (American Physical Society)	20	1963
ACS (American Chemical Society)	16	1961
IEEE	15	1963
AIEE	20	1956–1962
IRE	15	1948–1962
Medicine (American Medical Association)	45	1963
Law (American Bar Association)	30	1963

Note: Some of these societies admit juniors and associates at lower dues than those shown. In those cases, the dues listed are those paid by members at approximately age 35 and/or those in the professional grades.

level (as was assumed when the merger plans and projections were announced in 1962), income from this source over the four years of the merger to date would have been $2.36 million greater than has actually been realized. *This is the principal factor underlying the decision to increase the dues.*

The sharp reduction in advertising income is not unique to IEEE. It has been experienced by nearly every publishing organization in the electronics and communications field in the United States. It has been caused, first, by a substantial increase in the number of publications sharing the advertising expenditures. There were three such publications in 1945, six in 1955, 38 in 1965. Second, a change in policy in the administration of U.S. government contracts has disallowed nearly all forms of advertising expense, and this has caused many advertisers to restrict their budgets.

Support from dues

• *The present dues support only 35 percent of the operating expenses of the Institute.*

As shown in Fig. 1, the dues income in 1966 amounts to 37 percent of the income, which is equivalent (since a deficit of income is projected, Table IV) to 35 percent of the expenses for the year. Figures 2 and 3 show that this situation has been generally true since 1962, with an even lower percentage in 1964. This fact explains the high "leverage" that fluctuations in other income exert on Institute operations.

• *The Institute presently has the lowest dues of the engineering societies.*

Table II lists the dues of the major engineering societies in the U.S., as well as those in physics, chemistry, medicine, and law. The majority of these have raised their dues since the last increases of IEEE's parent societies (1948 for IRE, 1956 for AIEE). Adjustments in the dues of other societies have occurred for reasons similar to those affecting IEEE.

Most important is the fact that professional societies, no less than other activities, have been affected by steady inflation of costs. Since most expenditures of the IEEE are made in the U.S., it is appropriate to use U.S. statistics. The Bureau of Labor Statistics' figures, for the "services" category in which IEEE's expenses largely occur, show an inflation of three percent per year compounded over the postwar years to the present. The

III. Income from dues and entrance fees

Year	Income from Dues and Fees	Reduction from Premerger (1962) Income
1962 (AIEE/IRE)	$2 139 630	$.......
1963 (IEEE)	1 868 407	271 223
1964 (IEEE)	1 912 743	226 887
1965 (IEEE)	1 902 669	236 961
1966 (IEEE)	1 964 000 (est.)	175 630
	Total reduction accumulated	$910 701

total effect has been a 74 percent increase in costs, due to inflation, since IRE's dues were increased in 1948, and a 33 percent increase since the AIEE dues were raised in 1956. The added income derived from those increases has, in fact, been offset by inflation. Moreover, as noted below, the AIEE dues were reduced at the time of the merger. Only the growth of the electrical and electronics industries since World War II, with a corresponding increase in IEEE membership, has permitted the Institute to maintain its services without a dues increase.

The end of the growth curve was evident prior to the merger, as indicated by the fact that both AIEE and IRE incurred deficits in 1962 totaling $227 150 (see Table IV). It is reasonable to suppose, therefore, that if the merger had not occurred our parent societies would have faced, individually, the situation that IEEE faces today.

• *The dues of approximately 50 000 members were reduced at the time of the merger. The net reduction in income over four years has amounted to over $900 000.*

When the merger was planned, the dues were set at $15, the lower of the dues of the parent societies. As a result, approximately 40 000 members (in grades other than Student) have paid from $5 to $10 less dues each year. In addition, there were nearly 10 000 members (other than Student members) who had been members of both AIEE and IRE. These have been relieved of paying the second dues, which ranged from $15 to $25 depending on the grade and length of membership.

After the merger, the provision for lower dues during the first three years of membership (which had been in effect in both parent societies and in IEEE until 1964) was discontinued. This, as well as reductions in arrears,

IV. Income and expenses—1962–1966

Year	Income	Expense	Deficit
1962—AIEE	$1 701 514	$1 711 937	$ 10 423
1962—IRE	4 285 720	4 502 447	216 727
1962—Total	$5 987 234	$6 214 384	$ 227 150
1963—IEEE	5 284 484	5 708 593	424 109
1964—IEEE	5 824 398	6 138 698	314 300
1965—IEEE	5 314 019	5 557 953	243 934
1966—IEEE (est.)	5 377 000	5 577 000	200 000
			$1 409 493 Total
			$1 182 343 (IEEE only)

V. Reserve funds of the Institute

Year Ending December 31	Operating Fund Balance as of December 31 (Book Value)	Deficit Incurred During Year
1961	$4 026 951	
1962	3 799 801	$ 227 150
1963	3 375 692	424 109
1964	3 061 392	314 300
1965	2 817 458	243 934
1966 (est.)	2 617 458	200 000
Net change (1961–66)	$1 409 493	Accumulated deficit $1 409 493

has provided additional dues income to offset the reductions just cited. The net effect, detailed in Table III, has been a reduction of dues income of $910 701, accumulated over the four years of the merger to date. This reduction was projected in the merger plans, and the figures actually realized closely approximate the projection.

Reduction of reserve funds

• *The reserve funds of the Institute have been drawn upon to meet an accumulated deficit of over $1.2 million over the four-year period.*

Table IV shows the income, expense, and deficits that accrued to AIEE and IRE in 1962 and to IEEE since 1963.

The large deficit (over $400 000) due to merger starting costs in 1963 was anticipated in the merger planning, as were the reduced deficits in the following years. However, the unforeseen reduction in advertising income has prevented achieving the balanced budget that was forecast would occur in the third year. It is now clear that the budgeted deficit of $200 000 in 1966 will continue in future years, and will, in fact, be increased by the continuing effects of inflation.

The effect of the accumulated deficit of $1.4 million since 1961 on the Institute's reserve funds is shown in Table V.

The value of the reserves is reported at cost, in accordance with the established accounting practice. The reserves have been reduced from about $4.0 million at the close of 1961 to an estimated $2.6 million by the end of 1966.

The value of the Institute's securities has benefited from appreciation and reinvestment by about $0.8 million during this period. The book value of the 79th Street Headquarters property was realized when it was sold in 1964, and the proceeds of that sale have been converted to holdings in the IEEE portfolio of securities.

Reduction in expenses

• *This deficit has been accumulated in spite of operating economies of $1 million during that period.*

The record during the four years of the merger, shown in Tables I, III, IV, and V, can be summarized as follows: Compared with the premerger (1962) levels, the accumulated reductions in income have been $2.36 million in advertising and $0.91 million in dues, a total of $3.27 million. Offsetting reduction in expense, such as sales expense, printing and mailing costs, associated with the reduction in advertising, have amounted to approximately $1.05 million, leaving a net reduction in income of $2.22 million. The actual deficit during this period (Table IV) has been $1.18 million.

The difference between the net reduction in income and the actual deficit is $1.04 million. This is a direct measure of the operating economies that have occurred during the merger years to date as measured against the immediately preceding levels of activity.

The effect of budget control is shown in the expense column of Table IV. Except in 1964 and 1966, the operating expense each year has been lower than in the preceding year. The projected expense for 1966 (including inflation costs) is approximately the same as that actually incurred in 1965. In 1964 there were three major nonrecurring expense items, totaling $0.54 million, associated with the transfer and consolidation of all Headquarters staff activities at the United Engineering Center, and the installation and record conversion costs of the IEEE Computer Center. This result has been achieved despite the effects of inflation, which have increased the cost of nearly all purchased items, as well as salaries, by three percent per year.

Need for additional income

• *In spite of these economies, continuing deficits are forecast at present levels of income and service rendered.*

IEEE dues—past and present

Until January 1, 1967, the annual dues for all grades other than Student shall be fifteen dollars ($15.00) per year. The dues for Student Grade shall be five dollars ($5.00) per year. Student members transferring to a higher grade, without interruption of membership, shall pay dues of ten dollars ($10.00) for the first year of membership in the higher grade.

Effective January 1, 1967, the annual dues for all grades other than Student shall be twenty-five dollars ($25.00) per year. The dues for Student Grade shall be five dollars ($5.00) per year. Student members transferring to a higher grade, without interruption of membership, shall pay dues of fifteen dollars ($15.00) for the first year of membership in the higher grade.

(*IEEE Bylaw 109.2, as passed unanimously by the Board of Directors, March 25, 1966.*)

The consolidation of staff, economies from combined purchases and competitive bidding, and other economies planned for the merged operations of IEEE are now in full effect, and there is little prospect of further reductions. On the contrary, since continued inflation of costs is in prospect, an increase in the deficit must be anticipated to cover inflation, unless the present level of activity is further curtailed.

The reduction in reserves that has occurred since 1961 requires rebuilding the reserve fund against such future contingencies as taxation. The "prudent rule" for organizations such as IEEE is that they should maintain a reserve fund approximately equal to one year's operating expense. At the end of 1965, the Operating Fund Balance of $2.8 million (Table V) represented only 51 percent of the operating expenses in that year.

Increased IEEE activities

• *Increased support of IEEE publications, Groups, Sections, and Chapters has been requested by the membership and recommended to the Board by the officers responsible for these activities. Allocation of new resources to the expansion of these activities will be made on the advice of the responsible boards and committees that have been asked to make recommendations to the Board.*

In 1967 and thereafter, funds will be available for new and expanded services, over and above the funds necessary to absorb the operating deficit, provide for inflation, and replace reserves.

Those Standing Committees whose activities represent substantial outlays—particularly the Publications Board, the Technical Activities Board, the Sections Committee, and the Student Branches Committee—have been asked to present to the August meeting of the Board of Directors their recommendations regarding new or expanded activities. They have also been requested to analyze the income and expense associated with each proposal. The Executive Committee will then develop budget provisions for such of these activities as can be accommodated in the 1967 Operating Budget. These budget items will be reviewed by the Board at its November meeting.

Since it is anticipated that the recommendations will exceed the funds available, each Standing Committee has been asked to establish priorities and sliding scales for the proposed new expenditures. In this way, a balanced program of new and enlarged services will be established and announced to the membership through the pages of SPECTRUM.

To the proposals already under discussion will be added those submitted by many hundreds of members in answer to my letter of April 8. Among those now under consideration are the following:

Increased and improved editorial content of IEEE SPECTRUM, PROCEEDINGS, and STUDENT JOURNAL.

Expanded Special Issues of PROCEEDINGS.

Development of an IEEE program in automatic information handling.

Additional support for Group publications.

More frequent publication of the *IEEE Directory*, at a lower price.

Increased member rebates for Sections.

Rebate for Section Technical Conferences (already in effect).

Establishment of Section rebate for Group Chapter memberships (already planned, but amount not yet decided).

Establishment of a Continuing Education Program, Lecture Series, Lecture Bureau, in support of Sections.

Dues to include a partial subsidy of a Group membership fee or a subscription to PROCEEDINGS.

IEEE and the individual member

Because the profession of electrical and electronics engineering is in itself large and complex, the Institute of necessity must serve a very large membership and its sheer size tends to make its operations remote and impersonal. The problems of maintaining the vitality of a local organization in which the members are immediately and personally involved are much more readily understood. In a society as large as the IEEE, it is necessary that the members delegate to an elected Board of Directors the responsibility for balancing the support of expressed needs against available resources. The diverse and complex interests of our far-flung membership demand intelligent and concerned compromise.

A society like ours can serve as an effective instrument only if its service is sufficient as judged by the community at large. Resources too limited to support presently expected service and to allow the flexibility necessary for adapting to changing needs will, in the long run, restrict IEEE's ability to perform its vital role. Thus, the Board has had the important responsibility of judging when a critical point has been reached in this relationship.

The Directors have given the most careful consideration to the decision to increase the dues. They have done so in the belief that the Institute has a vital role to play in maintaining the competence of each member to serve society. The course of action on which it has embarked will provide, I am firmly convinced, a sound fiscal basis for the growth of the Institute and will enable it to fulfill its obligations to the membership. It is the hope of the Board that the members will feel that it has acted responsibly in their behalf.

Authors

Articulated science teaching and balanced emphasis (page 49)

William Shockley (F) received the Ph.D. degree in physics from the Massachusetts Institute of Technology and, in 1936, joined Bell Telephone Laboratories. His work included vacuum tube and electron multiplier design, studies of various physical phenomena, alloys, radar development, and solid-state physics. During World War II, he served as director of research for the U.S. Navy AntiSubmarine Warfare Operations Research Group and as expert consultant of the Office of the Secretary of War. He returned to Bell Labs to supervise a research group on solid-state physics. One result of this work was the invention of the transistor in 1948, for which Dr. Shockley, John Bardeen, and W. H. Brattain shared the Nobel Prize in physics. In 1953 he was appointed director of transistor physics and, the following year, he became visiting professor of physics at the California Institute of Technology. Later he worked in private industry in the field of semiconductor devices. At present, he is executive consultant on applied research and development of electronic components at Bell Labs and Alex M. Poniatoff professor of engineering sciences in the Department of Electrical Engineering at Stanford University.

Worldwide color television standards (page 59)

Francis Charles McLean (SM) currently holds the position of deputy director of engineering with the British Broadcasting Corporation, London, England. He was graduated from the University of Birmingham, England, in 1925. Following graduation, he became associated with the Radio Department of Standard Telephones. He remained with Standard Telephones, working in various European countries, until 1937, when he joined the Transmitter Department of the British Broadcasting Corporation in London. In 1963, he became the deputy director of engineering; in this position he has charge of all engineering activities in television, radio, and short-wave broadcasting.

Mr. McLean has represented the British Broadcasting Corporation at many international conferences. At the present time, he is serving as chairman of the European Broadcasting Union Subgroup on Color Television Receivers. In addition, he is a member of the Council of the Institution of Electrical Engineers (Great Britain). He has presented a number of technical papers on the subjects of radio and broadcasting to the Institution of Electrical Engineers and has been awarded three IEE Premiums.

Compatible color signal transmission and European modulation methods (page 60)

Richard Theile received the doctoral degree in physics from the University of Marburg, Lahn, Germany, in 1938. He then joined the Research Laboratories of the Telefunken GmbH, Berlin, where his work was concerned with electronics and electrooptical developments for television. After the war, he became a lecturer at the University of Marburg. He subsequently spent three years doing research at the television laboratories of Pye Ltd., in Cambridge, England. Since 1953 he has been connected with research and development for the West German broadcasting companies as organized within the ARD and the ZDF. He is now director of the Institut für Rundfunktechnik in Munich and honorary professor at the Technische Hochschule, Munich.

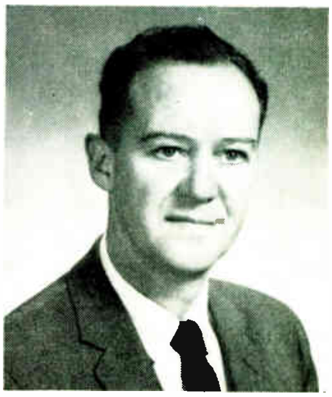
Prof. Dr. Theile is a member of a number of scientific and technical societies in Germany and other countries. He serves on the executive board of the NTG, FTG, and DKG. At the present time, he is chairman of the Ad Hoc Group on Color Television of the European Broadcasting Union. He received the Louis Sterling Premium of the British Institute of Radio Engineers in 1952 and the EMI Premium of the Television Society, London, in 1958.

Hydrocarbon–air fuel cell systems (page 69)

C. Gordon Peattie (M) is director of the Energy Research Laboratory, Texas Instruments Incorporated, Dallas, Tex. He received the A.B. degree from New York State College of Teachers in 1940, the M.S. degree from the University of Oklahoma in 1949, and the Ph.D. degree from the Massachusetts Institute of Technology in 1952. From 1940 to 1947 he was with the Texas Company, Beacon, N.Y., as leader of a group working on the analysis of petroleum products. Later he held the positions of teaching assistant and university fellow at the University of Oklahoma, teaching assistant and DIC fellow at M.I.T., and research associate at Harvard Medical School. From 1953 to 1956 he led a mass spectrometer group at Shell Development Company, Houston, Tex., and later spent six months with General Electrodynamics Corporation, Dallas, directing a research program on photoconductive television camera pickup tubes. He joined Texas Instruments in 1956 and, since then, has supervised research in electrochemistry. Prior to his present position, he served as branch manager.

The contourograph (page 77)

George Newton Webb (M) received the B.E. degree from the University of Toledo, Ohio, in 1947 and joined the faculty as an instructor in electrical engineering. In 1949 he became an instructor in electrical engineering and research engineer at North Carolina State College of Agriculture and Engineering, earning the M.S. degree in 1952. He then joined Johns Hopkins University as electronics engineer, Biophysical Division, in the School of Medicine's Department of Medicine. He became an assistant in medicine in 1957, an instructor in medicine in 1961, and assistant professor in biomedical engineering in 1964. Mr. Webb has served as a part-time consultant in biomedical engineering to Smith Kline Instruments Company and, in biophysics, to the Veterans Administration's Mt. Alto Hospital in Washington, D.C., and to the Psychophysiology Laboratory of the VA hospital in Perry Point, Md. He is a member of the International Federation for Medical Electronics and Biological Engineering, American Association for the Advancement of Science, Eta Kappa Nu, and Sigma Xi.



Richard E. Rogers (M) is an instructor in the recently formed Department of Medical Instrumentation of the University of Miami School of Medicine and, in this position, is responsible for the operation of the electronics facilities and for the initiation of a teaching program in medical electronics. He received the B.S.E.E. degree from Johns Hopkins University and, from 1957 to 1965, was associated with the Departments of Medicine and Biomedical Engineering at the Johns Hopkins School of Medicine. His work there was concerned with the area of physiological data acquisition and display, with major emphasis on the analysis of electrocardiograms and heart sounds. Investigative projects with which he was associated include the analysis of heart sounds, utilizing frequency–time–amplitude displays, acquisition and analysis of intracardiac sounds, and development of adult-patient and premature-infant monitoring systems. In addition, he devoted some time to work on the development of magnetic-tape systems for physiological data and analysis

Progress in ionized-argon lasers (page 88)

Roy A. Paananen (A) received the B.S. degree in electrical engineering from Michigan State University in 1949 and the E.E. degree from the Massachusetts Institute of Technology in 1953. During his last year at M.I.T. he worked at the Lincoln Laboratory on government classified research and development. He joined the Microwave and Power Tube Division of Raytheon Company in 1953 and was part of the group responsible for the original development of "M"-type backward-wave oscillators in this country. In 1960, he became associated with Raytheon's Research Division, working on projects directed toward advancing the microwave tube art, such as the application of ferrite tuning to medium-power klystron oscillators and the development of an understanding of the properties of dielectric-loaded interdigital delay lines. Since 1961 he has been concerned with the field of gas lasers and has published material on gas laser design, gas laser amplifiers, Zeeman effect in gas lasers, and discovery of laser action in chlorine gas. His most recent efforts have been devoted exclusively to the exploitation of ionized-argon laser potentialities.

Articulated science teaching and balanced emphasis

The experiences that stay in our memory are those that are dramatic, or are exaggerated, or are unusual in some other way. Teaching experiments have shown that this principle can provide us with a useful scientific thinking tool

William Shockley *Stanford University and Bell Telephone Laboratories, Inc.*

The set of scientific thinking tools described is based on dramatized, exaggerated patterns of logical structures. What these patterns have in common is an overall invariance relationship that makes something true even though the details may differ in different situations—and a knowledge of that relationship leads to simple solutions of apparently complex problems. Articulated science teaching is defined as a method of getting the invariance concept across to the student by using selected puzzles and games as exercises in logical thinking. The end result should be to improve rational reasoning ability both in science and in everyday life.

Prologue

"One objectionable feature of the instruction of most colleges is the step by step method. One subject is taken up, by application of sufficient time and energy pushed through, and then after passing an examination dropped to take up another subject. It is true that by steady application to one subject a great deal can be learned and splendid results derived in the examination papers, but all that is learned in this manner is just as rapidly forgotten. To understand a matter thoroughly, so as really to have a lasting benefit from it, and not merely make a good showing in examination papers, requires several years' familiarity. *Therefore, any subject that is not kept up during the whole college course might just as well be dropped altogether and the time spent therein saved*. . ."

"What then, should an ideal electrical engineering course contain?

"Of mathematics, plane and some solid geometry, arithmetic, and a good knowledge of algebra. Plane trigonometry—no spherical trigonometry required—and a thorough understanding of analytical geometry and of calculus, but no memorizing of integral formulas, etc.; they can always be looked up in a book when needed.

"A thorough knowledge of general physics, *especially of the law of conservation of energy*, which I am sorry to say is not yet an integral part of our thinking, and of chemistry, especially theoretical chemistry and the chemical laboratory, in the latter freely using tables, etc.

"The electrical laboratory work should be taken up right at the beginning. Even before taking up the theory of apparatus, the theoretical investigation of alternating currents, self-induction, transformers, etc., the student should have met these phenomena and handled the apparatus in the laboratory. Only after seeing the effects of self-induction, for instance, on the alternating current circuit in the laboratory, will the theory of self-induction, have any meaning or *make an impression on the student*. What the student has seen practically in the laboratory he will then take up in the theoretical course, and understand it; he will learn to calculate and control, and afterwards, going back once more to the laboratory, to apply it. It is here that the average college graduate is inferior to the practical man, and in spite of his handicap regarding theoretical knowledge, the latter frequently pushes ahead of the college graduate because of his superior understanding of the phenomena, based *on his familiarity with them*. . ."

Charles Proteus Steinmetz expressed the views quoted above (except for the emphasis added by the present author) in his 1902 presidential address to the AIEE. The reader will find much in common between Steinmetz' views, especially the emphasized phrases, and the article that follows. Any common conclusions are quite independent; this article was completed before the writer had read the Steinmetz address.

Principles of invariance

Around 1925 I had one of my earliest and most memorable experiences with the power of scientific thinking tools. I was marked wrong on a physics problem that I thought I had solved correctly. The problem consisted of calculating the propelling force required to keep a rowboat in motion. This force was produced by a man pulling with a 25-pound force on each oar handle. The distance from oar handle to oarlock was 2 feet and that from oarlock to blade was 5 feet.

The correct way to do the problem, my teacher insisted, is as shown in Fig. 1. First, regard the oar blade as a fulcrum. The lever arm for the man's 25-pound pull is then 7 feet. Next, calculate the force F_o exerted on the oar by the oarlock; the law of levers gives $25 \times 7 = F_o \times 5$ so that $F_o = 35$. This would give 70 pounds pushing the boat forward, but the man pushes back on his seat with 50 pounds. Therefore, the net force on the boat is $70 - 50 = 20$ pounds.

I calculated F_B, the push of the water on the oar blade, using the oarlock as fulcrum: $25 \times 2 = F_B \times 5$, so that $F_B = 10$. However, the push on the blades is the force pushing the boat–oars–man system forward. This force is thus 20 pounds—the same answer.

This solution was not acceptable to the teacher; it did not actually calculate forces on the boat. We argued. I tried to bring in an outside authority, a university professor whom I knew, but he was a professor of education and did not wish to become involved. I remained convinced that my method was sounder and simpler than the teacher's.

This is but one of a number of similar examples I remember from high school. They all had a common feature—some overall principle permitted an answer to be reached that was *always true no matter* what the details.

One of the best known of these problems was that involving the bee and the motorcyclists: two motorcyclists start from a common point and proceed at 30 mi/h in opposite directions on a circular racetrack. Simultaneously, a bee starts from the same point and flies at 60 mi/h in precalculated straight-line chords (not pursuit curves), intercepting the cyclists alternately until it is crushed between their noses when they meet. The circular racetrack is approximately 10 miles in diameter and the circumference is exactly 30 miles. What is the total length of the bee's path? That is, what is the sum of all the straight-line chords? (You may, of course, consider the idealized limiting case of point travelers and infinite accelerations.) As I recall the situation, I solved this problem on my own, thereby impressing the college student who had posed it to me. This was a dramatic experience. He was a big man in my eyes, farther along in college than I was in high school. I was impressed with my own accomplishment.

A similar problem that I was given at about the same time involved a string around the earth's equator. Suppose this string makes a snug fit. Cut it and add 3 feet, then form it into a new perfect circle at a height h above the ground. Can you insert a pencil under it? I began by converting 8000 miles to feet, multiplying by 2π, then adding 3 feet, and before I finished (I probably bogged down) someone told me how to use the algebra I knew.

The rowboat, the motorcyclists and the bee, and the string around the equator are what I now call dramatized, exaggerated patterns of logical structures. What do they have in common? In each case, there is an overall invariance relationship that makes something always true, no matter what the details may be. These invariance relationships become more apparent when they are expressed in parallel forms as follows:

In the rowboat situation, it is *always true* that action (the back push of the boat system exerted on the water by the oar blade) equals reaction (the forward push of the water on the boat system) *no matter* what internal forces exist between man and seat, man and oar handles, oars and oarlocks, etc. (*but not if* turbulent drag is included as part of the force on the boat; this obviously was not intended in the problem).

In the racetrack problem, for every unit of distance traveled by a motorcyclist, the bee travels two units. This invariance relationship shows that it is *always true* that the bee flies 30 miles while each cyclist does 15 miles *no matter* whether the bee flies chords, pursuit curves, or hexagons in the middle of the racetrack, *but not if* the bee's time does not match that of the cyclists or if the acceleration of the bee must be included.

The parallel sets of emphasized phrases exhibit the structure of the qualified law form discussed more fully later in this article. They emphasize that having a principle well under control means knowing not only *what* relationships it asserts, but also *when* it does apply in terms of specific examples and, furthermore, *when not*, again in terms of specific examples.

For the string around the earth, the key relationship is "circumference equals 2π times radius." Consequently, it is *always true* that each 2π feet of increased circumference gives 1 foot of height *no matter* how big a circle we choose, *but not if* our precision of discussion should include curved space effects due to the earth's gravitational field.

The invariance principles in these examples lead to simple solutions of apparently complex problems. The power of conservation laws and invariance principles to dispose of details and go directly to the answer has always seemed to me to be among the more appealing esthetic aspects of theoretical sciences. Can this combination of elegance and utility be taught? Can these methods be made more useful to science and engineering students? Can a sympathetic appreciation of them be imparted to nonscience majors?

Transfer and the articulated science approach

About seven years ago I tried to help a science student who was studying PSSC physics. The problem in question involved index of refraction. I shall give the discussion in brief; understanding the details is not essential to following my general line of thought. In trying to impart understanding, I pointed out that since $n_1 \sin \theta_1 = n_2 \sin \theta_2$, the quantity $n \sin \theta$ is conserved as a light ray crosses a plane of discontinuity of refractive index n at angle θ away from perpendicular incidence on the plane. The same will be true for a set of parallel plane slabs. Hence a light ray entering from air into a stack of parallel slabs will emerge at the far side into air on a path parallel to its entrance because $n \sin \theta$ is the same for incident and emergent rays, and so is n. Therefore, so is θ, and equal θ's mean parallel rays.

In summing up our original extended discussion I said, "It is like a conservation-of-energy principle." This went over like a lead balloon.

Articulated science teaching, as defined in the remainder of this section, is my attack on getting the invariance or conservation law concept across.

Articulated science has a logical parallel to a device in electrical engineering. This device, named the *dissected amplifier* by Mason and the writer,[1] shows that negative resistance plus nonreciprocal circuit elements can be combined to construct amplifiers. The dissected amplifier is composed of two spatially separated parts: ac power generation and one-wayness. In the amplifier these parts are articulated in combinations that have, for example, the four-terminal characteristics of an ideal pentode.

Somewhat similarly, articulated science teaching dis-

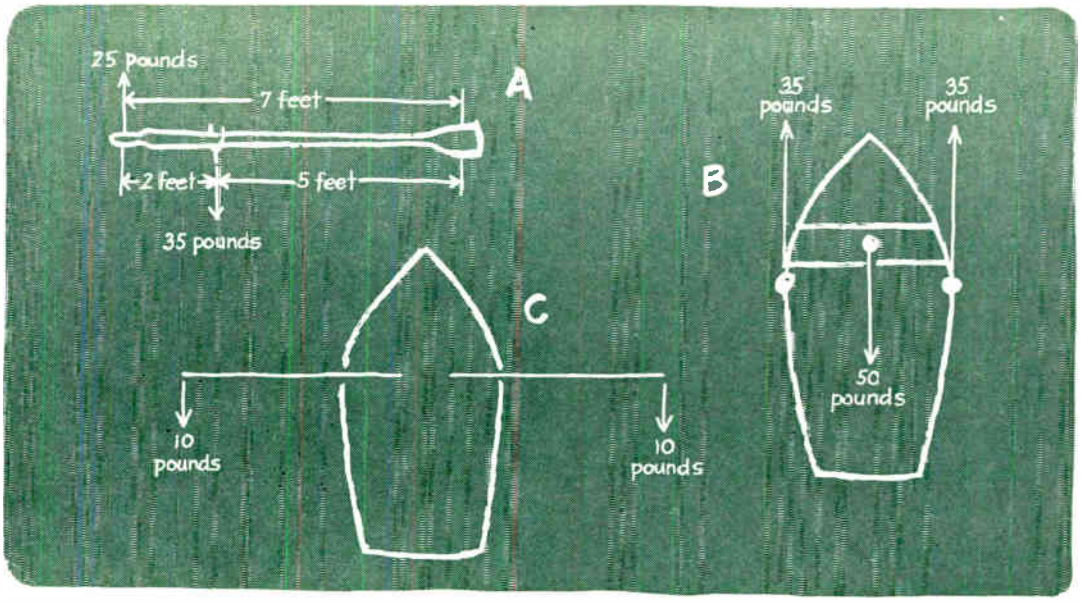


Fig. 1. The rowboat problem. A—Moments on oar with blade as fulcrum. B—Net forces on boat. C—Forces of oar blades on water.

sects the pedagogical problem into several smaller steps. Logical structures are introduced that have conservation principles.[2] These deal with very simple subject matter, such as counting piles of matches or paper clips. Then concepts of quantitative attributes of nature are developed, such as weight, length, or moment. The student can transfer the concepts of invariance from the subject matter of games and puzzles to the scientific subject matter. The recognition of a significant quantitative attribute such as a moment becomes more immediately meaningful; thus the logical structure concepts are articulated with the concepts of quantitative physical attributes of nature. An additional aim is to provide an experience of active participation in developing scientific concepts. The new teaching approach uses the tried-and-true technique of dividing and conquering. Articulated science teaching dissects the pedagogical problem into two smaller parts.

Provoking and observing a transfer of the concept of conservation in games and puzzles to the concept of conservation of energy gave me one of my most exciting teaching experiences. In the summer of 1963 I had taken four or five volunteer eleventh grade students through the logical structures of three or four selected games and puzzles. These (one of which I shall describe more fully presently) were the tennis tournament problem, the *n*-acquaintance problem, and the checkerboard problem. After searching out and stating in qualified law form the principles of these problems, I had led the students to establish experimentally the law of levers.

With these experiences as a basis, I then posed some hypothetical problems in which a set of weights was to be rearranged on a set of shelves. The conceptual operations to be done on paper were restricted to those that could be performed using balanced levers. In other words, some weights were lifted at the expense of equal losses of potential energy of other weights. I chose the first three problems so that there was no change in total potential energy. The students soon worked out solutions. Then I gave them an impossible problem in which the potential energy increased significantly.

These students had had little, if any, previous experience with quantitative energy concepts. Potential energy had not been mentioned in my class.

After five or ten minutes the conversation went like this:

Kathy: "I don't think it can be done."
Teacher: "Why?"
Kathy: "It doesn't add up."
Teacher: "What doesn't add up?"
Kathy: "I think the weights that go down times the distance they go down must equal the weights that go up times the distance they go up."

Another student, who had not concluded anything about the problem so far, observed: "Those are foot pounds. That is a form of energy."

My conviction is that Kathy had realized that this problem had the feature, "You can't do it, no matter what." This she recognized from the game experiences as a clue that a conservation law was involved. The games had been deliberately selected so as to exhibit conservation laws. They had the "you can't ever, no matter what" feature common to efforts to make perpetual motion machines, or to produce gold by alchemy, or to beat a gain–bandwidth product. Kathy sensed this situation and found the relevant invariant attribute, namely, sum of weight times height, or at least the significance of its changes. By carrying the concept of conservation from the games to potential energy, Kathy made my teaching experiment a success. Later I learned that this process is called *transfer* in educational psychology.

My teaching experiments have been directed toward imparting dramatized, exaggerated patterns of logical structures for transfer in scientific thinking. How can this be done? Why this emphasis on exaggeration and dramatization?

People remember things that are dramatized, exaggerated, or grotesque, or are striking or unusual in some

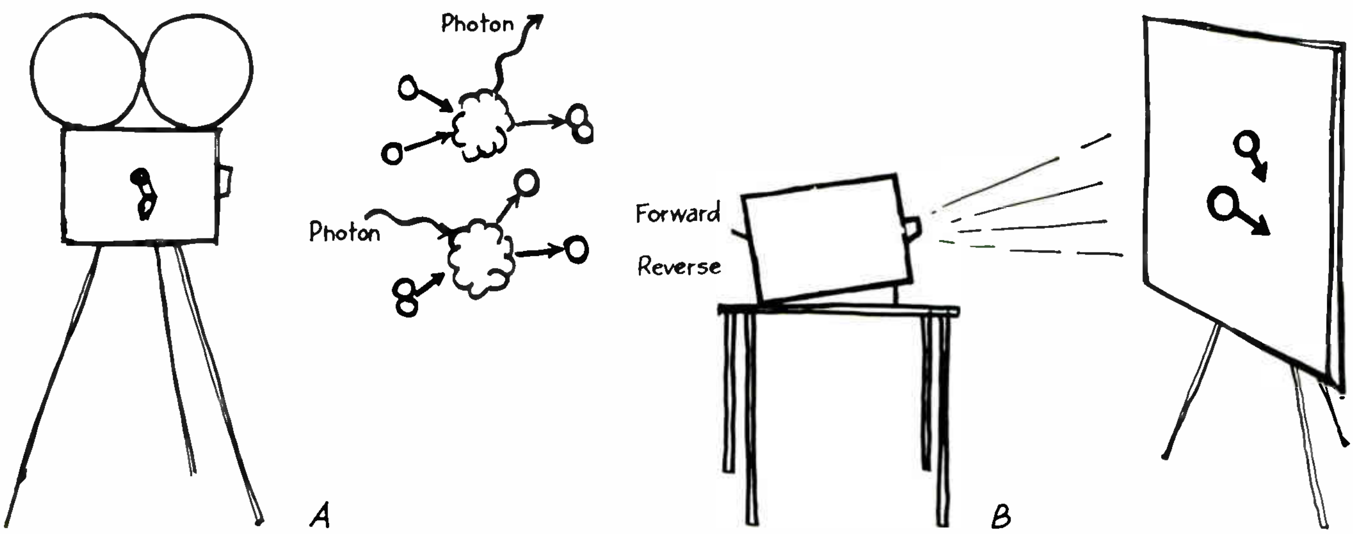


Fig. 2. The principle of detailed balance. A—The dramatized, exaggerated, and even grotesque idea that a movie can be taken of atomic processes, including the flight of photons. B—A similar idea that the movie can be projected and you can't tell by looking at it whether it is being run forward or backward.

other way. This fact is exploited in the techniques of memory stunts.

My favorite example of a dramatized, exaggerated pattern involves the principle of detailed balance. This has been one of my most useful scientific thinking tools. As I remember hearing it in a statistical mechanics course from my M.I.T. thesis professor, J. C. Slater, the basic idea is this: When a system has come to thermal equilibrium, it is going nowhere. Past and future are alike to it. On the average, for every time two atoms collide to form a molecule, a molecule will split into two atoms. Forward and reverse reactions balance in detail.

Suppose, as shown in Fig. 2, you could take a movie of this process (a grotesque, exaggerated idea indeed). If you projected this movie, you would not be able to tell whether it was running forward or backward. This invariance to reversal leads at once to equations that often are new, useful, and publishable results. Einstein's stimulated emission coefficients[3] and the Nyquist[4] derivation of Johnson noise are classic examples.

The reader prepared for more technical details may be interested in the following applications to the Shockley–Read recombination theory[5] in semiconductors: Suppose empty traps in a semiconductor capture electrons at a rate $nc_n n_t f_p$ (where n = electrons/cm^3, c_n = capture constant cm^3/s for empty trap, n_t = traps/cm^3, f_p = fraction of traps empty). Suppose also that the rate of emission of electrons out of traps is $e_n n_t f (e_n$ = emission constant for full trap, $f = 1 - f_p$ = fraction of traps occupied by electrons). Then detailed balance says if in one second in one cm^3 you observe $nc_n n_t f_p$ electron captures, you must also observe exactly the same number of emissions. Why? Because if you run the movie backward, capture looks like emission and emission like capture. If the rates aren't equal, then you can tell the difference between forward and reverse movies, and according to detailed balance you can't be at equilibrium. Hence you can conclude that

$$nc_n f = e_n(1 - f)$$

This is actually one way of deriving the Fermi–Dirac distribution, as may be seen by first solving for f:

$$f = \frac{1}{1 + (nc_n/e_n)}$$

Next, show that this is the Fermi–Dirac factor by noting that for nondegenerate electrons $n = N_c \exp (F - E_c)/kT$, where N_c is conduction band density, F is the Fermi level, E_c is conduction band edge, and kT is thermal energy. Then f is seen to be correctly the Fermi–Dirac distribution function. This is a simple example of the power of the detailed balance principle. My own work has used detailed balance numerous times.[6,7]

Three scientific thinking tools

Examples of logical patterns that have been tried in a ninth grade articulated science teaching experiment[8] are contained in two "take-away" games that resemble the famous "take-away" game of Nim. These games are played with one or more piles of objects such as paper clips or matches.

"Two pile" is played with two piles of matches with, say, ten matches in each pile to start. (This comes out even with a 20-match book of paper matches.) Two people play. For his play, a player may remove one match from either pile, or one from both; in other words, he has three possible plays. If there is only one pile left, he must take one from that pile. The winner is the one who takes the last match, leaving his opponent with no possible play.

"One to five" is played with one pile of, say, 20 to start. On his play, a player must take at least one match, and may take two, three, four, or five. Again, the winner is the one who takes the last match, leaving none for his opponent. (Alternate variations of both games make the loser the person who is forced to take the last match.)

Each game has its own principle, and knowledge of that principle enables one to win unless his opponent also knows it or happens to be lucky. The games are taught to dramatize the power of mastering a logical principle. Half a class learns one game, and half the class the other game. Students then pair off and play the two games alternately.

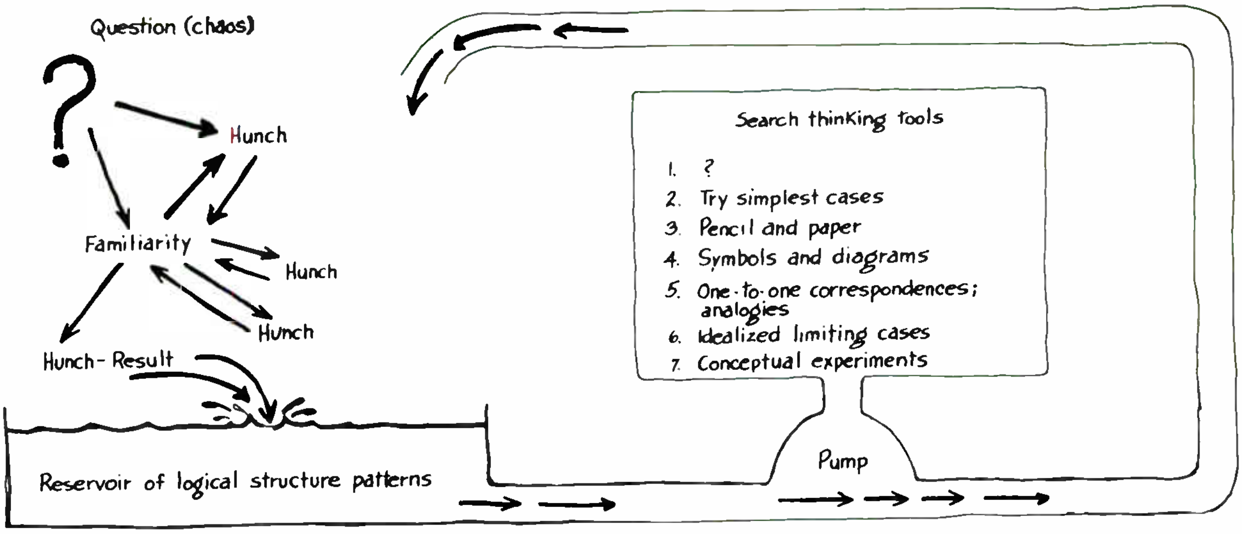


Fig. 3. The creative search pattern and seven search thinking tools.

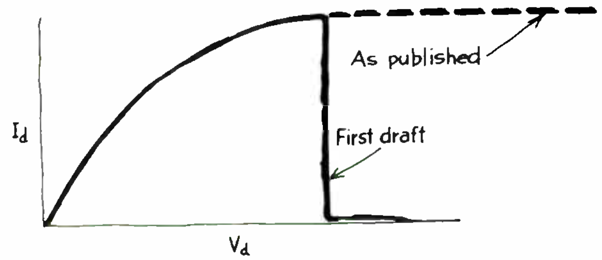


Fig. 4. A bum hunch about pinch-off in the field-effect transistor.

These games can be used as dramatized, exaggerated examples of three important scientific thinking tools: (1) CSP—the creative search pattern (including the "search thinking tools"); (2) ACOR—(key *a*ttribute)-(*co*mparison operation)-(result, an orderly *r*elationship pattern), of which postulate–proof–theorem is an example; and (3) QLF—the qualified law form for a result.

The creative search pattern[2,8] of Fig. 3 has a family relationship to similar forms in various writings on the creative process. To a greater degree than any others I know, it portrays the disorderly, frustrating aspect so often present in creative work. I can illustrate the role of the hunch of Fig. 3 by telling some stories on myself.

When I first worked out the gradual approximation for the field-effect transistor, I had a wrong but very useful hunch about pinch-off. (Pinch-off is now, I think, perfectly clear and obvious in the literature.[9,10]) My hunch led to the first plot of drain current I_d vs. drain voltage V_d, as shown in Fig. 4. I thought this characteristic would make a wonderful switch, and I circulated a draft memorandum in which I explained how it worked and prophesied its utility.

The more I thought about my idea, and became familiar with pinch-off concepts, the more unreasonable it seemed. Then the key attribute of saturation current became obvious; the influence of gate and drain voltage effects was compared by mathematical operations involving exponential attenuation of drain voltage effects. This was the needed ACOR pattern. I rewrote the memorandum and published it.[9]

The whole creative search process ends with a result. In this case the orderly functional relationship between drain current and drain and gate voltages constituted the now-classic analysis of the field-effect transistor. The CSP history involved alternating between hunches and familiarity for about six years. The hunch that a field-effect transistor really might be worth analyzing followed, as I recall, the stimulus of advancing semiconductor technology. The published output was a significant new ACOR pattern.

When I heard a few years ago that a copy of my "bum-hunch" draft memorandum still existed, I was distressed. But now I use its existence as a teaching tool. After I have told the story to a graduate seminar, I can refer to it to give perspective to comparable blind spots in the thinking of my students. I can say, "That reminds me of my first field-effect transistor manuscript." Most students can then be objective and carry out some valuable and constructive analysis of their own thought processes.

The story also illustrates another of my precepts for doing research: Don't expect to avoid mistakes; acquire familiarity and get better hunches by taking action and doing something to correct your mistakes before you publish (or spend a lot of money on development).

The validity of the question-familiarity-hunch-result aspect of research problems shown by the field-effect transistor can be easily demonstrated by using my selected puzzles and games. The tennis tournament problem is a good example. It is stated as follows: If there are 137 entrants in a tennis tournament, what is the maximum number of matches that may have to be played to decide the winner? The tournament should be scheduled so as to be as fair as possible.

A preliminary hunch, occurring before any real familiarity has developed, is often, "This is obviously a problem for geometric series." A less-frequent hunch is, "Use mathematical induction." A first step is usually to give one player a bye and play a first round of 68 matches, producing 69 survivors for 34 matches and one bye in the second round. When this is all added up, the sum is usually within a few matches of 137.

Then I say, "Use thinking tool 2." (Thinking tool 1 is left as a problem for the reader at this point.) The simplest case, the student usually says, is three players; less often does he say two and only very rarely one. Then thinking tools 3 and 4 lead to a table such as this:

P_o = number of players at start	1	2	3	4	5	...	8
M_T = matches required for tournament	0	1	2	3	4	...	7

Somewhere along the line, the hunch $M_T = P_o - 1$ is stated. I ask: "But is it *always true* that $M_T = P_o - 1$ *no matter* how the byes are scheduled? Will you bet me a dime against a dollar that I can't produce a counter-example?" (This challenge has cost me very little.)

If the decisive hunch that leads to the final ACOR structure has not occurred, I suggest: "Consider the quantity P_E, the number of players eliminated at each stage of the tournament." This quantity is the key attribute of the second scientific thinking tool, and usually recognition that the key postulate is "one match eliminates one player" follows. It is not easy to state this postulate properly; if you doubt this, try to do so before you look ahead at the QLF for Ohm's law or read the footnote.*

The logical structure of the tennis tournament problem is shown in Fig. 5. It is an example of the (key attribute)–(comparison operation)–(result) pattern. This logical structure pattern is harder to get across to students than the principle that the number of matches is one less than the number of players. But it *is* the logical structure that clearly displays the conservation or invariance principles. Nonscience college freshmen have enjoyed mastering it. (This conclusion is based on two years of experimentation in teaching mechanics at San Jose State under the direction of W. A. Gong. Professor Gong has organized a program in the six departments of the school of science to utilize on a unified basis the three scientific thinking tools of this section.) Figure 5 contains two key postulates for the problem. (I used a number of similar diagrams in *Electrons and Holes in Semiconductors*. They helped my thinking; but there has been little feedback from readers.) The first postulate is the elimination principle, and the second, in which P_R is remaining (not eliminated) players, is the conservation of people, in this case, players. Step 1 uses a boundary condition postulate so obvious that it is not shown: before the tournament starts $M = 0$ and $P_E = 0$. Step 2 is "substitute equals for equals." Step 3 uses the definition: M_T corresponds to $P_R = 1$.

The tennis tournament and other problems described have been chosen to give the "you can't change it, no matter what you do" feeling characteristic of important scientific and engineering principles. The final principle, $M_T = P_o - 1$, is often grasped before the logic is understood. That is why the student lacks the necessary conviction to risk his dime to get my dollar. The dollar approach, or its equivalent in classroom prestige, dramatizes the power of the logical structure.

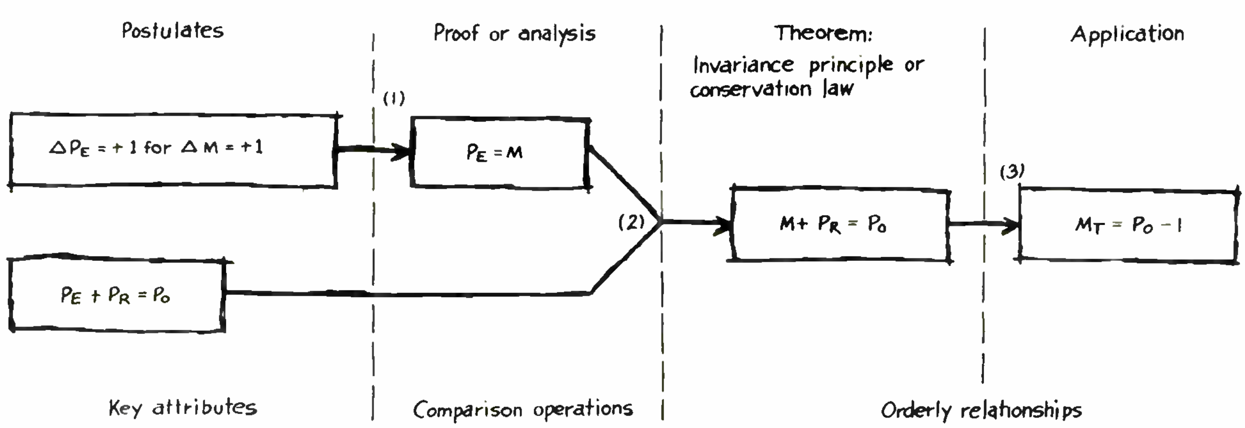
Once a bright student has been taken through one or two of these exercises, he can ask leading questions and provoke a fellow student to accomplish the creative search process. The use of a student-teaches-student program has many advantages: It puts the quick learners to work in roles of leadership such as they may well have in later life and it keeps them active at solving the real problem of imparting knowledge to a colleague. It overcomes the "convoy" effect in which the speed is determined by the slowest by making new tasks available to the fastest. The slower students like the program because they receive personal attention and they need not worry about asking foolish questions of the teacher.

The limited experimentation that has been carried out convinces me of the soundness of the student-teaches-student program, but I fear that it will not receive the contract support it deserves. It does not have the glamour of integrated circuits. There is no profit-motivated incentive for contractors to promote such programs as there is for computerized programmed material.

Let me remark that flashy performance on a problem like the tennis tournament problem is far from a sure sign of competence in progressing through a CSP. I have seen a junior high school student get the answer in two minutes; in contrast, I have also succeeded in getting eminent scientists to pore over it for ten minutes or more. Nevertheless, there is no doubt but that the scientists are better able to bring order out of chaos than high school students.

* The qualified law form should include the *when not* condition to allow for fewer matches in case of default or withdrawal. This rules out "each match eliminates one player" because that is true no matter whether there are defaults or not. The best form I have found is, "It is *always true* that to eliminate a player requires a match *no matter* how many players, how the byes are arranged, how many matches are played simultaneously, *but not if* a player defaults." This postulate is an ACOR structure.

Fig. 5. Outline of the logical structure for the tennis tournament problem.



The scientist's training has conditioned him so that he is convinced that he can solve the silly problem if it's worth his time—and, furthermore, has given him a battery of search thinking tools, such as those discussed in the next section, to search out the order hidden in the chaos. The high school student without luck flounders helplessly, unhappily, and may quit. However, I have seen this behavior improve when the student becomes analytically aware of the realities portrayed by CSP and of the importance of finding how key attributes can be fitted into orderly relationships of the postulate-proof-theorem form.

The separation of one's reasoning into the postulate-proof-theorem form of reasoning shown in Fig. 5 is a form of analytic introspection that is also useful in research and in technical conferences. However, one often encounters a rebuff when one attempts to find out what postulates are actually being used. The questioner is not in the know. He doesn't belong to the club. The articulated science approach is intended to encourage a more constructive and exploratory attitude.

The third scientific thinking tool is the qualified law form,[2] which was illustrated in the footnote to the tennis tournament problem. In science, QLF is particularly important in stressing that most laws have both wide ranges of application and at the same time serious limitations. An example is Ohm's law:

(*What* is true) It is *always true** that for a two-terminal electronic conductor with majority carriers only, V/I is constant,

(*When* it's true) *no matter* whether holes or electrons are involved or if ac or dc conditions prevail,

(*When not* true) *but not if*† the power furnished alters the specimen temperature or the field is high enough to produce hot-carrier effects or the frequency is high enough that a displacement current is comparable to a conduction current or if collision times are comparable to the ac period.

The main feature brought out by this example is not perfection of statement but emphasis on three aspects: (1) *what* it says—a useful relationship is stated; (2) *when* it is practical—you can use it in many circumstances; and (3) *when not* applicable—it does have its limitations.

The qualified law form gives the perspective lacking in the usual definitions of a law of science: "an established fact," "a universal truth."

Before discussing the remaining tool on our list, the ACOR pattern, the role of the search thinking tools in Fig. 3 will be explained.

Search thinking tools and frustration

One of the objectives of the articulated science approach is to increase general awareness of the power of useful principles and of rational reasoning in everyday life.

At present, teaching of postulate-proof-theorem logical structures is largely confined to geometry. However, new-math geometry does not do its best in suggesting how rational reasoning applies to real life. For geometry the postulates now selected in modern mathematics are characterized by mathematical neatness rather than correspondence to objective experimental reality. This basis for selection of postulates inhibits transfer of the postulate-proof-theorem structure to real life. It loses the advantage of Euclid's appeal to the realities of rulers and stretched strings. I fear that the emphasis on mathematical nicety rather than relationship to everyday phenomena may be analogous to the suppression of phonetics a decade ago. We may develop Roberts who can't reason about space just as we had Johnnies who couldn't read.

As an example of the lack of connection between mathematics and reality in present-day education, I recently had a well-above-average graduate student in electrical engineering who argued that we can't really prove that space is three dimensional. He argued that this was only a matter of definition. He now knows more about Bridgman's operational definitions and I use his "bum hunch" as another thinking tool. I also point out how the anecdote about his difficulties with the reality of three dimensions for space helps much as did my bum hunch about the field-effect transistor. As further evidence of the lack of tight connection between various parts of mathematics that I observe from my experience, I estimate that not as many as ten percent of superior high school seniors can easily prove Pythagoras' theorem or prove that $A = \pi r^2$ can be derived from $C = 2\pi r$; perhaps mathematicians in education consider such mundane matters too elementary to discuss beyond the eighth grade. I think these two particular examples of human ingenuity should be selected for dramatized, exaggerated presentation so that they will always be available to the student. (Here I agree with Steinmetz; they should be taught every year.) These are two items to be considered in my recommendations given later.

The articulated science approach is intended to improve rational reasoning ability both in science and in ev ryday affairs. This aspect is dramatized by the competitiv real-life situations of the "take-away" match games. Stu lents frequently exploit the power that the principles give hem by defeating their older siblings after school. (This fe ure might be explored as a partial remedy for the self-est em problems of students with underprivileged backgroun s.)

The random, frustrating aspects brought out by he creative search diagram are intended to reduce feelings of inadequacy on the part of students who feel they are n t working out problems in the way they are supposed t . I have some hope that this aspect of the new approac might reduce the percentage of mathematics haters.

The search thinking tools on the list are supposed to b of help in getting a student under way. If he understands tools 2, 3, and 4, he will look for simplest cases, use pencil and paper, and try to represent quantitative and logical relations symbolically. By doing so he may pump up, as represented in Fig. 3, some relevant ACOR structure out of his reservoir of experience. By properly associating this ACOR with his familiarity with his problem, he may develop the hunch that leads to washing a new logical structure out of a chaotic situation and into his reservoir, as shown in Fig. 3. In this way he can increase the number of reference patterns in his reservoir.

What in Fig. 3 is search thinking tool no. 1? Perhaps you will think that it is, "define the problem." In the tennis tournament, defining the real problem of relating

* Some semanticists object to modifying *always true* by *but not if*. Prof. Gong and I also considered simply *true* and *indeed true*. Our choice of *always true no matter what* gives a dramatic ring we like. Obviously the merit of the QLF does not depend on exact phrasing.

† Originally,[2] I used *provided* something does *not* happen; *but not if* developed while I was working with ninth graders.

matches played to eliminated players was most of the result. In alchemy the real problem was not how to make gold, but the meaning of *you can't, no matter what.* Clearly, defining the problem is not your first step. How then can you describe what you should do first?

The best word I have found for search thinking tool 1 is *action.* One of my standard actions when I am stuck at an equation while reading a technical paper is simply to copy mechanically the equation that blocks my understanding. In this way I load up my short-term-memory or familiarity bank with details about the equation. The action greatly increases the likelihood of a good hunch that will let me see through the problem.

The ACOR pattern

Our short-term, ready-access memories have limited capacity. Overworking them produces fatigue.

One familiar example is the important, long-distance telephone call with poor transmission. You hold a sentence of incomprehensible syllables in your short-term memory until a key recognizable word gives you the clue and these syllables suddenly click into meaning. Under such circumstances, you can't think effectively. After the conversation ends, you feel that you have overworked some brain machinery; you are tired. (When faced with such a situation, I conserve my energies and get a better circuit; at least if I am phoning in the United States.)

A very similar situation occurs when the communication involves foreign language difficulties.

Awareness overload is the name I propose to describe this effect. I interpret the phenomenon of awareness overload as a demonstration of the interference between two uses of the short-term or immediate-access memory banks. When I am holding uninterpreted sounds in awareness, this interferes with my ability to process concepts effectively. I conjecture that a set of neurons in the brain is being monopolized in trying to find associations between uninterpreted memories of sounds and familiar auditory word memories. They then cannot contribute to finding associations between higher level concepts.

The limitation of capacity shown by this example to store new concepts for ready access leads to a common frustrating teaching experience. You wish to impart to a student a demonstration of how A and B imply C. You impart concept A; then you impart B—but when you try to discuss the relationship that makes A and B lead to C, you find A has gone out of immediate awareness for the student. Concept A is apparently stored more remotely, if at all. By repeated processes both A and B finally become well enough established to be simultaneously available to awareness and the argument can be completed. Before this happens your teaching is blocked by awareness overload.

The familiarity in the creative search pattern of Fig. 3 consists of storing essential attributes of the question in a form in which they are available for quick recall. The hunches then occur by some now-mysterious association process that links them to logical structures previously stored. I like to compare this process to the common experience of trying to remember something particular about a person. You may remember that you liked or disliked him. You may remember that his name had two syllables, even that the first letter was D; then often, in a minute or so, the sound or even the spelling comes back. However, you may recall only part of the name. A hunch is like this partial recall. Often your hunch is vague but sometimes it is very specific about the exact relationship of a feature in the problem to one in a pattern you have stored.

My experiences in talking to lawyers, writers, and theatrical directors have impressed me with the generality of the ACOR and creative search patterns. I illustrate with a quotation from John Barrymore: "I love rehearsals. I hate performances." In terms of CSP, this means that Barrymore loved looking for the key attributes of the emotional states of the characters as portrayed by the playwright. He also enjoyed developing hunches during the comparison operations of the rehearsals and creating the orderly relationship that was the finished performance. This is the ACOR pattern of play production. Subsequent noncreative repetitions were what Barrymore hated.

This ACOR structure can be identified in very diverse situations as can be illustrated by expressing the three following cases in parallel form.

1. The number of eliminated players in a single elimination tennis tournament equals the number of matches played.
2. Potential energy defined as weight times height above the floor is conserved in any rotation of a balance lever.
3. Logical forms with their characteristic properties arise within the operation of inquiry and are concerned with the control of inquiry so that it may yield warranted assertions.

In each of the foregoing examples attention has been focused on a significant attribute of the situation:

1. The number of players eliminated and the number of matches played.
2. The potential energies of each of a set of weights.
3. Statements or propositions and their truth or falseness.

Furthermore, some *comparison operation* among the key attributes has been set up. For example, "Socrates is a man"; "All men are mortal"; therefore "Socrates is mortal" relates three statements.

Potential energies are defined in terms of a standard unit. Others are measured with respect to the standard by a comparison operation consisting of comparing weights and comparing heights and multiplying the ratios. (In fact, "standard unit" plus "comparison operation" is the meaning of physical measurement.)

If the lever rotates, all potential energies change in unison.

If a tennis match is played, one more player is eliminated. Increases in eliminated players are compared with increases in number of matches played.

As a neurological hypothesis I propose that the logical patterns of these three examples may all be stored according to one and the same scheme. Another example is the standard postulate-proof-theorem pattern of formal mathematics. The correspondences are illustrated by Table I.

Hunch development would then proceed as follows: Familiarity consists of storing relevant information about the question at hand in those memory banks available for immediate access. The search thinking tools set up opportunities for one-to-one correspondences to be discovered between features of the new problem and features of some stored logical structure. The "quick" and well-trained mind contains a variety of such patterns and makes associations quickly.

What does it take to do this? Why do some brains produce so many more patents and publications than others? I have proposed that this might be related to the number of concepts that they can interrelate at one time.[11] Can this be controlled by some genetic feature such as size of pyramidal cells or length of axones and dendrites in the cerebral cortex? Do the frontal lobes somehow add coding in classifying tags on stored concepts? A frontal lobotomy apparently prevents creation of new reference patterns although it leaves memories intact. How much can we increase effectiveness of rational reasoning by imparting dramatized, exaggerated patterns of logical structures available for transfer? And how can we measure the increases if we do produce them?

There are difficult but, I conjecture, not impossible problems to attack constructively. Answers may well emerge, as Wooldridge[12] emphasizes, from interaction between electronics and neurology.

Some thinking tools in electrical engineering

Superposition, linearity, and symmetry are thinking tools of great importance in solid-state electronics. Consider the Hall effect problem of Fig. 6(A). One analysis predicted that the Hall voltage between points A and B of an isotropic Hall effect plate would remain finite as the disk radius went to infinity. Buehler, Pearson, and the writer[13] concluded it would go to zero.

I tried the problem on Polycarp Kusch and he came to the same conclusion, pointing out the similarity to the one-ohm resistor net of Fig. 6(B). If each resistor is one ohm, the ohmmeter will read 0.5 ohm. Obviously Kusch had a relevant ACOR at hand.

Why half an ohm? Superposition of two simple conceptual experiments gives the answer. Let one ampere flow into junction E and out at infinity (put a large contact ring at infinity in this conceptual experiment), then by symmetry 0.25 ampere flows through the resistor connecting junction E to F. Turn off the source to E and take 1 ampere out of F; again 0.25 ampere flows from E to F. Superimpose the two cases. The ring at infinity plays no role now; throw it away. The superimposed currents produce 0.5 volt between E and F. Therefore the ohmmeter reads 0.5 ohm. Q.E.D.

Can you transfer to the Hall effect problem? Put a ring contact at infinity, pass 1 ampere into C, and use symmetry. The currents don't go radially; they spiral. However, the equipotentials do not. They are single-valued. They can only be circles. Current into C produces no voltage between A and B. Superimpose and you have the answer.

This obvious fact escaped the author of a paper, the reviewing facilities of his organization, and the referees of a reputable journal. The error was hidden somewhere in the mathematics. To transfer: The author and referees were confused by the complexity of scheduling byes in the tennis tournament or the multiple forces on the rowboat.

I consider that an important and challenging engineering educational problem is to produce a set of dramatized, exaggerated reference patterns of linearity, superposition, and symmetry; of dimensional analysis; and of other selected important electrical engineering topics. Their use might both expedite and improve the output of trained engineers—and so might emphasis on student-teaches-student, especially in so far as the important aspect of communicating results is concerned.

Recommendations for a balanced program

Even a superficial inspection of the recent texts prepared by nationally sponsored programs leads to a discouraging conclusion: Some of them are obviously the minimum retainable residue that would keep a writing committee from violent schism. They yield to the illusion that science has covered so much ground that our only hope is to increase the rate of cramming new information into the students' heads. (See the "Prologue.")

The result is an awareness overload of the intermediate-term memory banks. These intermediate-term memory banks have a time constant of the interval from one final examination to the next and discharge very efficiently thereafter.

The articulated science teaching approach can offset this trend. It can continue to emphasize, year after year, the fundamental verity of the ACOR pattern in the human mind:

1. Some things must be accepted; the important ones are key attributes of the problem at hand. This is the A of ACOR.
2. These attributes must be interrelated by some process of logical operations that the student is sure from experience work to yield warranted assertions. This is the CO of ACOR.
3. The outcome is an orderly relationship that makes thinking easier and recall simpler and more systematic. This is the R of ACOR.

I propose that a major aim of research in science education should be the development and perfection of a significant set of dramatized, exaggerated ACOR patterns. These patterns could then be more easily retained by the student. They would also be available for transfer to aid him both in discovery of new relationships on his own

I. Generalized logical structure pattern—ACOR pattern

Key Attribute	COmparison Operation	Orderly Relationship
Players eliminated and matches played	Equal increases per match	Players eliminated, equal matches played
Potential energies, weight times height	Conceptual balanced lever experiments; standard unit of potential energy, comparison with any weight at any height	Conservation of potential energy
Statements and their truth or falseness	Interlocking of parts of statements	Logical forms
Postulates and definitions	Proof; comparison operations among aspects of postulates and definitions	Theorem

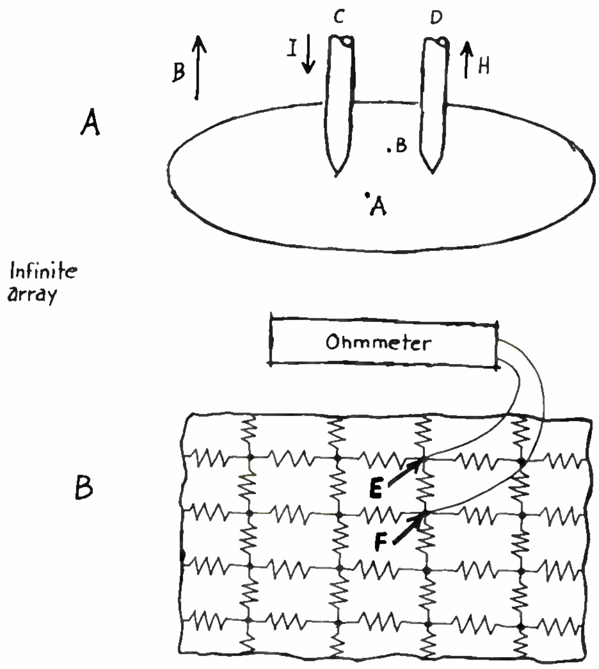


Fig. 6. Example of symmetry and superposition. A—The Buehler–Shockley–Pearson isotropic Hall effect plate. B—The infinite, periodic, two-dimensional square one-ohm resistor net.

and in interpreting scientific principles in his studies.

Our modern educational trend toward a new model of our subject matter each year does not take into account that the human brain is not provably different from what it was 20 000 years ago. (Many now worry that it may be becoming worse, but that is my subject elsewhere.[14-16]) Are we wise to discard Robert the Bruce's spider that showed the value of persistence, the boy who cried wolf, and Euclid? In my opinion, we are not. The educational encyclopedias produced by writing teams are capable of overloading the awareness of any brain save the "idiot savant" who is all memory.

What is the answer? It is selection and *balanced emphasis*. By balanced emphasis I mean: Pick a small, highly selective sample of undeniably important facts and develop optimum exaggerated, dramatized techniques to impart these well. Do these patterns right. Treat superficially all the remaining wonderful variety of accomplishments of modern science and show them as a survey.

I feel it is practical to surmount the objections that no agreed-upon set of first-priority scientific principles can be found. The linearity, superposition examples given in the preceding section are a small existence proof. Slater's detailed balance pattern is another example. With detailed balance a high school student can visualize how a catalyst cannot disturb the end state of a chemical reaction. He does not have to repeat incomprehensible generalized concepts, for example: "Net reaction stops when equilibrium is reached—our operational definition of equilibrium. This must mean that the reaction tendency, called free energy (ΔF), is zero." The fact that Avogadro's number is the number of molecules in 18 unit volumes of water (i.e., cubic centimeters) is the significant factor, not that "a mol is 6.03×10^{23} particles of any kind... atoms... or even people, dollars, or calories." Any trace of an ACOR pattern in this statement is imperceptible. If you think the mol is understood, ask a high school honor student what to change to make a mol 48.24×10^{23}.

Balanced emphasis is my expression for ideas similar to those covered by the "posthole" method or "islands of knowledge." Whatever it is called, developing exaggerated, dramatized, or in any event unmistakably emphasized, experiences on specially selected specific topics will give the student a chance to know what it is to master a logical orderly relationship. By contrast, the survey of the rest of the field of science will show him how vast and exciting is the field of things he has admired from an intellectual distance. And he will know the difference.

This I think is vitally important: He will be able to understand the difference between knowing well on the one hand, and on the other, knowing superficially enough to know the vocabulary and the general purposes but not the tight relationships. With this knowledge the student can have the mental tools to decide better what he wants to do, to know how to do it and when he has done it.

Unrealistic? I, with many others, once worked for three years on a device that one of my colleagues said should be called a "persistor," if it ever became practical!

This material was prepared for the 37th Memorial Steinmetz Lecture 1966 at Union College. It was also presented at the Seventh (1966) National Creativity Conference and at the Second (1966) International Symposium on Communications Theory and Research. Portions of this work were supported by Office of Education Contract OE 4-10-216, Project S-090.

Editor's note: A treatment of "mechanics" incorporating the articulated science teaching approach described in this article is being prepared by Dr. Shockley, with W. A. Gong, for publication in September in the Merrill Physical Science Series, published by Charles E. Merrill Books, Columbus, Ohio.

REFERENCES

1. Shockley, W., and Mason, W. P., "Dissected amplifiers using negative resistance," *J. Appl. Phys.*, vol. 25, p. 677, May 1954.
2. Shockley, W., "Proposed important mental tools for scientific thinking at the high school level," *Science*, vol. 140, p. 384, Apr. 26, 1963.
3. Einstein, A., "On the quanta theory of radiation," *Phys. Z.*, vol. 18, pp. 121–128, Mar. 15, 1917.
4. Nyquist, H., "Thermal agitation of electric charge in conductors," *Phys. Rev.*, vol. 32, pp. 110–113, July 1928.
5. Shockley, W., and Read, W. T., Jr., "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, pp. 835–842, Sept. 1, 1952.
6. van Roosbroeck, W., and Shockley, W., "Photon-radiative recombination of electrons and holes in germanium," *Phys. Rev.*, vol. 94, pp. 1558–1560, June 12, 1954.
7. Shockley, W., and Moll, J. L., "Solubility of flaws in heavily doped semiconductors," *Phys. Rev.*, vol. 119, pp. 1480–1482, Sept. 1, 1960.
8. Shockley, W., and McDonald, F. J., Rept. on Contract OE-4-10-216, Proj. S-090, U.S. Dept. of Health, Education and Welfare.
9. Shockley, W., "A unipolar 'field-effect' transistor," *Proc. IRE*, vol. 40, pp. 1365–1376, Nov. 1952.
10. Prim, R. C., and Shockley, W., "Joining solutions at the pinch-off point in field-effect transistors," *IRE Trans. on Electron Devices*, vol. ED-4, pp. 1–14, Dec. 1953.
11. Shockley, W., "On the statistics of individual variations of productivity in research laboratories," *Proc. IRE*, vol. 45, pp. 279–290, Mar. 1957.
12. Wooldridge, D. E., *The Machinery of the Brain.* New York: McGraw-Hill, 1963.
13. Buehler, M. G., Shockley, W., and Pearson, G. L., "Hall measurements using Corbino-like current sources in thin circular disks," *Appl. Phys. Letters*, vol. 5, no. 11, pp. 228–229, 1964.
14. Shockley, W., in *U.S. News and World Rept.*, Nov. 22, 1965.
15. Shockley, W., in *Stanford M.D.*, Stanford Medical Alumni Assn., Jan. 1966.
16. Shockley, W., *Genetics and the Future of Man.* Amsterdam: North Holland Publishing Co., 1966.

Worldwide color television standards

Francis C. McLean *British Broadcasting Corporation*

The choice of a color television standard will increasingly affect both the broadcasting authorities and the public in most countries of the world. Although there are considerable differences of opinion about what should be done, those countries that have already reached a decision are very fortunate to be spared much of the heart-searching taking place in those countries that have not yet done so. In Europe, work on these problems is in full swing; in fact, several countries already have announced plans for the start of regular color service in the second half of 1967. How many systems will be used will not be clear until after the Oslo Conference, being held in June 1966; but it seems almost too much to hope for that a single standard will result—desirable and logical as this would be.

Television program exchange all over the world is advancing very rapidly, with techniques now available for relaying by satellite that were virtually undreamed of a decade ago. Not only will these facilities increase at a very rapid rate, but direct broadcasting from satellites will probably also be achieved in the near future. This increased use of television program exchange adds strength to the argument for the greatest possible degree of agreement on world standards of color television, or at least for the provision of adequate means for translating one set of television standards to another and one color television system to another.

In Europe, there is unity on the 50-field approach, but not much unity in other matters—even in black and white. For historical reasons there are, at present, different line standards in various countries, although there is a move toward unification. Perhaps we can expect that a 625-line system will be universally used in Europe before too many years have passed. The world will then settle down for a number of years to 625- and 525-line pictures. However, it looks as though the world will be split into the 50-field area and the 60-field area for a very long time to come, although eventually we may have a uniform line frequency in all areas. This was one of the fortunate provisions of CCIR's television planning some 15 years ago. A common line frequency greatly simplifies many aspects of the problems of program exchange.

Program exchange in the future is much more likely to be on video tape than on film, and so three conversion problems will exist: (1) changing the field frequency; (2) changing the line frequency—a problem largely solved by recent developments; and (3) conversion between television systems. There is so far only one color television system in use in the world: the NTSC 525-line system, employing a 6-Mc/s band. However, other possibilities include a 625-line NTSC system with an 8-Mc/s band, or one or two additional systems using the basic principles of the NTSC system in the separation of the chrominance and luminance signals but with different methods of modulation on the chrominance subcarrier. There are, therefore, program conversion problems of considerable complexity, but the work that will be outlined should go a long way toward resolving them. There is also, of course, widespread hope that the number of television systems to go into service will be smaller than now seems possible. All countries realize the importance of reducing standards to a minimum, and this aim will be in the minds of all the delegates to the discussions at Oslo.

Another matter of the greatest importance in the choice of a color system is the cost of the color receiver, its ease of manufacture and maintenance, and its reliability in the hands of the public. The results of experience regarding these factors will come from the United States. This information—though of course derived from the operation of an NTSC system—will be of great interest in connection with other systems as well, because 95 percent of the receiver is quite independent of the coding technique used. The information that we receive, therefore, will be of particular interest since much of what happens in the receiver applies equally to all systems, based as they are on the NTSC system's fundamental concept of the separation of the signals into a luminance and a chrominance signal.

Full text of a paper presented at the IEEE International Convention, Session 25, New York, N.Y., March 21–25, 1966. Also being published in IEEE Transactions on Broadcast and Television Receivers, May 1966.

Compatible color transmission and

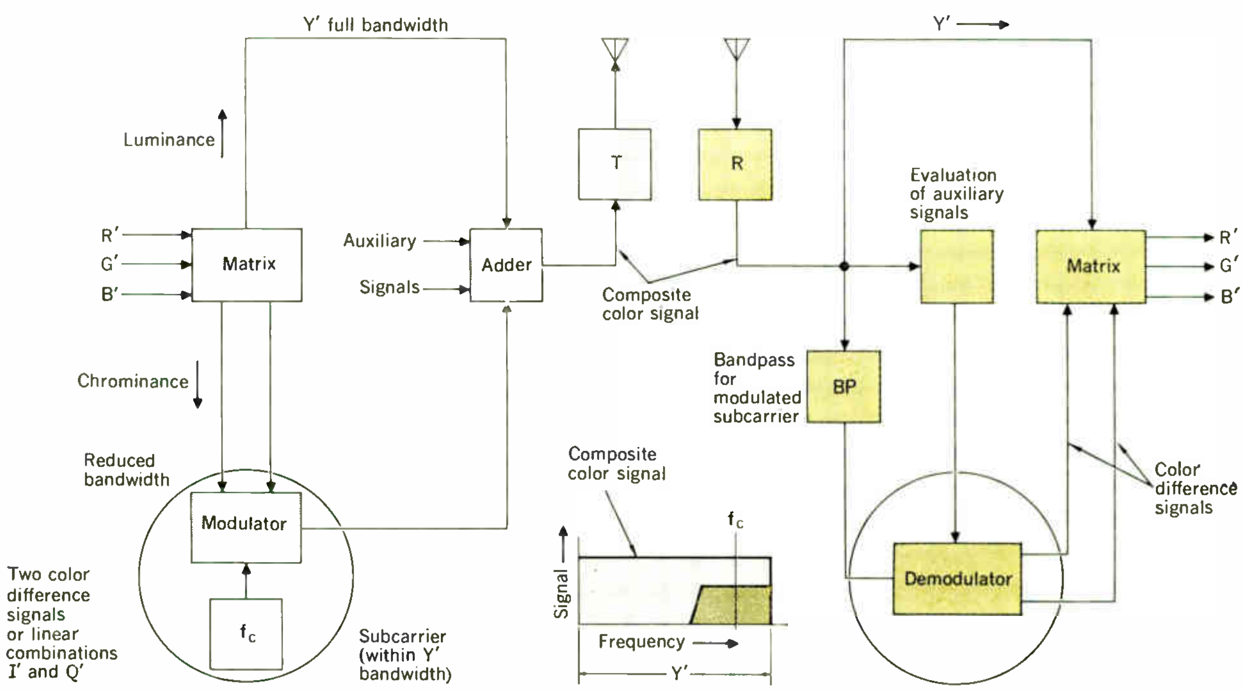
Attempts are now under way in Europe to formulate standards for color television transmission. A number of methods have been devised for the transmission of the chrominance and luminance color signals. These techniques are all characterized by signal separation, the use of difference signals, and interleaved frequency bands; in other words, they are all based on the National Television System Committee proposals adopted when a standardized color television system was established for the United States some years ago. Their differences lie only in the means employed for subcarrier modulation and demodulation.

The stages in the development of color television may be characterized by the different ways of assigning signals to the color parameters. It is manifestly relevant and logical to relate one signal to the brightness of a given picture element (the luminance signal), and to assign two other signals to represent the "kind of color" or chrominance—that is, the hue and degree of color saturation.

However, no picture source (camera or film scanner) in existence directly provides signals in this form, nor is there at present any picture display device that can be directly actuated by such signals. Therefore, the color information is analyzed and reproduced within three "subspectra" (red–green–blue) of the total visible spectrum. There is an important stipulation here: The color television chain must always begin and end with three primary signals E_R, E_G, and E_B. What still remains to be settled and optimized is the suitable form of color signal transmission through the telecommunication channel to the receiver.

It will be readily appreciated that the early development of color television started with direct transmission of the three primary signals; the field sequential version of this method had great attraction from the point of view of technical simplicity. However, this technique has not been adopted in television broadcasting because of its fundamental disadvantages. The main drawbacks are that

Fig. 1. General schematic diagram, showing the compatible transmission of color television, with separation of the luminance and chrominance signals and insertion of frequency bands of color modulation in the luminance band.



European modulation methods

Various groups and individuals in Europe have proposed several techniques for subcarrier modulation and demodulation of color television signals. All of these approaches employ the compatible principles of the NTSC system used in the United States

Richard Theile *Institut für Rundfunktechnik, Munich*

it is essentially wasteful in its bandwidth requirements and that it is not compatible with the usual monochrome system.

Fortunately, a new line of development was progressing at the right time, based on studies of the behavior of the human eye for color vision, which proved that the perception of color and brightness differ materially from each other with regard to small details. To put this knowledge into effect—i.e., to use it for frequency bandwidth saving—new correspondence of signals to the color parameters and new methods of coding and channeling signals had to be devised. By an admirable series of developments, such methods led step by step to the NTSC (National Television System Committee) system of the United States.[1-3] Since these novel principles are basic to all the variants to be discussed, it will be useful to summarize them briefly.

The first step that needed to be taken to reduce the information transmitted was to effect the transformation of information relating to brightness and color content as obtained from the three primary color signals into new signal forms, separated and directly related to luminance and chrominance.

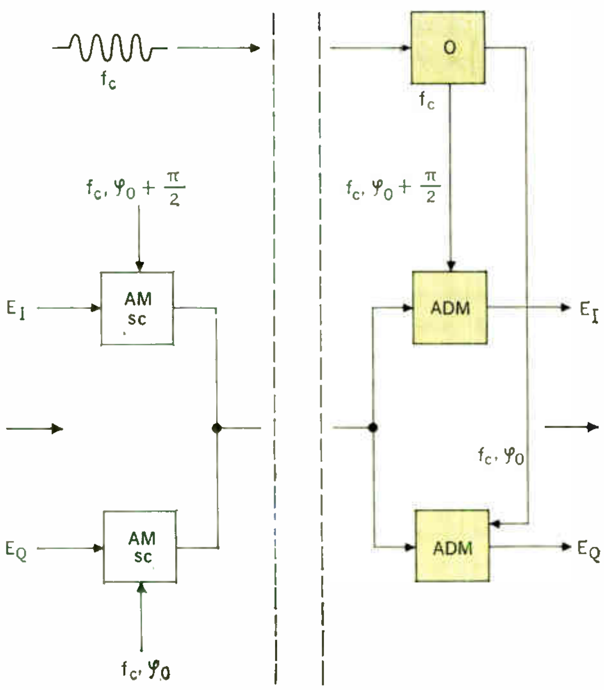
Another significant step was the choice of color difference signals whose amplitudes increase and decrease with the color saturation, having the great advantage that they are zero if no color is present.

The separation in luminance and chrominance signals allowed the allocation of different frequency bandwidths for each, and resulted in significant overall bandwidth saving because the chrominance requires much less bandwidth than the luminance.

Another important step in the development of the NTSC system was the bold concept of the simultaneous transmission of both luminance and color information by intercalating their frequency spectrum within the video band—i.e., the "band-sharing" technique. For this it was necessary to choose a suitable subcarrier frequency in an offset position, accommodated at the high-frequency part of the band. Finally, an efficient AM-quadrature-modulating technique with suppressed carrier was introduced to transmit the two color signals with one subcarrier only. Figure 1 shows in schematic form the practical accomplishment of this kind of color television transmission.

The techniques introduced with the NTSC system have justified themselves in practice, and even the expert continues to be impressed by the way in which this system manages to imitate the thrifty information-handling mechanism of our visual sense organs. Practically all subsequent developments followed along the same line, characterized by signal separation, the use of difference

Fig. 2. Block diagram showing modulation and demodulation of the color signals in the NTSC system.



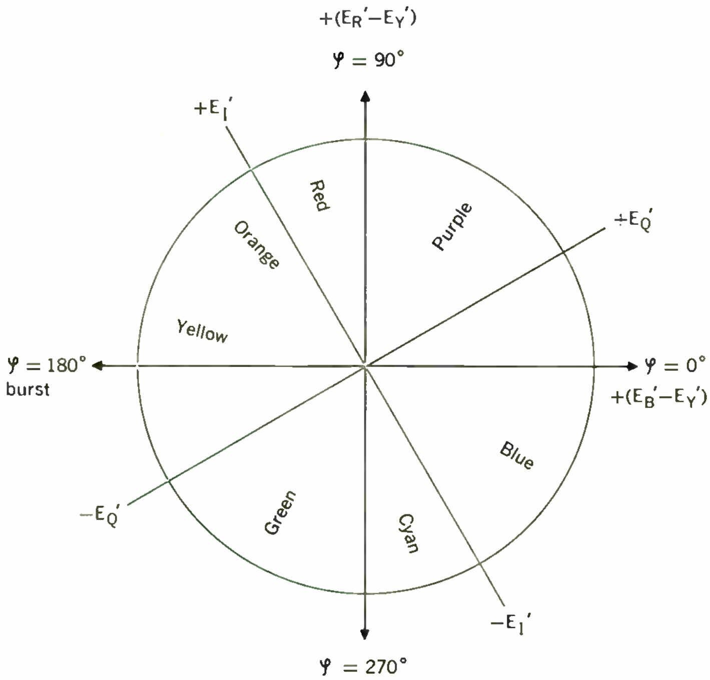
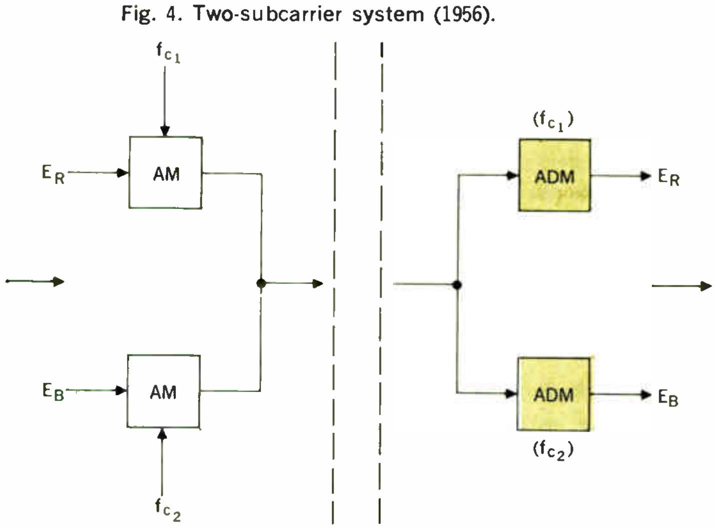


Fig. 3. Vector diagram of the amplitude double modulation of the color subcarrier in the NTSC system, showing the position of the reference axes for the original and transformed color difference signals.

Fig. 4. Two-subcarrier system (1956).



signals, and interleaved frequency bands. The differences in the variants to be described solely relate to the technique for transmitting the two color signals—that is, to the type of subcarrier modulation and demodulation.

As Fig. 2 shows, transmission was accomplished in the original NTSC system by a simultaneous double amplitude modulation of the subcarrier f_c in precision offset, differing in phase by 90°—that is, by quadrature modulation. The amplitude of the subcarrier (shown as "sc" in Fig. 2) varies in proportion to the intensity of the color signal and is therefore zero in the absence of color information. The synchronous detection process at the receiver necessarily requires a control signal, which is transmitted in form of the "burst" in the line-blanking interval.

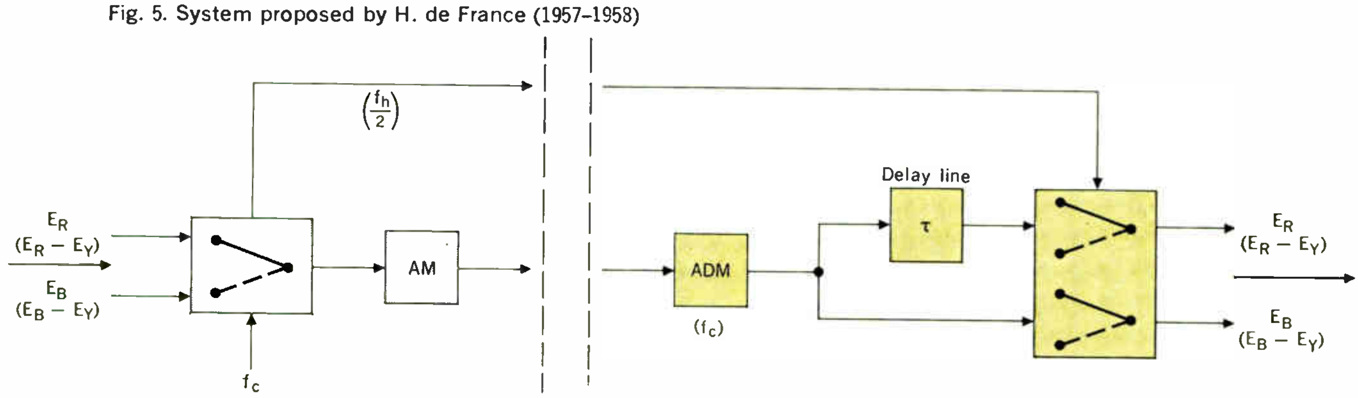
This type of modulation produces an amplitude modulation and a "zero phase modulation" of the subcarrier. In a very reasonable correspondence, the amplitude is proportional to the color saturation and the phase to the hue. Figure 3 indicates the corresponding vector diagram. It is important to note that the phase angle of the combined chrominance signal determines a highly significant quality parameter, the hue of the color.

An important practical consideration is the influence of distortions. If there is, for example, a phase variation within the modulation range (a so-called "differential phase error"), noticeable changes in hue can occur; and, since hue is a very decisive factor in the picture quality, some criticism has arisen from the susceptibility to distortion of this kind. Also, the quadrature modulation is sensitive to other forms of distortion, such as errors in amplitude symmetry within the range where the sidebands of the two signal components overlap, a distortion that gives rise to crosstalk, causing colored transient effects.

Of course, the importance of all these effects on the quality of the color television picture depends on the degree and statistical occurrence of the distortions. This subject could promote endless discussions. It is true that, in the light of technical progress, these distortions can be kept within narrow tolerances in the studio, distribution network, transmitter, etc. There is, however, always some uncontrollable influence along the path from the transmitter to the receiver. Therefore, the problem was also treated from the other side—that is, by investigating various modifications of the color signal modulation technique in an attempt to achieve greater immunity against certain distortions.

Many workers in Europe applied their efforts to this problem, and new proposals for alternative solutions were

Fig. 5. System proposed by H. de France (1957–1958)



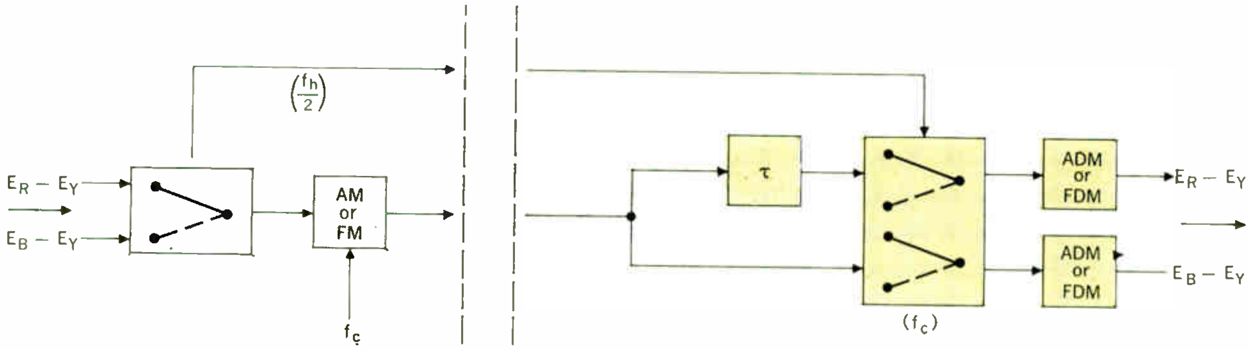


Fig. 6. SECAM system. Amplitude modulation and demodulation were used in the 1959–60 system; frequency modulation and demodulation in the 1960 system.

made, such as use of separate subcarriers for each color signal and the application of different modes of modulating one subcarrier for each signal (e.g., AM and FM). Use of a delay line for a one-line period resulted in the principle of sequential transmission of the two color signals line by line (which necessitates decoding information from the preceding line). There are also the principle of phase reversal of one color signal line by line, color-phase alternation, and, finally, the addition of a reference signal to the luminance signal, either continuously or with interruptions sequentially in every other line.

Some of the proposals that have shown promise and have been tried experimentally will be discussed and shown schematically (in approximate chronological order of the available reports of these developments). Because these variants refer only to the color signal transmission, we shall consider here only those parts of the overall modulation system that are depicted within the circles in Fig. 1.

A proposal (Fig. 4), developed and tested some ten years ago by the Philips Laboratories at Eindhoven, Netherlands,[4] sought to avoid the difficulties of double modulation by using two different subcarrier frequencies, f_{c_1} and f_{c_2}, for two color signals (these being the ordinary primary color signals R and B). Conventional amplitude modulation is used (without carrier suppression and without a burst signal). However, new problems were encountered in connection with the compatible monochrome picture, with the stabilization of the ratio of the two color components (which determines hue), and with limitations in the signal-to-noise ratio for color because of restrictions in the subcarrier amplitudes, etc.

At about the same time, an interesting proposal by Henri de France[5] was being studied (Fig. 5). This technique also avoids the use of double modulation. Only one color signal at a time is impressed upon a subcarrier at frequency f_c, but this subcarrier is modulated by sequential switching, alternately by the first and second color signal. Thus, at any given instant, one of the color signals is absent. Conventional amplitude modulation was used, originally with primary color signals E_R and E_B and later with the difference signals $E_R - E_Y$ and $E_B - E_Y$. The transmission of a burst signal is unnecessary with simple envelope detection at the receiver, but a synchronizing signal is transmitted in the field-blanking interval to control the phase of the electronic switch in the receiver. The switching process follows a line-to-line sequence, and so the recurrence frequency is half the line frequency f_h.

At the receiver, a delay line with a delay equal to a line period is needed for holding the color information transmitted during earlier line periods to ensure a continuous flow of the signal information necessary for the reconstitution of the primary color signals. There may be small errors (in the receiver) in the vertical direction; however, they are within acceptable tolerances, since the reduced bandwidth of the color signals implies diminished resolution in the horizontal direction. In the early stages of development, the delay line was designed to operate at video frequencies, and was placed immediately after the detector.

Further developments of this method were carried out by the Compagnie Française de Télévision in Paris.[6,7] An important advance was the successful solution of the problem of designing a cheap delay line. The solution took the form of an ultrasonic line that could be used at carrier frequencies. Figure 6 shows the modifications made in the sequential switching technique, in which the new type of delay line is placed before the detector. This system was given the name SECAM (a derivation of "séquentiel à mémoire"). A further stage of development was a change from conventional amplitude modulation to frequency modulation.

At the same time, various remedial measures were introduced in order to reduce the effects of interference from the subcarrier—always present in FM—upon the compatible monochrome picture, to reduce the sensitivity to noise, and to minimize crosstalk effects from the luminance components.

Optimum results were found with phase reversal of the subcarrier frequency from field to field and in every third line. Special signal-forming processes, as shown in Fig. 7, were also used to reduce the effects of interference from the subcarrier. Video pre-emphasis and de-emphasis (PR and DE) are introduced before modulation and after demodulation, and amplitude-shaping (SH) is employed after modulation in the form of progressive amplitude change disposed symmetrically above and below the subcarrier frequency, with corresponding de-emphasis correction shaping in the receiver before demodulation in the feed to the delay line and the electronic switch. In addition, to reduce crosstalk effects from luminance signal components, the chrominance signal is caused to vary in amplitude by a modulator AM, which is con-

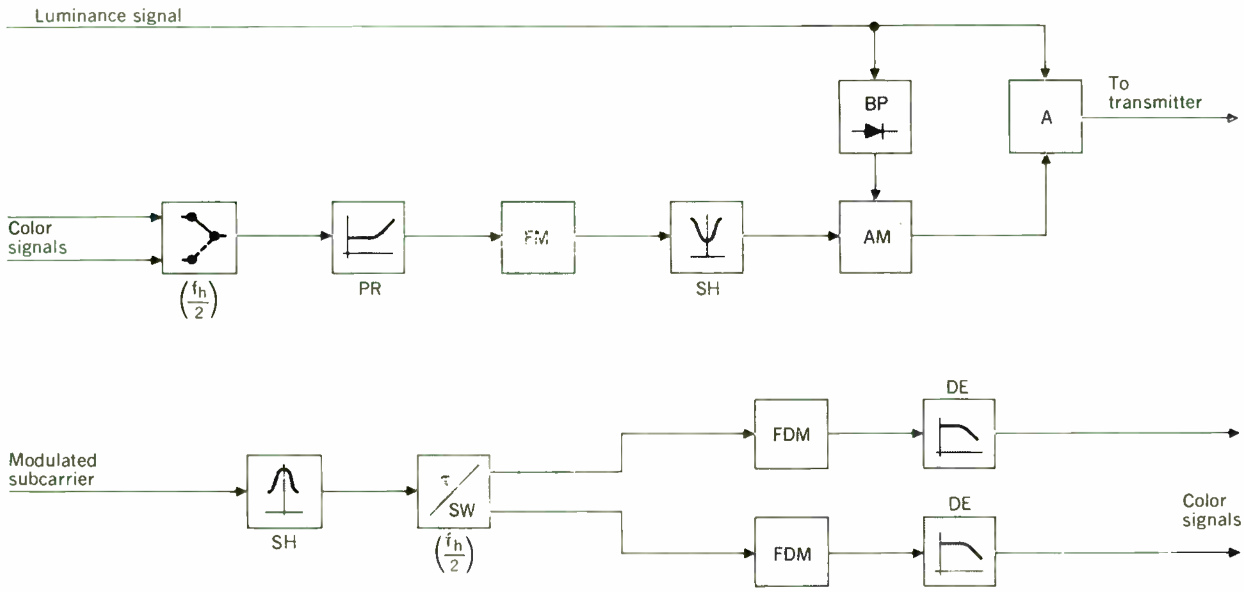


Fig. 7. Details of the formation of signals in SECAM.

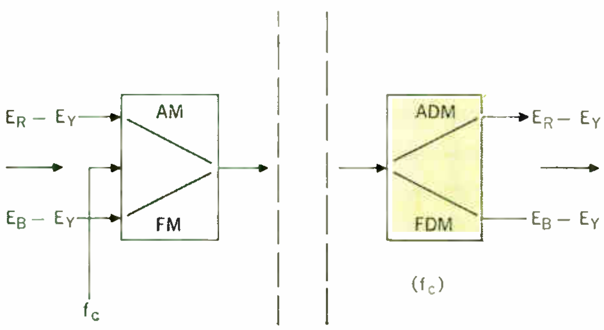


Fig. 8. FAM system (1960).

trolled by those parts of the luminance signal that lie within the band of the color carrier modulation, derived by means of a bandpass filter and a rectifier.

The SECAM color signal transmission method was one of the variants selected for the study and for comparison work of the Ad Hoc Group on Color Television of the European Broadcasting Union (EBU).

Figure 8 shows an interesting technique, studied in detail by N. Mayer of Institut für Rundfunktechnik (IRT), Munich,[8] which eliminates sensitivity to differential phase error effects by use of different forms of modulation for the two color signals. Conventional amplitude and frequency modulation are used. Synchronizing signals are unnecessary since conventional detector circuits are used in the receiver, so this method is very attractive from the point of view of technical simplicity. There are, however, other problems, such as the visibility of subcarrier in the compatible picture and dependency of the hue on the amplitude ratio of signals transmitted from different forms of modulation, with correspondingly different reaction to distortions.

The proposal that is shown in Fig. 9 was made by W. Bruch, of Telefunken A.G., as an attempt to combine the SECAM sequential switching system in some way with the efficient method of modulation used in the NTSC system.[9] Instead of the amplitude or frequency modulation of the SECAM system, suppressed carrier AM is introduced. This makes a burst signal necessary for the receiver. As in SECAM, only one color signal at a time is transmitted and the phase sensitivity of quadrature modulation is avoided. An auxiliary synchronizing signal is also necessary to control the phase of the electronic switch at the receiver.

Further developments in sequential transmission with suppressed carrier modulation led Bruch to what is known as the PAL (phase alternation line) variant (Fig. 10).[10] Here the sequential switching principle is concerned not with the selection of one of the signals to modulate the color subcarrier but with the periodic phase alternation of one of the two simultaneously transmitted signals. In this sense, the technique resembles that of the NTSC system, since the form of modulation amounts to quadrature amplitude modulation of the subcarrier. The difference lies only in the periodic switching from line to line of the phase of one of the component signals between $+90°$ and $-90°$. This switching corresponds to alternate changes in sense of the color vector. The errors in hue due to differential phase effects in transmission will therefore alternate in opposite senses and, in a suitably designed receiver, will be averaged out.

The basic idea of an alternating transmission of this kind was already well understood when the NTSC system was being developed. It had been described by Loughlin, Brown, Hirsch, and others, as oscillating color sequence and color phase alternation.[11,12] Phase alternation was duly considered for reducing crosstalk between wide-band asymmetric-sideband quadrature-modulated signals. In view, however, of the alternative of two signals differing in bandwidth with a region of channel overlap within which both sidebands of both signals (*I* and *Q*) are transmitted, it was not pursued further in the United States.

The proposals of Bruch resulted in material improvements in technique. Attention was concentrated upon line-to-line switching and upon the choice of a suitable

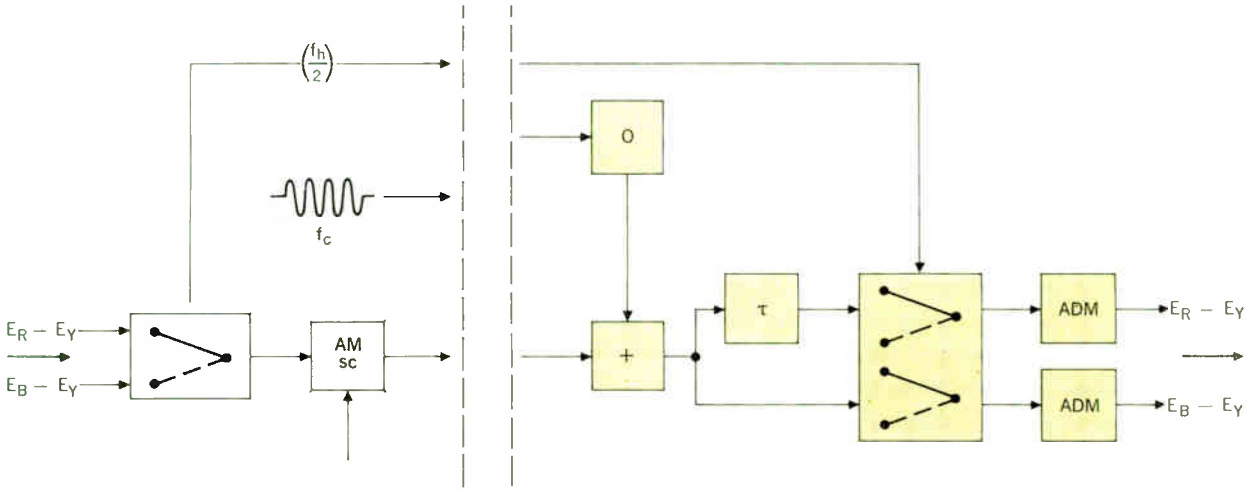
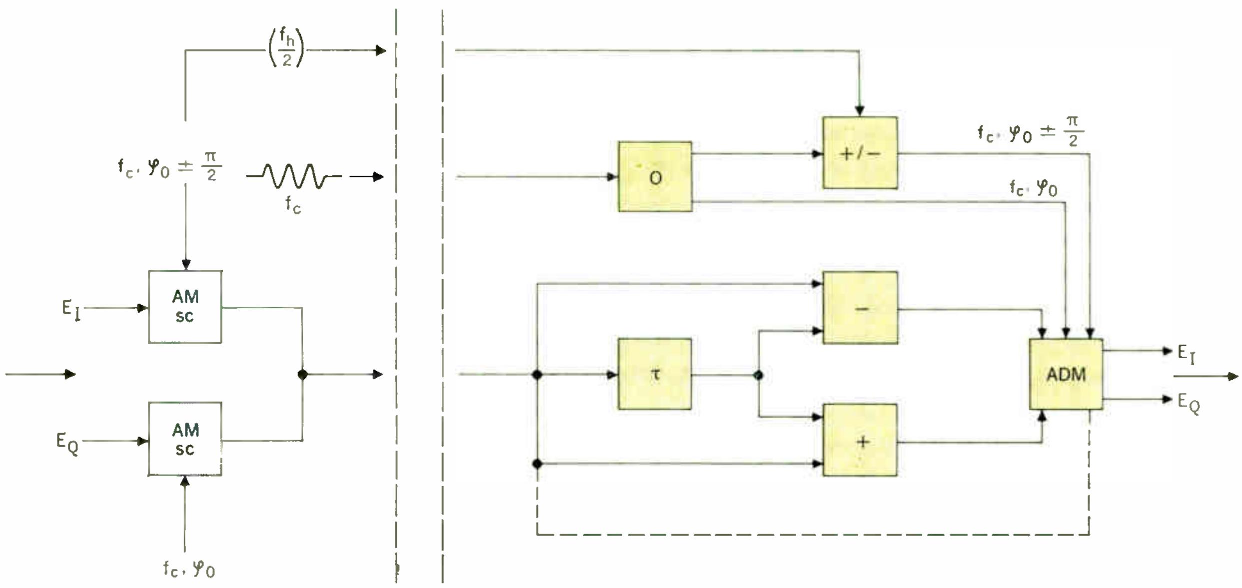


Fig. 9. SECAM-NTSC system (1961–1962).

Fig. 10. PAL system (1962).



new subcarrier frequency f_c in quarter-line offset

$$f_c = (284 - 1/4) f_h + 1/2\, f_v$$

where f_h = horizontal frequency and f_v = vertical frequency.

Another important novelty was the utilization of color information contained during the preceding line period for deriving hue information free from the effects of transmission errors in several respects.

At the receiver, synchronous detection is essential and the phase-alternation signals will need to be demodulated. If the receiver is restricted to making use of the color signals as transmitted, the effects of transmission errors will be opposite in sense from line to line and the opposite effects will be integrated by the eye. Reception of this kind, indicated by the dashed lines in Fig. 10, is quite feasible ("simple PAL"), but in the presence of differential phase errors, highly saturated colors give rise to noticeable stroboscopic "venetian blind" interference patterns. Therefore, it is advantageous to associate the PAL principle with an ultrasonic delay line, as used and developed for inexpensive mass production in the SECAM system. Such a delay line, when accurately set to the correct group delay and used in the circuitry indicated in the figure, makes it possible for the mean value of the two color signals over two adjacent lines to be utilized at the receiver. Thus, the error integration takes place during viewing, and stroboscopic disturbance is avoided.

For PAL, two auxiliary signals are required: the burst of subcarrier frequency f_c and the identification signals at frequency $1/2 f_h$ for phase-locking the electronic switch. The latter signal can be transmitted as in SECAM, during the field-blanking interval.

Another proposal that was tried out led to the latest version of PAL, shown in Fig. 11. Here, the burst signal is made to perform both synchronizing operations. The phase of the train of oscillations of the reference frequency f_c is displaced alternately from +45° to −45° relative to the reference axis, which is the $-(E_B - E_Y)$ axis. The receiver adjusts itself to the correct average (reference) phase. Furthermore, the color synchronization signals can be applied to the oscillator-controlling discriminator, from which can be derived the signal controlling the phase of the electronic switch, as shown on the right-hand side of Fig. 11. Also, it was agreed recently to modulate with the simple $E_R - E_Y$ and $E_B - E_Y$ difference signals rather than the transformed color difference signals I and Q, with a maximum bandwidth for both signals not exceeding that of the I signal. This specification will lead to greater flexibility and lower cost of the design of receiver circuits.

The PAL method of color signal transmission was the other variant selected by the EBU color group for careful study.

To complete the survey, the technique of reference signal addition has to be mentioned. As seen in Fig. 12, normal NTSC color signal transmission is used with suppressed carrier amplitude modulation. However, according to the proposal developed by N. Mayer and G. Holoch[13] of IRT, a low-amplitude pilot or reference signal is superimposed on the outgoing complete television signal (chrominance + luminance). It, therefore, shares in any level, depending upon the distortion to which the latter may be subjected. After recovery, it is used to control the oscillator of the synchronous quadrature detector, thus continuously maintaining correct phase lock. This technique is referred to as "additional reference transmission," or ART.

The superimposed reference signal must be distinct from the color subcarrier, so that the two sets of information can be separated—that is, so that picture modulation cannot effect control of phase. As an identifying characteristic, a periodic line-to-line phase reversal is chosen. The ART method can provide instantaneous steering, whereby with a delay line in the receiver, the reference signal information is extracted and operates directly upon the demodulator. Alternatively—using a cheaper arrangement without delay line—the reference signal information, after periodic reversal, is applied via a selective filter to the synchronous detector. In this case, however, the error correction is averaged over several line periods. Synchronization may also be effected, if so desired, by means of a burst signal (see dashed lines in Fig. 12); thus, if the need should arise, receivers either of conventional NTSC design or specifically intended for ART can be interchangeable.

I should like now to add some unofficial remarks. Recently another idea for a color transmission system was

Fig. 11. PAL system (1965).

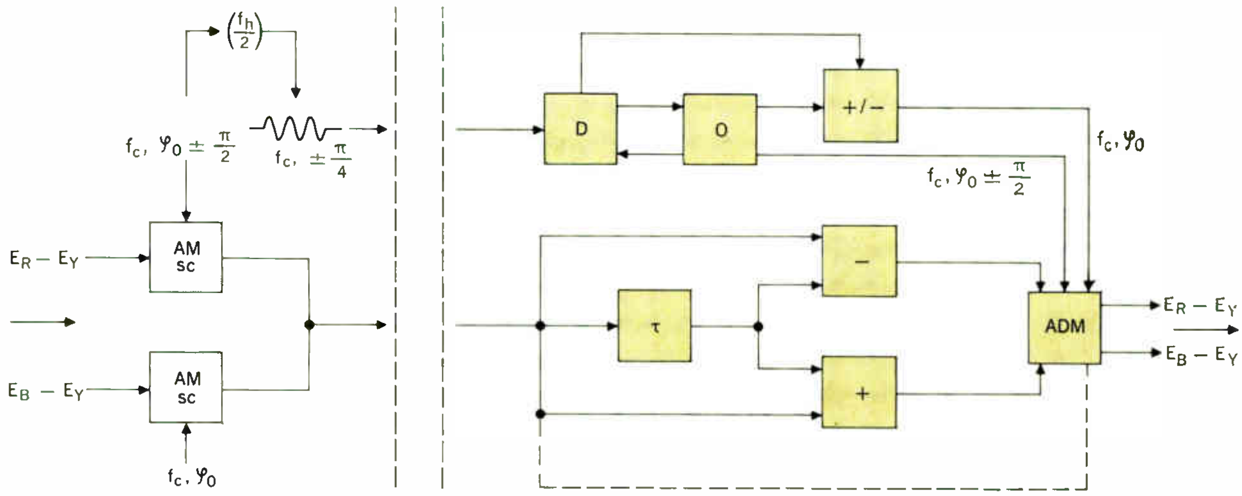
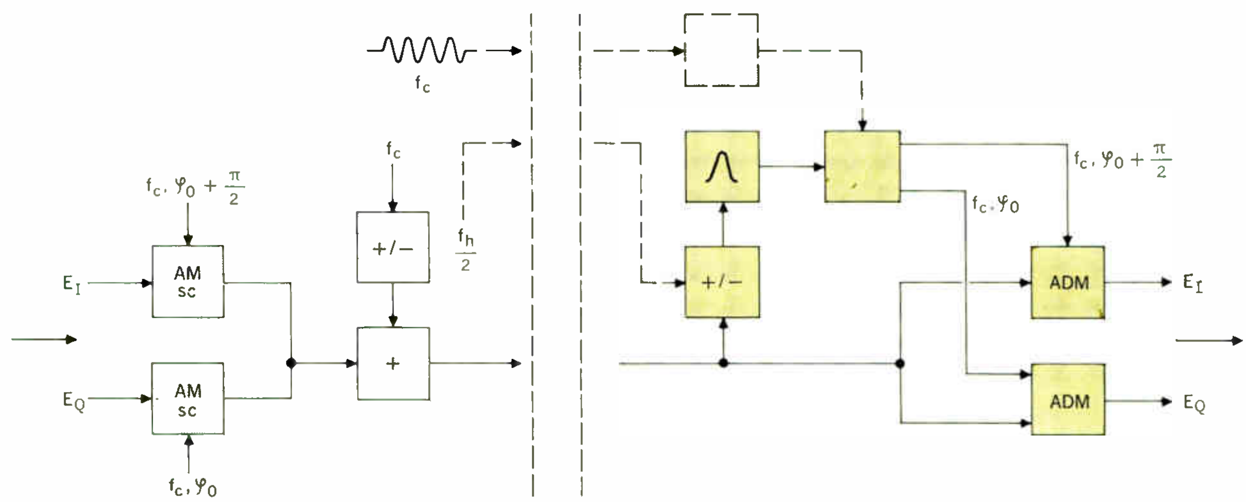


Fig. 12. ART system (1963).



announced. It is a sequential version of the principle of reference signal addition, originating in the U.S.S.R., also mentioned in an internal report of B. W. B. Pethers of the BBC, London, written in 1963. Authoritative information on the details is not available, and this variant has not been officially proposed so far. However, the fundamental principle is known, for example, from some press reports.[14,15] Figure 13 indicates how it works. The basic idea is to transmit in one line an NTSC-type chrominance signal in quadrature modulation and, in the following line, a reference signal of constant phase, modulated with the saturation signal amplitude. The reference signal again rides up and down on the luminance information as explained for ART and meets the same kind of distortion as the chrominance signal in the neighboring line, which is practically at the same level in normal cases. A switch operated at one-half line frequency (½ f_h) connects to the output, either to the chrominance signal or to the reference, so that the signals transmitted on alternate lines of any one field are the luminance signal plus either an amplitude-modulated subcarrier f_c with null phase modulation φ (hue) or an amplitude-modulated subcarrier of constant phase φ_r, which can be the phase angle of the axis of one of the color difference signals. In the original version, the amplitude of the subcarrier is the square root of that of the NTSC subcarrier, but the usual linear relationship seems to be preferred now.

In the receiver, a delay line of the PAL type and a synchronized switch are necessary to regain in one branch the chrominance information only and in the other one the reference information only, which is then used for the demodulation process. The switching process does not allow the NTSC half-line offset of the subcarrier; therefore, as with PAL, a quarter-line offset condition has been used.

Finally, I should like to mention another interesting proposal (Fig. 14), made by N. W. Lewis[16] of the Post Office Research Station in London, to reduce differential phase errors by automatic control in the transmission chain. This proposal, therefore, is not for a different modulation technique, but it is of some interest for the discussions on distortion in general. As shown in Fig. 14, the normal burst signal is accompanied by two superposed signals of the same kind, which are set at gray and peak-white levels ("multiburst" technique). The distortion of these supplementary signals can be measured at

Fig. 13. Additional reference signal in sequence line by line.

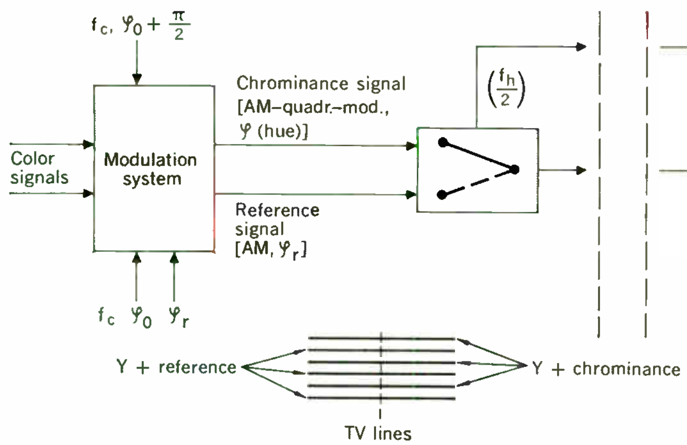
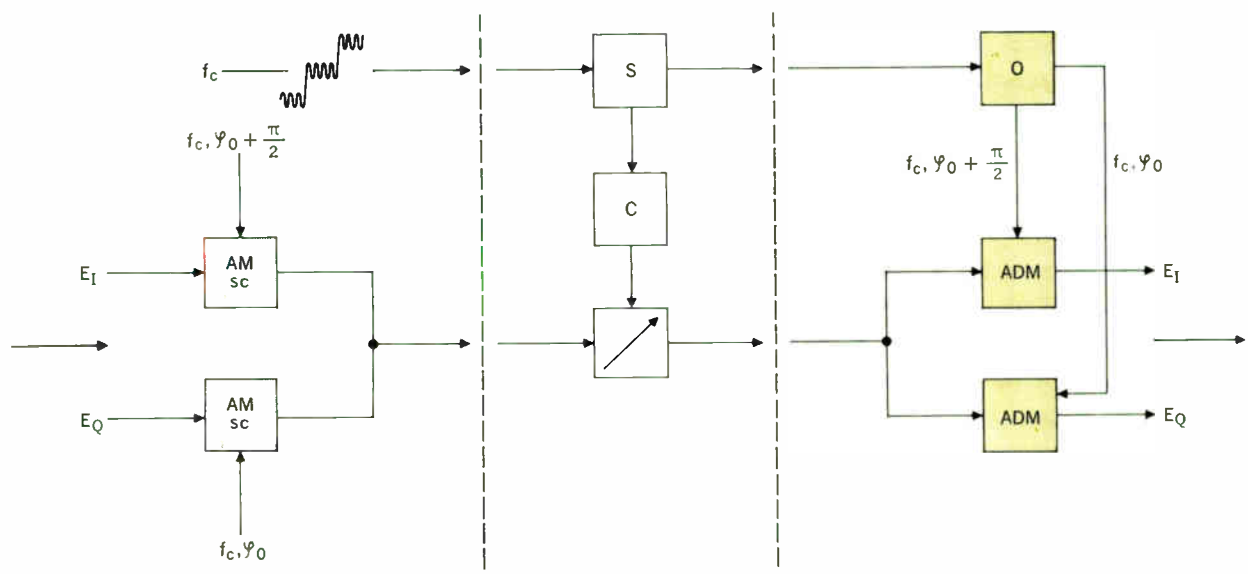


Fig. 14. Multiburst technique (1964).



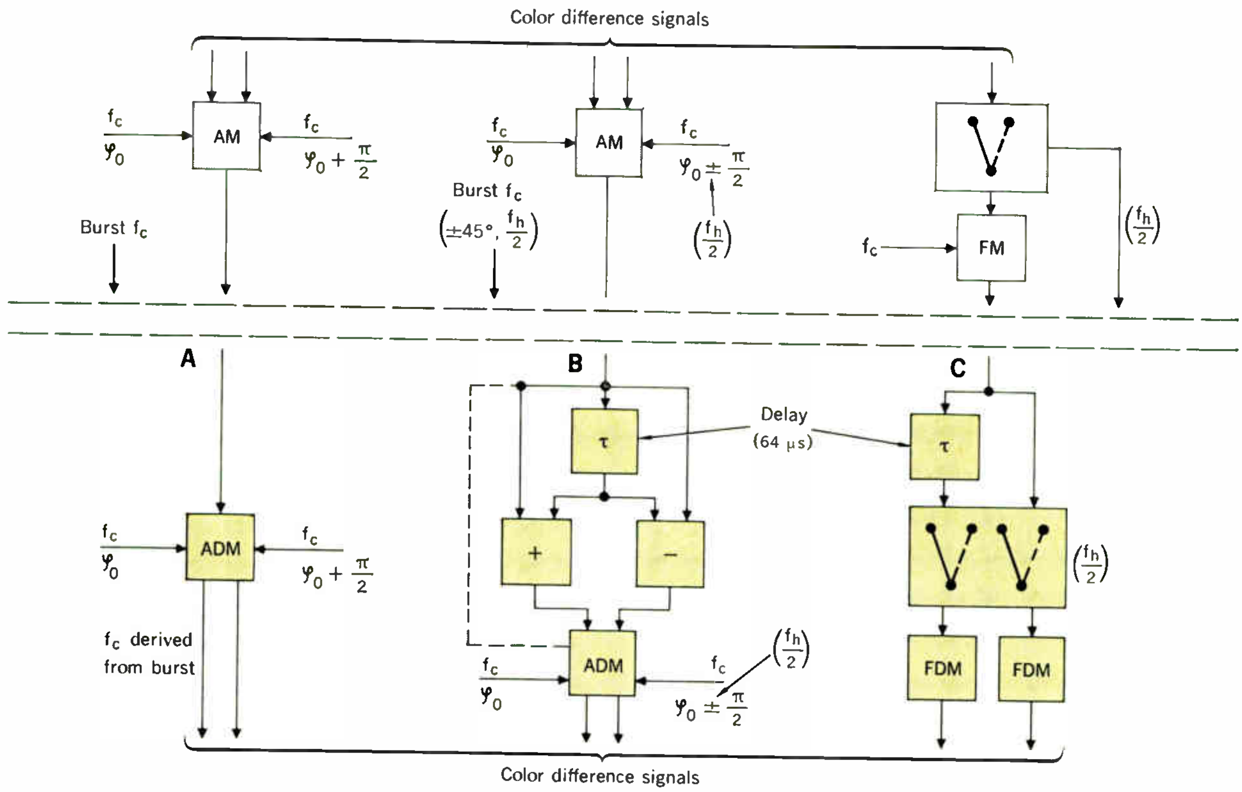


Fig. 15. Comparison of (A) NTSC, (B) PAL, and (C) SECAM systems.

different points on the transmission path by a device *S*. By means of another device *C*, correction signals can be made to control an equalizer.

In the endeavors during the past few years to reach agreement on a single color standard for Europe, a comprehensive study of the different variants in color signal transmission has been made. In particular, comparisons were made of the performance of SECAM and PAL with that of NTSC (Fig. 15). As you can imagine, assessing advantages and disadvantages was a problem of great complexity. The results have been discussed elsewhere.[17]

In conclusion, I should like to say that variants other than SECAM and PAL have also been considered in this survey, in order to show that many experts, laboratories, and organizations have contributed to this work. All of these contributions, directly and indirectly, have been fruitful in furthering progress. Those developments that have not been singled out for official proposal have made their contribution, too, by provoking and stimulating discussion. But, as I said before, the described variants refer to the color signal transmission technique and we must not forget that all of them use the main, basic, compatible principles of the NTSC system, the development of which in the United States was a unique and admirable contribution to the modern art of electronics.

Revised text of a paper presented at the IEEE International Convention, Session 17, New York, N.Y., March 21–25, 1966. Also being published in IEEE TRANSACTIONS ON BROADCAST AND TELEVISION RECEIVERS, May 1966.

REFERENCES

1. Fink, D. G., *Color Television Standards NTSC.* New York: McGraw-Hill, 1955.

2. Wentworth, J. W., *Color Television Engineering.* New York: McGraw-Hill, 1955.

3. McIlwain, E. E. K., and Dean, C. E., eds., *Principles of Color Television.* New York: Wiley, 1956.

4. Haantjes, J., and Teer, K., "Compatible colour-television," *Wireless Engr.*, vol. 33, pp. 3–9, 39–46, Jan., Feb. 1956.

5. de France, H., "Le système de télévision en couleurs séquentiel-simultané," *L'Onde Electrique*, vol. 38, pp. 479–483, 1958.

6. Cassagne, P., and Sauvanet, M., "Le système de télévision en couleurs SECAM," *Ann. Radioélectricité*, vol. XVI, pp. 109–121, Apr. 1961.

7. Townsend, G. B., "New developments in the SECAM color television system," *Television Soc. J.*, vol. 10, no. 6, pp. 172–190, 1963.

8. Mayer, N., "Farbfernsehübertragung mit gleichzeitiger Frequenz- und Amplituden modulation des Farbträgers (FAM-Verfahren)," *Rundfunktech. Mitt.*, pp. 238–252, Dec. 1960.

9. Bruch, W., "Some experiments with modifications of the NTSC color television system," presented at the 1962 Internat'l Television Symp., Montreux, Switzerland.

10. Bruch, W., "Das PAL-Farbfernsehsystem. Prinzipielle Grundlagen der Modulation und Demodulation," *Nachrtech. Z.*, vol. 17, pp. 109–121, 1964.

11. Loughlin, B. D., "Recent improvements in band-shared simultaneous color-television systems. Pt. II—Color-television systems with oscillating color sequence," *Proc. IRE*, vol. 39, pp. 1273–1279, Oct. 1951.

12. Brown, G. H., and Luck, D. G. C., "Principles and development of color television systems," *RCA Rec.*, pp. 144–204, June 1953.

13. Mayer, N., and Holoch, G., "NTSC-Farbfernsehübertragung mit additivem Referenzträger," *Rundfunktech. Mitt.*, pp. 157–165, June 1965.

14. "From Russia—A political compromise colour television system?" *Wireless World*, vol. 72, pp. 73–74, Feb. 1966.

15. "B.B.C. man anticipates Russian colour TV system?" *Wireless World*, vol. 72, pp. 75–76, Feb. 1966.

16. Lewis, N. W., "Proposed subcarrier pilot for NTSC-type color television," *Electronics and Power*, p. 207, June 1964.

17. *Report of the EBU Ad Hoc Group on Color Television.* Brussels, Belgium: EBU Tech. Center, 2nd ed., 1965.

Hydrocarbon–air fuel cell systems

Because they operate on air and "omnifuels" (such as gasoline, kerosene, jet fuel, or diesel fuel), hydrocarbon–air fuel cell systems appear promising for many general applications in the future; despite the many technical problems that must be solved before they become economically feasible

C. Gordon Peattie *Texas Instruments Incorporated*

The fuel cell is basically an energy-conversion device in which chemical energy is isothermally converted into dc electricity. Several types of fuel cell systems have been developed, including the hydrogen–oxygen system used in certain space missions, but many of them employ fuels that are expensive and not readily available. Considerable research and development work is being done on hydrocarbon–air systems, which use air and "real-world" fuels. In the direct-oxidation cells, the hydrocarbon fuel is oxidized directly at the fuel electrode. Because of reaction problems at the electrode, these cells are still in their early stages of development. In the indirect-oxidation cells, now in the systems engineering development stage, the hydrocarbon fuel is converted into an impure hydrogen, which may then be purified to a certain degree and injected into the fuel cell modules. Initial applications will probably be in the military, because of the fuel flexibility and ease of maintenance inherent in these fuel cell systems.

The Gemini flight in 1965 focused attention on fuel cell systems and their use as power sources for certain space missions. After the flight, there were announcements of intentions to apply these and similar space-power fuel cell systems to earthbound uses in the near future. In space, these systems are operated on pure hydrogen and oxygen. There are some special applications on land or on or under the sea that might justify the use of relatively expensive hydrogen and oxygen. There are probably more uses that could justify using hydrogen and air. However, for general use as sources of power in land or marine situations, fuel cell systems will have to operate on air and such real-world fuels as gasoline, jet fuel, kerosene, or diesel fuel. This requirement, imposed mainly by economics, brings with it technological problems that are different in kind from those posed by the space-use situation.

In trying to get around these problems, which stem both from the fuel and from the oxidant (air), the hydrocarbon–air fuel cell researcher can often create new problems that are just as hard to solve; nevertheless, progress in this work is being made. It is the purpose of this article to present the status of hydrocarbon–air fuel cell technology at the present time and to discuss some potential uses of such systems when they are finally brought to the marketplace.

Background

The fuel cell principle—that is, the conversion of chemical energy to electric energy by a path that can avoid the thermodynamic limitation on efficiency imposed by the Carnot relation*—has intrigued scientists and engineers for more than a century. In 1839–1842, Sir William Grove[1,2] probably invented the first fuel cell.† He used platinum-catalyzed electrodes to combine hydrogen and oxygen so as to produce electricity. It is interesting to note that almost 123 years later the Gemini fuel cell used the same catalyst, though probably in different physical form. One major trouble with Grove's cell was that its voltage fell off badly when an appreciable current drain was put on it.[3] In 1889, Mond and Langer[4] made a hydrogen–oxygen cell with perforated platinum-sheet electrodes, catalyzed by platinum black. This cell produced 1.46 watts at 0.73 volt at about 50 percent efficiency. However, it contained 1.3 grams of platinum and required pure hydrogen and oxygen. Thus, its capital cost made it a poor buy as an electric generator. In addition, to complicate things, it ran well only on pure hydrogen and oxygen.

*Carnot described an ideal engine, the efficiency of which cannot be exceeded by any heat engine that operates within the same temperature limits. His relation shows the maximum heat-engine efficiency to be directly proportional to the factor, $(T_{high} - T_{low})/T_{high}$, where T_{high} is the absolute temperature (°K) of the incoming working fluid and T_{low} is the absolute temperature of the working fluid rejected to the cold sink.

†In 1801, Sir Humphrey Davy had invented a cell that used zinc and oxygen in an electrolyte. This cell, the precursor of the air-breathing battery, or air cell, did permit the direct conversion of chemical energy into electric energy. If you want to consider zinc as a fuel, this Davy cell can be considered a fuel cell.

These economic factors, together with the successful demonstration by Edison in 1882 of central station power production and the invention of the first ac system of power generation and distribution by Lucien Gaulard and John Dixon Gibbs in Europe and by William Stanley in the United States (1885) as well as the invention of the induction motor by Nikola Tesla (1888), materially slowed the development of the fuel cell as a practical generator of electricity.

Nevertheless, considerable fuel cell research was done over the period from 1890 to 1930. An excellent résumé of the work done in these years is that by Baur and Tobler.[5] Baur was himself a fuel cell researcher whose work with high-temperature (to 1000°C, or 1832°F) cells was summarized by Allmand and Ellingham.[6]

In the 1950's, a resurgence of fuel cell R & D occurred. In the United States, the need for auxiliary power in space systems and for silent, lower-maintenance, more efficient power sources in the military was a prominent driving factor. In Europe—England, France, Netherlands, Sweden, Germany—the impetus came essentially from a desire for more flexible and efficient land-based generators of electricity, together with increasing discovery and subsequent production of hydrocarbons on the Continent. From 1955 to 1960, a large number of fuel cell types were proposed and investigated. Claims and counterclaims were made, some of them legitimate, others naïve, and a few smacking of charlatanism. Engineering ambition outran much-needed basic knowledge. As a result, the potential of fuel cells was exaggerated, if not oversold. Eventually, however, the hard facts of economic and technological life and just plain common sense tempered the situation. The necessary basic understanding is still lacking, though it is growing. As is often true with new systems, there is broad lack of agreement on what kind of specific impact fuel cells with have on the market place. It is generally conceded that they do have sufficient technological and economic potential to justify considerable industrial research effort. An overall view of fuel cell technology three years ago was summarized previously.[7]

Basic principles of hydrocarbon–air fuel cells

There are two types of hydrocarbon–air fuel cells: indirect-oxidation and direct-oxidation systems. Most development engineering is now being done on the former. In this system, the hydrocarbon fuel is first converted to hydrogen, which is then oxidized at the fuel electrode (anode). As shown in Eq. (1), free electrons, represented by the symbol e^-, are given up to the anode.

$$\text{Fuel electrode reaction: } H_2 = 2H^+ + 2e^- \qquad (1)$$

The flow of these electrons, collected by the anode and transferred to the external circuit, makes up the electric current produced by the cell. After they have done work in the external circuit, the electrons are consumed at the air electrode (cathode) in a reaction, represented by Eq. (2), involving the oxygen of the air.

$$\text{Air electrode reaction: } 2e^- + \tfrac{1}{2}O_2 = O^{--} \qquad (2)$$

The cell circuit is then completed by an ionic (*not* electronic) current through the electrolyte back to the anode. Adding Eqs. (1) and (2) gives Eq. (3), which symbolizes the overall reaction for the indirect-oxidation fuel cell.

$$\text{Overall cell reaction: } H_2 + \tfrac{1}{2}O_2 = 2H^+ + O^{--} \; (= H_2O) \qquad (3)$$

At this point, it is necessary to point out that Eqs. (1), (2), and (3) are not necessarily literally characteristic of any indirect-oxidation fuel cell systems now under development. However, to consider each kind of hydro-

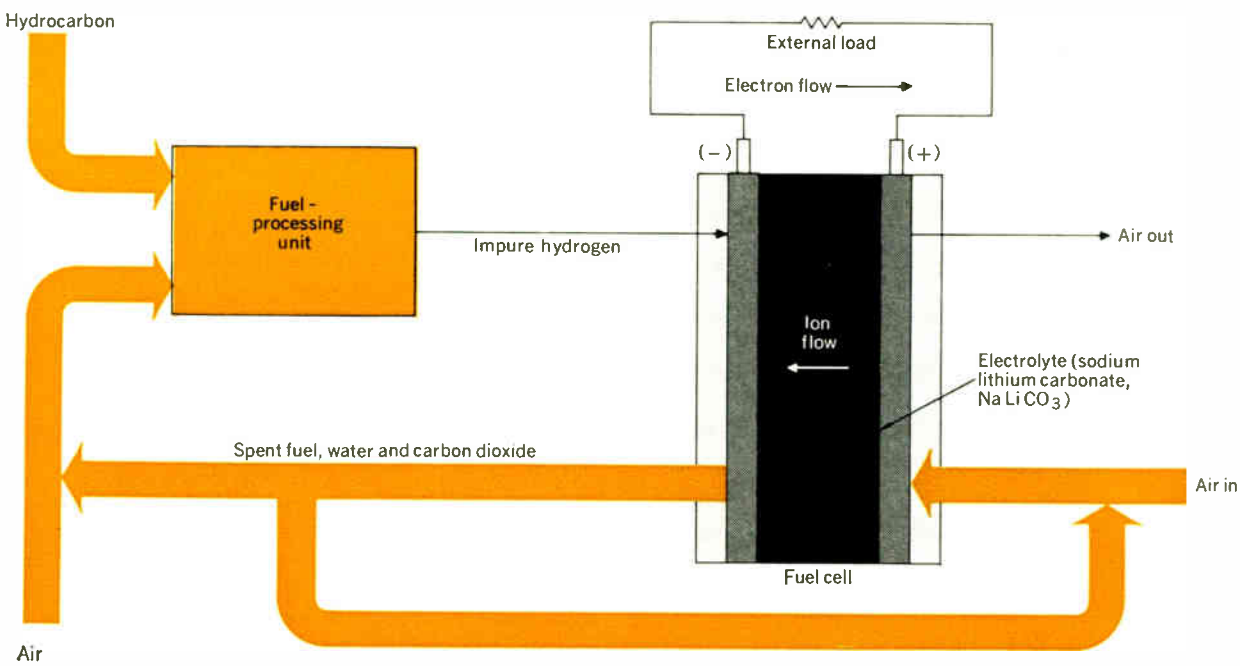


Fig. 1. Schematic diagram of indirect-oxidation hydrocarbon–air fuel cell system now under development. Cell voltage ≈ 1.0 volt; cell temperature = 600–650°C (1112–1202°F). (Elements are drawn out of scale for purposes of clarity and only pertinent parts of the system are shown.)

carbon–air indirect-oxidation system* in correct technical detail might well obscure the simple but important point that the basic operating reaction for this type of fuel cell is the air-oxidation of hydrogen to water.

A schematic diagram showing the operational sequence of one type of indirect-oxidation hydrocarbon–air fuel cells is shown in Fig. 1. This particular cell system, which uses molten electrolyte, is now in the systems engineering development stage.

Unlike the indirect system just described, the direct-oxidation hydrocarbon–air fuel cell system does not involve any prior processing of the hydrocarbon to form hydrogen. Instead, it is based on the direct reaction of the hydrocarbon at the fuel electrode, wherein the hydrocarbon fuel is oxidized and gives up electrons to the fuel electrode. These electrons then flow through the external circuit and are eventually consumed at the air electrode in the manner described. An example of such a direct, electrochemical oxidation of hydrocarbon fuel† is that observed for methane[8]:

$$CH_4 + 2H_2O = CO_2 + 8H^+ + 8e^- \qquad (4)$$

It can be seen that this reaction also involves the water of the cell electrolyte.

Systems now under development

Hydrocarbon–air systems now under development include three types of fuel cells: (1) those using aqueous electrolytes; (2) those using molten electrolytes; and (3) those using solid electrolytes.

Aqueous-electrolyte cell systems. Either acid or alkaline electrolytes are used in aqueous-electrolyte cells. The former generally employ sulfuric or phosphoric acid or acidic fluoride solutions. Alkaline electrolytes usually consist of solutions of sodium or potassium hydroxide. Aqueous-electrolyte cell types ordinarily operate in the temperature range of 60–150°C (140–302°F). From a viewpoint of reaction chemistry, these are relatively low temperatures. To compensate for the low electrode reactivity inherent in cell operation at these low temperatures, relatively active electrode catalysts are used.‡

An example of a sulfuric-acid-electrolyte cell system is the one developed by Shell Research Ltd.[9-11] Examples of an alkaline-electrolyte cell system are those developed by Allis-Chalmers,[12,13] Union Carbide,[14-16] and Justi.[17] These are all indirect-oxidation cell systems. Ion-exchange membrane cells[18] also fall in this class.

Developments in direct-oxidation fuel cells have recently been summarized by researchers at Esso Research and Engineering Company[19] and General Electric Company.[20]

Each of these two aqueous-electrolyte indirect-oxidation cell types has its system advantages and disadvantages. Unlike the alkaline-electrolyte cell, the acid-electrolyte cell is unaffected by the carbon dioxide content of the air (usually 0.03 volume percent). With an alkaline electrolyte, it is necessary either to remove the carbon dioxide from the air before admitting air to the cell or to provide a way of replenishing the alkaline electrolyte as it is consumed by the carbon dioxide. The former procedure is the better, since it avoids any possibility of clogging tiny electrode pores with insoluble carbonates. However, the alkaline-electrolyte cell usually presents a more efficient air electrode (cathode) and fewer corrosion problems than does the acid-electrolyte cell.

These particular fuel cells are categorized as low-temperature fuel cells. At their operating temperatures they require relatively active, and often expensive, electrode catalysts. The catalysts are usually sensitive to "poisoning" (a term used to denote a marked decrease in catalyst activity produced by certain materials) and therefore the hydrogen fuel should be purified before admission to the cell. At present, diffusion through a silver–palladium alloy is used to purify the hydrogen. Research is under way toward achieving catalysts that poison less easily. Recently, platinum-based catalysts have been developed that show a reasonable tolerance for carbon monoxide in acid-electrolyte cells. Since carbon monoxide is almost always present in any fuel made by processing hydrocarbons, this is a significant achievement. Poisoning by sulfur compounds, such as hydrogen sulfide, in the processed fuel is still a major catalyst problem.

The direct-oxidation cell is, of course, what everyone would like. It typifies the devices, described in the Sunday-supplement type of "things to come" article, in which you just add gasoline to a black box and get instant electricity. This is a legitimate goal, but we must work extremely hard to achieve it. Much of the work will involve understanding and eliminating the formation of polymer-like intermediates that can form when a hydrocarbon fuel is oxidized. These "polymers" can coat catalysts and block electrode reactions. Anyone who has tried to clean the gums and varnishes out of a lawnmower carburetor that he inadvertently left full of gasoline over the winter can appreciate the nature of this problem. The direct-oxidation cell also has an economic difficulty to overcome: At present, relatively large quantities of platinum are required to make it work.

Molten-electrolyte cells. The first of the two types of molten-electrolyte fuel cells uses a hydrate of potassium hydroxide as electrolyte, at a nominal operating temperature of 216°C (400°F). This cell system is being developed by the Pratt and Whitney Division of United Aircraft Corporation for use in the Apollo moon mission.[21] As it is used in the Apollo command service module, this cell system operates on pure hydrogen and pure oxygen. In order to use it with real-world fuels and air, it is necessary to purify the hydrogen produced from the fuel and to scrub the air free from carbon dioxide.

The second type of molten-electrolyte-fuel cell system uses a molten-carbonate electrolyte—usually a eutectic mixture (solid solution with sharp melting point lower than that of any of its components) of sodium, lithium, and/or potassium carbonates—operating in the temperature range of 500–750°C (932–1382°F). In the United States, it is being developed by the American Gas Association[22] and also by Texas Instruments Incorpo-

*An example of an indirect-oxidation nonhydrocarbon–air fuel cell system is the ammonia–air system. The ammonia is first thermally dissociated to nitrogen and hydrogen, and then the hydrogen reacts at the fuel electrode, as described previously. The inert nitrogen does not react; it only dilutes the hydrogen.

†Direct-oxidation nonhydrocarbon–air cell systems also exist. Two that show merit are hydrazine–air and methanol (methyl alcohol)–air cell systems.

‡These low temperatures need not necessarily mean low electrochemical reactivity. Hydrazine, a highly reactive material, shows high electrochemical activity at room temperature. Enzyme processes, on which our lives depend, take place at body temperature. Fuel cells using both hydrazine and enzymes are under development; however, hydrazine is expensive (see Table 1) and enzymes are too easily inactivated.

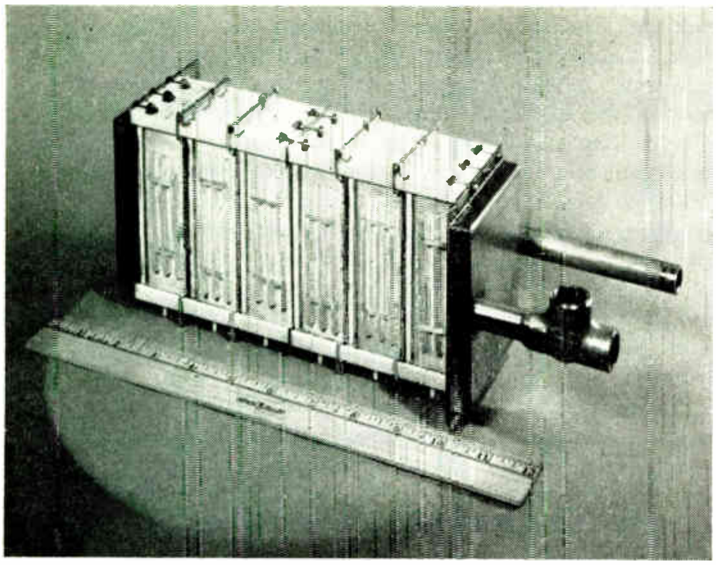


Fig. 2. View of 36-cell hydrocarbon–air fuel cell module.

I. Approximate costs of various fuel cell system fuels

Fuel	Cost, cents per pound
Hydrazine	100–105
Gasoline (92 octane)	1.9–5.5
Ammonia	4.6–4.7
Methyl alcohol	4.1
Hydrogen	3.6–4.0
Kerosene	1.5
Diesel fuel	1.4
Liquefied petroleum	0.86–1.7
Petroleum gas	2.0–2.4
Natural gas	0.40–2.6

Price variations are attributable to such factors as location, quantity, and taxes.

rated[23]; in Britain, by Energy Conversion Ltd.[24]; in the Netherlands, by Central Technical Institute TNO[25]; and in France, by Gaz de France[26] and Electricité de France.[27]

Molten-electrolyte cells do not require active electrode catalysts, because at their operating temperatures (500–750°C or 932–1382°F) the rates of electrode reactions involving activated processes proceed much more rapidly than at the temperatures at which aqueous-electrolyte cells operate. An associated characteristic of the higher-temperature operation is a lack of sensitivity to catalyst poisons. Therefore, low-cost electrode materials can be used and, in addition, fuel flexibility is achieved. These cells can also run on air containing carbon dioxide; in fact, they require the addition of carbon dioxide to the air reacted at the cathode (air electrode). With fuels such as kerosene, the fuel itself is the source of this carbon dioxide. The example of a molten-carbonate fuel cell module shown in Fig. 2 contains 36 cells in a series–parallel hookup.

An additional system advantage of molten-carbonate fuel cells is that their operating temperature is high enough to provide high-grade heat that can be used to increase overall system efficiency before it is rejected from the system. It should be realized, however, that all fuel cell systems, regardless of operating temperature, will probably not exceed 50–60 percent overall thermal efficiency; thus, a 100-kW system that has 50 percent efficiency must also get rid of 100 kW of heat. If high-grade heat, it can be used for heat-exchange purposes and thus increase overall system efficiency; with low-grade heat, however, this technique cannot be used effectively and the heat must simply be rejected.

The chief disadvantage is that molten-carbonate fuel cells present severe problems in corrosion and in high-temperature chemistry—problems that demand both basic understanding and good materials technology.

Solid-electrolyte cells. Cells employing solid electrolytes are based on the fact that as the temperature of certain solids is raised, an ionic species becomes mobile enough to carry reasonable electric currents through the solid. An example of such a cell system, now under development by Westinghouse Electric Corporation,[28] uses as an electrolyte stabilized zirconia (zirconium oxide containing 15-mole percent calcium oxide). The cell operates at 900–1000°C (1652–1832°F) and hence, like the molten-carbonate cell, is not subject to catalyst poisoning. It also has high-grade heat available with which to increase overall system efficiency. If the electrolyte resistance is to be kept low enough, the solid electrolyte must be thin (0.04 cm, or 0.016 inch); thus, structural fragility becomes a problem. Another major problem seems to lie in maintaining long-term bonding between the electrodes and this solid electrolyte.

Fuel cell fuels

Some practical considerations. If fuel cell systems are to have widespread use, they must run on low-cost, generally available fuels, such as gasoline and kerosene distillate (which includes kerosene, jet fuel, and diesel oil). Because of the almost universal availability of these fuels, we will call them "omnifuels." If the fuel cell system manufacturer wishes to create, make, and market a system for general use, he has two basic choices he can make with regard to fuel. He can develop either a system that will run on one or more of the omnifuels or one that may have considerably more power per unit weight and volume than the omnifuel system but which will require a fuel that is not generally available—e.g., hydrazine or pure hydrogen.

With the omnifuel system, the manufacturer will be taking advantage of an existing fuel distribution network that is practically worldwide. In addition, if it becomes necessary, his omnifuel system can run not only well but even better on such fuels as pure hydrogen or hydrazine. With the second system, if he wishes to tap broad markets with his product, he is faced with the problem of having to set up his own fuel distribution system. In addition, his system probably will not run on omnifuels without major modifications.

Also important from the customer point of view is fuel cost. Table I shows the range of prices involved. It can be seen that kerosene and diesel fuel are the lowest-priced omnifuels, unless we include gasoline bought under a tax exemption. It also can be seen why hydrazine-fueled systems will probably be limited to special uses.

Fuel preparation. It was pointed out earlier that indirect-oxidation hydrocarbon–air fuel cells systems require prior processing of the hydrocarbon fuel in order for it to be converted to hydrogen that the fuel cell can consume. Two fuel preparation processes can be used: steam reforming and partial oxidation. Both processes can be run at thermal efficiencies of 65 to 90 percent and produce usable hydrogen, but their approaches are significantly different, from the viewpoint of the fuel cell systems engineer.

In steam reforming, a widely used process in the petroleum industry,[29,30] steam and desulfurized hydrocarbon fuel are preheated and then reacted over a catalyst (usually nickel) at 538–871°C (1000–1600°F) and 100–200 pounds per square inch gauge. The reaction products for aviation jet fuel (JP-4), reformed at 650°C (1202°F) with three moles of water per carbon atom of fuel and then subjected to the water–gas shift reaction, are shown in Table II.* The steam-reforming reaction is highly endothermic; that is, large amounts of heat are required. This feature would seem to make this process a desirable subsystem to use in high-temperature fuel cell systems, where high-grade reject heat is available.

The catalyst used in this process can be poisoned by the sulfur that is usually present in hydrocarbon fuels. If the sulfur content exceeds 50 parts per million by weight, the reforming catalyst is usually so poisoned that the process breaks down. The result is that, to date, the steam reforming of omnifuels in a fuel cell system is not practical.

Partial oxidation is a process used to make hydrogen from hydrocarbon materials.[31] Basically, the process is just what it says, and the main products are hydrogen and carbon monoxide (CO), as contrasted with hydrogen and carbon dioxide (CO_2), the products of steam reforming, with shifting.

As used in industry to make hydrogen for ammonia-synthesis plants, partial oxidation is carried out by partially burning a hydrocarbon, such as diesel fuel, with pure oxygen at high pressures. As used in the Texas Instruments fuel cell system, partial oxidation is accomplished by partially burning a military fuel, such as JP-4 or diesel fuel, with air at a pressure only slightly above atmospheric (one inch of water above atmospheric pressure). The reaction product consists mainly of nitrogen (an inert diluent), hydrogen, carbon monoxide, and water, together with much smaller amounts of hydrogen sulfide, carbon dioxide, and some low-molecular-weight hydrocarbons. An example of such a product gas, made by partial oxidation of jet fuel (JP-4) with air, is also given in Table II. With such a product gas as fuel and using air as oxidant, the fuel cell module, shown in Fig. 2, has produced 40 watts dc (9.5 amperes at 4.3 volts).

It can be seen from Table II that there is a considerable difference in composition between the products of steam reforming and air–partial oxidation. This is shown even more clearly when the product compositions are compared on a dry basis. In this case, the steam reformate is about three times richer in hydrogen than the product of air–partial oxidation, mainly because of the diluent effect on the partial-oxidation product of the inert nitrogen, which results from the use of air instead of oxygen.

*Steam reforming of propane is represented by the equation: $C_3H_8 + 3H_2O = 3CO + 7H_2$. The water–gas shift reaction takes place at a lower temperature—e.g., 427°C (800°F)—than does steam reforming. Taking place over a catalyst (usually iron oxide), it may be represented thus: $3CO + 3H_2O = 3CO_2 + 3H_2$. The two processes—steam reforming and water–gas shifting, or shifting, as it is usually denoted—are almost always carried out sequentially.

II. Steam reforming and air–partial oxidation of JP-4 fuel

Component	Steam Reforming (650°C) Wet	Steam Reforming (650°C) Dry	Air–Partial Oxidation Wet	Air–Partial Oxidation Dry
Hydrogen	37.9	66.6	17.8	19.9
Nitrogen	...	...	50.0	55.8
Carbon monoxide	4.0	7.0	8.7	9.7
Methane	4.0	7.0	1.3	1.4
Carbon dioxide	11.0	19.3	11.8	13.2
Water	43.1	...	10.4	...

(Volume Percent)

For partial oxidation:
Feed rates are 0.522 pound per hour for JP-4, 0.538 pound per hour for water, and 46.3 standard cubic feet per hour for air.
Product rate is 730.0 standard cubic feet per hour.
Thermal efficiency of the process =

$$\frac{\text{Gross heating value of product}}{\text{Gross heating value of fuel}} \times 100 = 68\%$$

On both a water-free and nitrogen-free basis, the volume percent of hydrogen in the partial oxidation product would be about 68 volume percent. However, oxygen is hard to come by in normal field operations, whether military, industrial, or consumer. Hence, the fuel cell system developer must choose between obtaining a hydrogen-rich product with a high water content by a process that is sensitive to sulfur poisoning (i.e., steam reforming) or a nitrogen-rich product by a process that is free from catalyst-poisoning problems (i.e., air–partial oxidation).

Present situation in research, development, and technology

Availability of fuel cell systems. Progress in the hydrocarbon–air fuel cell field is at present limited to feasibility-demonstration systems and to laboratory test systems, ranging from 40 watts to several kilowatts. The emphasis is still on systems development. No one is yet offering for sale a line of hydrocarbon–air fuel cell systems, although one manufacturer has announced the possibility.[32]

Research and development. Programs are now under way in the United States, Great Britain, France, Netherlands, Germany, Japan, and the Soviet Union.

In the United States, the Army Mobility Command's U.S. Army Engineer Research and Development Laboratories at Fort Belvoir, Va., represent the focal point of the U.S. Army R & D effort in this area. This agency is both doing in-house work and funding basic studies and systems engineering development in direct- and indirect-oxidation hydrocarbon–air fuel cell systems. In addition, many previously mentioned private corporations have in-house funded programs on hydrocarbon–air systems. The American Gas Association is sponsoring a research program at the Institute of Gas Technology in Chicago.

In Europe, government-funded development programs for hydrocarbon–air systems exist in Great Britain, France, the Netherlands, and Russia. Production of natural gas from the recently discovered North Sea field has served as an added incentive to some of these programs. Corporate capital funding of hydrocarbon–air fuel cell systems is represented in England by Energy Conversion Ltd.; in Germany, by Battelle-Institut (Frankfurt am Main). A considerably larger fraction of the European work is devoted to molten-electrolyte cell systems than is the case in the United States. The status of the European programs is outlined in a recent summary.[33]

Systems engineering research. The main need at present in hydrocarbon–air fuel cell technology is in systems engineering. It has been finally recognized that the successful demonstration of a single fuel cell is only the first step in a series of increasingly complex system processes that must be optimized before one can even think of a practical fuel cell system. It is in this systems area that performance and reliability data are lacking. This gap must be filled and eventually will be, but it requires considerable time, money, and top-notch engineering talent.

The basic systems problem is to combine three subsystems into a single, efficient, and reliable system. The subsystems include (1) the hydrocarbon fuel-processing unit, (2) the fuel cell stack, and (3) the regulator–inverter–converter unit. Overall system efficiency is the product of the efficiencies of the three. Approximate efficiencies (power in/power out) being sought for these three subsystems are: fuel-processing unit, 80 percent; fuel cell stack, 50–60 percent*; and regulator–inverter–converter, 80 percent.

Major problem areas in fuel processing include (1) accurately metering small flow rates of liquid fuel—e.g., 1 to 3 lb/hr; (2) preventing deposition of gums and carbon that can clog small orifices; and (3) maintaining an even vaporization of fuel. These are tough problems in chemical engineering technology, but they do not require the acquisition of new, basic information.

The fuel cell stack, which is a series–parallel arrangement of modules, presents the problem of module equalization. This problem, which is well known to the battery industry, primarily means trying to bring every module making up the stack to the same power output. Variables involved can range from reproducibility of module fabrication to the extent of wetting of electrodes by the electrolyte. This particular problem represents one area in which more basic knowledge is needed, as is effective communication between the electrochemist and the system engineer.

Problems in regulator–inverter–converter subsystems, which are almost invariably solid-state, mainly involve increasing the efficiency while decreasing the weight. Although they demand ingenious solutions, they do not involve the need for new, basic knowledge.

Thus, we see that each of the three subsystems has its own problems. These are superposed on the overall system problem of integrating the subsystems into a practical hydrocarbon–air fuel cell system. Basic understanding is still needed in the area of the electrode–electrolyte interface; however, even in its absence, effective hydrocarbon–air fuel cell systems can, and will, be made. As a matter of fact, the building and testing of complete systems, even though imperfect, can facilitate obtaining these basic inputs by permitting them to be put in proper perspective—that is, by letting the system indicate the key areas in which increased understanding is essential.

Some possible early uses

First uses for a hydrocarbon–air fuel cell system will probably be in military or in military-related installations where remote power is needed. This need is now met by batteries and by gasoline- and diesel-fueled engine generators. For power needs to 100 watts, batteries are usually used; from 100 watts to 10 kW, gasoline-engine generators predominate; and from 10–300 kW, diesel-engine generators are commonly employed.

The following power system performance factors, either singly or in various combinations, can be critical to the success of certain military missions: silence, fuel flexibility, fuel economy, absence of RF interference, simple maintenance, and lack of catastrophic failure. The long-term hope of hydrocarbon–air fuel cell developers is that their systems will excel in all these performance characteristics. The potential for doing this exists; the major questions are when and at what cost.

There will be some military missions important enough to justify paying more for a power supply. Naturally, these first, special uses will not represent a large dollar market. They will, hopefully, result in a growing appreciation by military personnel at the field level of some of the advantages of this new power system. Then, as the initial cost of hydrocarbon–air fuel cell systems decreases, they will begin to be used in the broad spectrum of military power applications from 1.5 to 300 kW. Typical uses will include battery charging (1.5 kW), general lighting (3 to 5 kW), and powering missile installations (10 to 60 kW). In general, the volumes and weights of these systems should be comparable with those of engine-generator systems having the same output power.

With the increasing lifetime that will result from military uses of fuel cell systems, we can begin to look for penetration into the consumer market. The first uses will probably be the civilian analogs of the military applications discussed above. In other words, fuel cell systems will be used to furnish remote power, where a minimum of maintenance is desired. Examples of such installations are power units for radio repeater stations. Another remote-location application will be to furnish electricity for remotely located plants or villages to which the cost of running utility electricity would be prohibitively expensive.

At this stage, if their initial costs and operating expenses have been brought low enough, fuel cell systems will probably begin to be used to power off-the-road vehicles, such as bulldozers, graders, and excavators. In consumer uses of these vehicles, silhouette and weight are usually not critical factors; however, in military uses they can be, especially if the vehicles have to be airborne. Such vehicles would probably be driven by dc motors, one on each wheel. The high torque that is characteristic of dc motors, plus the ability to drive each wheel individually, could be advantageous factors in this type of vehicle. However, it should be pointed out that costs associated with present off-the-road vehicles are low enough to require that the fuel cell be at a highly developed and competitive stage of economics before this use can be considered seriously.

Another potential use almost invariably cited for fuel cell systems is to power the industrial, or fork-lift, truck. Here, the cell will be competing against present-day technology in gasoline engines and storage batteries. A cost comparison[34] made in 1960 by graduate students at the Harvard School of Business Administration shows that hydrocarbon–air cell systems will be competitive

*This assumes a fuel utilization (conversion of the electron content of the fuel to electricity) of 70–86 percent.

when fuel cell life reaches five years. This kind of operating life is not yet in sight. Beyond this, it is difficult to project uses to power vehicles with any degree of accuracy. With improving technology and decreasing costs, the application of fuel cells to power military tanks might become practical. A fuel-cell-powered tank would have certain tactical advantages, especially at night (in the absence of enemy radar and/or infrared-sensing systems).

Whether or not fuel cell technology and its economics will ever qualify the cell to power vehicles cannot now be stated with any reasonable amount of certainty. A good review of the factors to be considered here has been given by Giacoletto.[35] Commenting on the present status of internal combustion engines, he concludes: "In the case of automobiles and airplanes, the combination of large power/mass factors of the conversion equipment and large energy/mass factors of the fuel produces a highly effective overall system. Although the conversion efficiency of such a system is not very high, it is sufficient to make the system attractive for most mobile uses." (Giacoletto gives 16.7 per cent as the overall utility efficiency of a 120-hp automobile.) Considering these statements in the light of present hydrocarbon–air fuel cell technology, one concludes that, barring breakthroughs, if fuel cells are to be used to power vehicles, the changeover from the internal combustion engine will be a slow, evolutionary process. As such, it probably lies beyond 1980.

It should be noted that the potential uses discussed above are essentially substitutional in that in many cases the hydrocarbon–air fuel cell system will merely be replacing another power-generating unit. Although these situations may simplify the job of making fuel cell market projections, they by no means represent the only areas in which hydrocarbon–air fuel cell developers hope to market their systems. These developers argue, and probably rightly so, that here is a new kind of power system with a unique combination of characteristics that will enable it to create wholly new and presently unanticipated markets.

Competitive systems

In the power range below one kilowatt, batteries, windmills, and gasoline-fueled engine generators are now used as power sources. The choice of system or combination of systems depends on the kind of use. Batteries are widely used at the lower end of this power range. Automobile batteries are a classic example of this type of use of secondary (rechargeable) batteries, while transistor radio batteries represent a widespread use of primary (nonrechargeable) batteries. Another use often served by secondary batteries is to power remotely located radio repeater stations. In this use, the battery is usually kept charged by another power source. Gasoline or diesel engines are often used. In appropriate locations, windmills are effective. A propane-fueled thermoelectric generator has also been used successfully to charge the nickel–cadmium batteries that powered a remotely located 10-watt VHF-FM radio repeater station.[36] For very-long-term, remote, unattended operation, nuclear-isotope-heated thermoelectric generators seem practical. Examples include power supplies for ocean-bottom navigational aids or polar weather stations. Hydrocarbon–air fuel cells cannot compete here.

Among the energy conversion systems now being developed, both thermoelectric and thermionic generators will compete with hydrocarbon–air fuel cell systems, at least to the 250-watt level. With their high power density (watts per square foot), thermionic generators could conceivably be used to one kilowatt if an effective way to heat them should be developed. Both thermoelectric and thermionic generators are externally heated engines. With slight changes in their fuel-processing units, these systems could probably handle several omnifuels. Another omnifuel system being investigated for use in this end of the power spectrum is the steam-engine-powered generator.

The power range from one to 300 kW is now served mostly by engine generators. Gasoline-fueled engine generators predominate to 10 kW and diesel-fueled engines from 10 to 300 kW. It is these systems that hydrocarbon–air fuel cell systems will try to supplant. It will not be an easy job, despite some of the inherent advantages of the hydrocarbon–air fuel cell system. Engine-generator technology is far from stagnant, as evidenced by three advances announced in 1965. Westinghouse Electric Corporation introduced an engine generator,[37] using an engine developed by Curtiss-Wright, that is based on the Wankel principle. Westinghouse states that the specific weight (pounds per kilowatt) of this engine-generator system is much lighter than that of comparable diesel generators and nearly equal to that of turbine–generators. A multifuel engine, claimed to run equally well on gasoline, kerosene, and diesel oil, was recently described.[38] Field testing of a 3-kW Stirling engine-powered generator was announced.[39] The working fluid is a gas that moves in a closed cycle; air, nitrogen, helium, and hydrogen have been used. Since the working fluid is heated externally, the engine is an omnifuel system. It can also be made to operate at a low level of audibility.

Competitive developments in gas turbines and their uses are also being made. These systems have a high power-to-volume ratio. They are being used to drive electric generators in so-called "total energy" installations that involve on-site generation of power, together with efficient use of the rejected heart.[40] The gas industry is a strong advocate of this idea, which it would like to see developed to include also natural gas–air fuel cell systems, especially adapted for residences. These latter systems would hopefully free the homeowner from his present dependence on utility electricity. The economic and reliability aspects of this proposed use seem rather formidable at present. In addition, simple substitution of the fuel cell system for the utility lines neglects the innovative possibilities to be gained by taking advantage of the inherent characteristics of the hydrocarbon–air fuel cell system.

The fact that the hydrocarbon–air fuel cell produces high-current low-voltage dc electricity makes it especially attractive for potential use in the electrochemical industry, where low-voltage dc electricity is in high demand. At present, it is usually obtained by rectifying ac sources with mercury ignitrons or silicon rectifiers. The electrochemical industry is highly competitive, and both initial and operating costs of the electric power systems can be critical factors in the profit equation, especially in aluminum production. Hence, except for special-use situations, it will probably be many years before fuel cell systems will be competitive enough to gain wide acceptance in this industry.

Summary

Indirect-oxidation fuel cell systems are now in the systems engineering development stage. In these cell systems, the hydrocarbon fuel is first converted to an impure hydrogen, which may then be purified and injected into the fuel cell modules. The degree of purification of the hydrogen depends on the particular type of indirect-oxidation fuel cell being used, together with its sensitivity to certain impurities in the hydrocarbon-derived hydrogen. Temperature of cell operation can range from 60 to 1000°C (140 to 1832°F). Increasing cell operating temperature usually means cheaper electrode catalysts, if any, and fewer catalyst-poisoning problems. Along with these advantages come corrosion problems characteristic of high-temperature technology. Long-term, reliable operation of these complete systems is the chief goal now being sought. After that, reduction of system cost will be in order.

Direct-oxidation hydrocarbon–air fuel cells have not yet reached the systems engineering development stage. There are still some basic reaction problems to be solved at the fuel electrode. After that, a major effort will be needed to reduce the platinum content in this type of cell.

Initial use of hydrocarbon–air fuel cell systems will probably be in the military to furnish remote power in the range of 0.1 to 300 kW. The competitive factor being pushed here, but yet to be realized, is the significantly decreased and simpler maintenance inherent in fuel cell systems. An excellent summary of power system needs of the U.S. Army was recently given by Looft.[41]

After the technology has been refined in the military, these systems will start to move into consumer markets. Initially, their use will probably be the same as in the military—namely, to furnish remote-location power. With improving technology and decreasing costs, however, hydrocarbon–air fuel cell systems may be used to power off-the-road vehicles.

Applications beyond these will depend mainly on (1) improvements in technology and in system reliability, (2) cost reductions, and (3) proliferation of wholly new uses that will arise from the uniqueness of a combination of properties of this new kind of power system.

REFERENCES

1. Grove, W. R., "On voltaic series in combinations of gases by platinum," *Phil. Mag.*, vol. 14, pp. 127–130, 1839.

2. Grove, W. R., "On a gaseous voltaic battery," *Phil. Mag.*, vol. 21, pp. 417–420, 1842.

3. Rideal, E. K., and Evans, W. R., "The problem of fuel cells," *Trans. Faraday Soc.*, vol. 17, pp. 466–485, 1921–22.

4. Mond, L. L., and Langer, C., *Proc. Roy. Soc. (London)*, ser. A, vol. 46, pp. 296–304, 1890.

5. Baur, E., and Tobler, J. "Sammelreferat. Brennstoffketten," *Z. Elektrochem.*, vol. 39, pp. 169–180, 1933.

6. Allmand, A. J., and Ellingham, H. J. T., eds., *The Principles of Applied Electrochemistry*. New York: McKay, 1924, pt. 3, chap. 13, pp. 220, 226–227, 232–233.

7. Peattie, C. G., "A summary of practical fuel cell technology to 1963," *Proc. IEEE*, vol. 51, pp. 795–806, May 1963.

8. Grubb, W. T., and Michalske, C. J., "Electrochemical oxidation of methane in phosphoric acid fuel cells at 150°C," *Nature*, vol. 201, pp. 287–288, Jan. 18, 1964.

9. Williams, K. R., "The fuel cell for transport," *New Scientist*, vol. 8, pp. 338–340, Aug. 4, 1960.

10. Garner, P. J., and Williams, K. R., "Gaseous fuel cells," British Patent No. 844 584, Aug. 17, 1960.

11. Williams, K. R., Sutton, L., and Gregory, D. P., "Gaseous fuel cells," U.S. Patent No. 3 116 170, Dec. 31, 1963.

12. Kirkland, T. K., and Smoke, W. G., Jr., "5-kVA hydrocarbon–air fuel cell system tests," *Proc. 19th Ann. Power Sources Conf.*, pp. 26–32, 1965.

13. Engle, M. L., "A hydrocarbon–air fuel cell for military application," presented at 58th Nat'l AIChE Meeting, Philadelphia, Pa., Dec. 5–9, 1965.

14. Kordesch, K., "The hydrogen–oxygen (air) fuel cell with carbon electrodes," in *Fuel Cells*, G. J. Young, ed. New York: Reinhold, 1963.

15. Kordesch, K., "The 'Union Carbide' hydrogen–oxygen battery with carbon electrodes," in *Cold Combustion Fuel Cells*, E. W. Justi and A. W. Winsel, eds. Wiesbaden, Germany: Franz Steiner, 1962, pp. 353–360.

16. Kordesch, K., Raub, S. H., and Uline, L. J., "Fuel cell battery containing flat carbon electrodes," U.S. Patent No. 3 188 242, June 8, 1965.

17. Justi, E. W., and Winsel, A. W., *Cold Combustion Fuel Cells*. Wiesbaden, Germany: Franz Steiner, 1962, pp. 1–59.

18. Johnson, T. K., "Ion-exchange fuel batteries," *Proc. 18th Ann. Power Sources Conf.*, pp. 25–28, 1964.

19. Epperly, W. R., "Direct hydrocarbon fuel cell," *Proc. 19th Ann. Power Sources Conf.*, pp. 43–46, 1965.

20. White, E. R., and Maget, H. J. R., "Direct hydrocarbon fuel cell," *Proc. 19th Ann. Power Sources Conf.*, pp. 46–49, 1965.

21. Morril, C. C., "Apollo fuel cell system," *Proc. 19th Ann. Power Sources Conf.*, pp. 38–41, 1965.

22. "Fuel cell converts natural gas to electricity," *Elec. Eng.*, vol. 82, no. 9, p. 587, Sept. 1963.

23. Truitt, J. K., and Gray, F. L., "Development of the molten carbonate fuel cell system," presented at 58th Nat'l AIChE Meeting, Philadelphia, Pa., Dec. 5–9, 1965.

24. Tantram, A. D. S., paper presented at the American Chemical Society Symp. on Hydrocarbon–Air Fuel Cells, Atlantic City, N.J., Sept. 13–15, 1965.

25. Broers, G. H. J., and Schenke, M., "Investigation of high-temperature fused carbonate cells with paste electrolyte," *Comp. Rend., Journées Internationales d'Etude des Piles a Combustible*, pp. 54–63, 1965.

26. Lovy, P. "Pile à gaz monobloc à électrolyte fondu," *Comp. Rend., Journées Internationales d'Etude des Piles a Combustible*, pp. 82–87, 1965.

27. Millet, J., and Buvet, R., "Pile à électrolyte libre alimentée en air et gaz de reformage," *Comp. Rend., Journées Internationales d'Etude des Piles a Combustible*, pp. 49–54, 1965.

28. Archer, D. H., Sverdrup, E. F., and Zahradnik, R. L., "Coal-burning fuel cell systems," *Chem. Eng. Progr.*, vol. 60, pp. 64–68, June 1964.

29. James, G. R., "Which process best for producing hydrogen?," *Chem. Eng.*, vol. 67, pp. 161–166, 1960.

30. "Bigger bargains in hydrogen. Steam-reformed hydrogen is key to cheaper petrochemicals," *Chem. Week*, vol. 97, no. 26, pp. 18–22, Dec. 25, 1965.

31. Slater, W. L., and Dille, R. M., "Partial combustion of residual fuels," *Chem. Eng. Progr.*, vol. 61, pp. 85–89, Nov. 1965.

32. "Leesona mulls production of commercial fuel cells," *Wall Street J.*, Nov. 1, 1965.

33. *Comp. Rend., Journées Internationales d'Etude des Piles a Combustible*, Brussels, Belgium, June 21–24, 1965, Société d'Etudes, de Recherches et d'Applications pour l'Industrie, Brussels, 1965.

34. Adams, D. R., *et al.*, *Fuel Cells—Power for the Future.* Boston: Purnell Co., 1960, pp. 105–110.

35. Giacoletto, L. J., "Energy storage and conversion," *IEEE Spectrum*, vol. 2, pp. 95–103, Feb. 1965.

36. Marihart, D. J., and Strong, E. M., "Low-cost pole-mounted VHF radio station for Bonneville Power Administration," *IEEE Trans. on Power Apparatus and Systems*, vol. PAS-84, pp. 191–197, Mar. 1965.

37. "A new concept in portable engine generator systems," Surface Power Systems, Aerospace Electrical Division, Westinghouse Electric Corp., Lima, Ohio, 1965.

38. "Multifuel engine concept revealed by 2 engineers," *Dallas Morning News*, p. 13A, Jan. 16, 1966.

39. "Our Stirling engine: just add heat," Institutional adv., General Motors Research Labs., *IEEE Spectrum*, vol. 2, p. 88, May 1965.

40. "The power company's new rival," *Business Week*, pp. 80–83, June 13, 1964.

41. Looft, D. J., "Electric power needs reviewed for advances in weaponry," *Army Res. Develop. News Mag.*, vol. 7, no. 1, pp. 38–45, Dec. 1965–Jan. 1966.

The contourograph

Already a useful tool for compressing masses of physiological data into compact form for simplified visual review, the contourograph should also find wide application in nonmedical studies

George N. Webb *The Johns Hopkins University*
Richard E. Rogers *The University of Miami*

Utilization of the three-dimensional aspects of the cathode-ray tube—vertical signal deflection, flexible time base, and intensity modulation—makes it possible to organize certain types of data into a meaningful format, greatly simplifying the interpretation of these data. This article describes the basic principles of organizing and displaying data in the contourographic form, indicates several degrees of complexity in providing the instrumentation for this method, and presents a number of contourograms illustrating various applications.

The contourogram[1] is a display of nonidentical, quasi-periodic data in a compact form. It is usually a time-exposure photograph of a number of oscilloscope traces. The data are organized about an event common to each cycle, in order to emphasize and present for study the similarities or dissimilarities between cycles, collected over extended periods of time. The device that produces this display is called a contourograph.

Many physiological signals fall within the general class we are calling nonidentical quasi-periodic, and serve to illustrate the contourographic technique. Many other data that have similar characteristics could be usefully displayed in the contourographic format.

An example of the contourographic form of display is given in Fig. 1. The electrocardiogram (EKG) is quasi-periodic in that the cycle repeats itself, varying normally

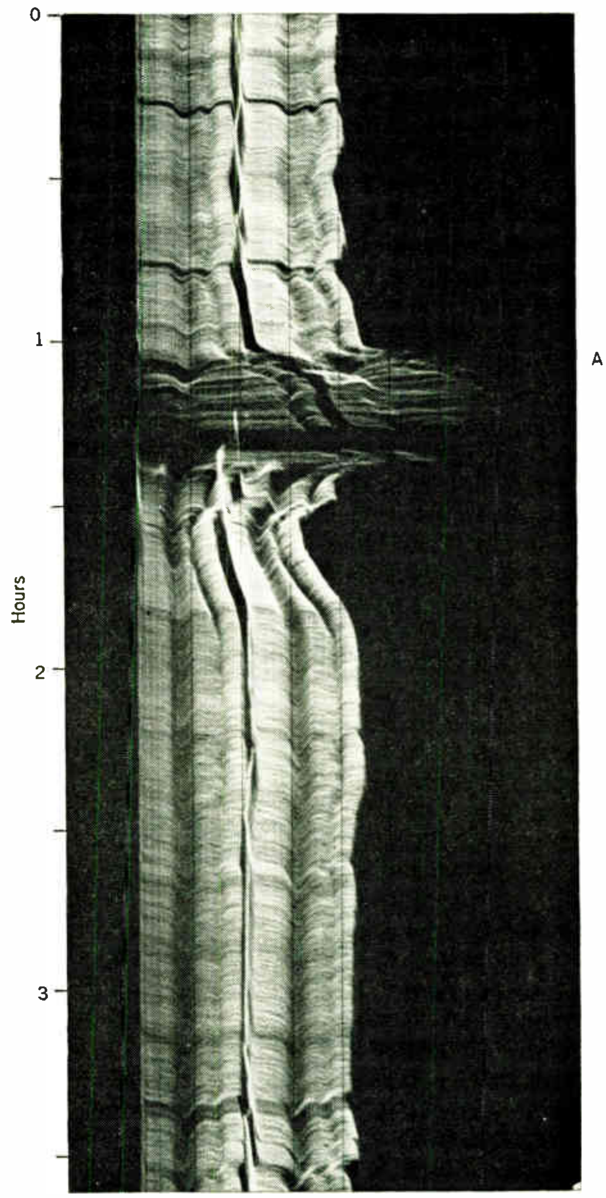


Fig. 1. Contourogram of an EKG of a newborn infant. Linear horizontal sweep, left to right, is triggered by alternate R waves of the EKG. Vertical signal deflection and intensity are proportional to the amplitude of the EKG. Vertical linear sweep, top to bottom, is produced by a moving film. Note the drastic slowing of the heart rate and waveform change at A. The blank area indicates lack of triggers, not necessarily lack of heartbeat. Also note the uniformity with which the heart rate is maintained or modified.

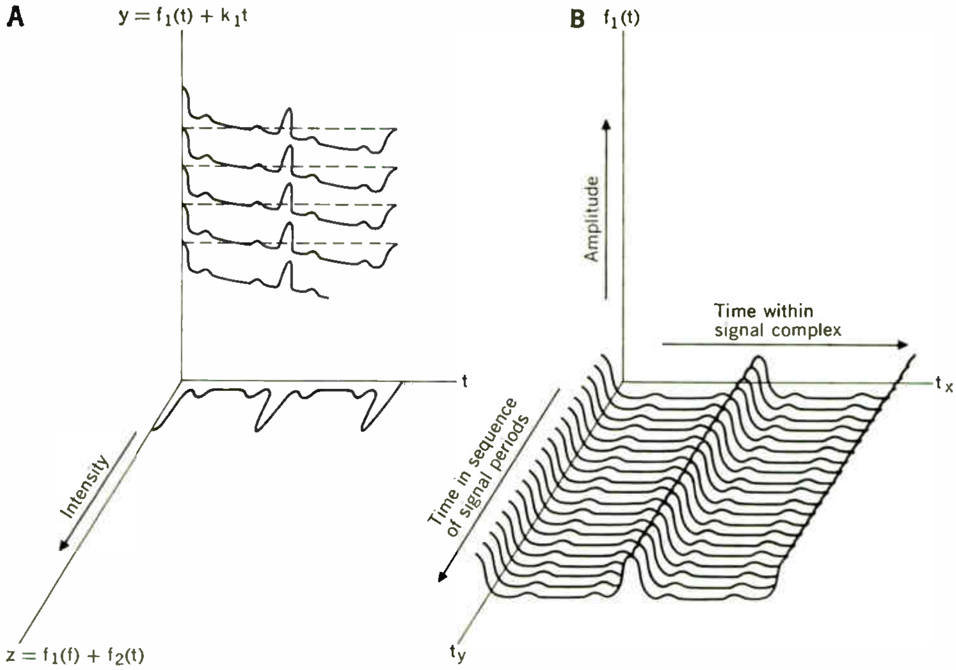


Fig. 2. Generation of the contourogram. A—Electrical signal applied to the oscilloscope, showing that the linear horizontal trace is triggered on alternate R-wave epochs. B—Pseudo-three-dimensional display as it appears to the viewer.

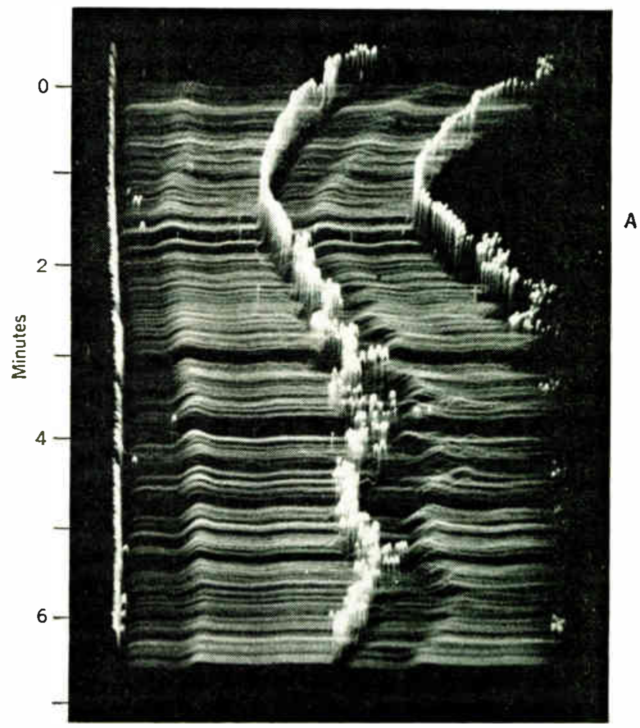


Fig. 3 (below right). EKG contourogram. formed as a raster, using a stationary film and a linear vertical sweep in addition to the vertical signal deflection. Heart rate increase at A is the result of exercise. Note depression of the waveform following R wave after exercise.

from 70 to 180 times a minute; however, it seldom has many sequentially identical waveforms. Here the contourograph organizes and presents the data so that the beat-to-beat variations (both in periodicity and waveform), as well as the long-term trends, may be seen.

In general, the cathode-ray oscilloscope (CRO) has been used in the laboratory in the same mode as a pen or galvanometer recorder, merely operating faster. The writers believe that intensity modulation of the cathode-ray tube, as used in television and radar, can also be used to advantage in the laboratory. Through exploitation of this "third-dimensional" aspect of the cathode-ray tube (CRT), the utility of the oscilloscope should be further advanced.

The basic idea is simple and requires only a small departure from the normal use of the CRO. However, a whole new viewing dimension is created which proves to be very useful. The contourograph creates a signal-controlled raster display, which has a three-dimensional appearance. The signal applied to the y axis is the sum of a time-varying input signal to be displayed, $f_1(t)$, and one linear with time. Thus the vertical deflection signal has the form $y = f_1(t) + k_1t$. The electrical signal applied to the x-axis input increases linearly with time. The signal applied to the intensity axis is of the form $z = f_1(t) + f_2(t)$, where $f_2(t)$ may be a function of the input signal (its derivative, for example) or may be another function which is time-related to the input signal; see Fig. 2(A). The display that results from the application of these signals gives the viewer the impression of a spatial representation of two time axes and an amplitude axis; see Fig. 2(B).

Principles of the contourograph

Signals that may usefully be shown in this form are wave complexes that (1) recur, either regularly or with periodicities not varying by more than a ratio of about five to one, (2) have a common event in each complex from which an organizing trigger can be derived, and (3) have variations in periodicity or waveform from which useful information may be derived. The complete contourograph includes the following characteristics:

1. The CRT horizontal linear sweep is triggered by the common event (epoch).
2. This linear sweep may be retriggered even if the beam has not traversed the full CRT screen width.
3. Normal vertical deflection of the primary signal to be displayed is provided.
4. A vertical deflection, linear with time or number of sweeps, is added to the normal vertical deflection (3) to form the raster.
5. Intensity control (level and contrast) is made proportional to one or more signals, primarily to

(Continued on page 83)

the amplitude of the vertical deflection signal (3).

6. Provision is made for simultaneously viewing a group of sequential sweeps.

Many of the foregoing functions are provided in any normal oscilloscope. The opposite of the second function is generally found; that is, the sweep circuit has features that prevent retriggering until the beam has completed a trace across the entire tube face. Although the contourograph can operate in this mode, visual enhancement and gain in information in some cases is provided with retrace on demand.

The greatest modification that must be made to a normal oscilloscope to make it into a contourograph is in the intensity circuit. Our application necessitates careful control of the beam current, signal mixing, frequency response, and contrast.

Linear vertical deflection can be obtained by use of a slow linear sweep added to the vertical deflection amplifiers or created in the recording process by a moving film.

Contourograph operations

Contourography requires the simultaneous viewing of many sweeps. This is made possible by a time-exposure photograph of a CRT display, a cathode-ray storage tube, or a CRT phosphor having a persistence long enough to show one full raster scan. With the slower physiological signals few phosphors are available to achieve this persistence, particularly for real-time display. A bistable cathode-ray storage tube oscilloscope (Tektronics 564) has been tested, but the packing density is limited due to bleeding of the image.

Vertical deflection. The simplest form of the contourograph is obtained by adding a linear vertical sweep to the normal vertical signal deflection. Since most oscilloscopes have differential amplifiers, at least in the final stages, only a relatively simple modification is necessary to add this sweep. For example, the vertical signal may be derived from a motor-driven continuous-rotation potentiometer.

Two forms of potentiometers have been used. The first one makes a symmetrical triangular sweep, which rises and falls linearly. This method is useful for real-time viewing with long-persistence phosphors and in cases in which the vertical retrace cannot be synchronized with the horizontal sweep. The second, a more conventional potentiometer, provides a sawtooth with linearly rising voltage and quick retrace. It can be synchronized with a single-frame automatic photographic system for producing nearly continuous contourograms, as shown in Fig. 3. These motor-driven vertical sweeps can also be useful for continuous monitoring, to avoid stationary patterns and to minimize burning of the phosphor.

Intensity. The next feature to be added is the intensity controls. Few oscilloscopes provide convenient means for accurately controlling the intensity of the trace, chiefly because the demands for this feature have been few and because the high negative potential between ground and grid–cathode circuit of the electrostatically deflected CRT does not lend itself readily to precise application of the necessary grid–cathode signals required for intensity modulation. The magnetically deflected CRT operates the grid–cathode circuit within 100 volts to ground; hence grid–cathode signals can be easily applied. However, the cost of the circuits for driving the magnetic deflection coils for the arbitrary deflection signals encountered in oscilloscope applications has restricted the use of this tube to special applications. The advantages to be gained for contourography by use of magnetically deflected tubes—small spot size, greater choice of precision tube force, ease of accurately controlling intensity modulation—may be worth the added cost involved in the higher accelerating potentials and in magnetic-deflection drive circuits.

The Tektronix 561A oscilloscope presently used for the contourograph display has two regulated high-voltage supplies, one for the cathode and one for the grid. There is a floating supply to the grid, and the sweep retrace blanking signal is normally coupled to the grid by a signal near ground, moving the whole power supply. Thus, accurate steady-state values may be applied to the grid by driving this power supply from the contourograph. For the sweep rates that are used for the EKG contourograph, intensity-modulation signals of about 4 volts are needed.

It is sometimes useful to have facilities for mixing several signals with individual level and contrast controls; for instance, it may be desirable to add a derivative of the input signal to the intensity circuit. In this case the positive slope of either negative or positive signal spikes would be intensified and could readily be detected as bright spots.

Figures 4 and 5 illustrate the addition of a second intensity signal, which is neither the vertical deflection display nor is it derived from it. The first, Fig. 4, is a physiological signal (heart sounds) during a stress test. The marker pips on the left indicate the trigger derived from the *R* wave of the EKG and show the movement of the sounds with respect to the *R* wave. Figure 5 shows another physiological signal (myogram), which is the electrical response of a muscle produced by an electrical stimulus. The stimulus triggers the horizontal sweep and appears as the bright line at the center and right.

Horizontal deflection. Modern oscilloscopes have provision for triggering the horizontal sweep from either the internal vertical signal or an external signal. Also usually provided are hold-off circuits, which prevent a second trigger from restarting the trace until it has crossed the face of the tube. This latter feature is acceptable for the contourograph only if there is an orderly progression between complexes to be shown. However, with an arrhythmic EKG, as in Fig. 6, it may be necessary to show each and every beat in its entirety. Thus the next degree of complexity calls for retriggering the sweep before it has completely crossed the CRT face. Because of the nature of the signal, trigger hold-off may be required before a second trigger is accepted, but this will probably be about the first ten percent of the trace, or slightly less than the shortest period to be expected.

In some instances the start of a cycle of the input waveform may not correspond to the portion of the waveform from which a trigger is most easily derived. This is the case with the EKG signal which has been illustrated. The trigger is derived from the largest deflection, the *R* wave; however the cycle of interest starts before the *R* wave and extends past it. To display this cycle properly the *R* wave must appear near the center of the screen, as drawn in Fig. 2. Thus the sweep runs for two *R*-to-*R* intervals and is started on alternate *R*-wave triggers.

Figure 6 illustrates a type of irregular rhythm that may be studied more easily if each complex, not just alternate waves, is shown on the left and in the middle. For this, Fig. 6(B), two interlaced sweeps are employed, each sweep still running for two *R*-to-*R* intervals with one or the other sweep starting on every *R*-wave trigger.

signal introduced at this point drives the CRT grid power supply with respect to ground and provides accurate intensity control.

Applications of the contourograph

There are three major areas of usefulness for contourographic display. First, it provides a tractable means for quickly reviewing masses of data to find regions of interest that require more detailed study. Of course, this application implies that the data from which the contourogram was made are in storage (on magnetic tape, for instance) and can be used for the detailed study. This application is particularly adaptable to physiological data obtained during continuous monitoring of hospital patients. A person with a fast heart rate of 150 beats per minute will have approximately a quarter of a million beats per day. In this type of monitoring, hours may go by before a significant trend or variation occurs, hence some means for quickly scanning the data is extremely important. Figure 8 shows continuous records, each made over an eight-hour period, from a portable magnetic tape recorder[2] and an electrocardiograph amplifier carried by the subject. Records such as these can be scanned by the eye rather rapidly to find areas of interest.

The second advantage of this display accrues from the time pattern of waveform complexes that is formed. Whereas the individual waveform is read by the physician for significant detail, this overall pattern can be read for detail in relation to the environmental circumstances under which it was taken.

Finally, this method of displaying nonidentical data may be helpful for teaching purposes. Well-known characteristics can be shown dramatically and perhaps better remembered than from word descriptions. It is commonly known, for example, that a pulse wave has finite velocity as it travels down a blood vessel. This phenomenon is shown quite well in Fig. 9, which shows the result when a pressure transducer is pulled back through an artery. The pressure waveform is displayed in the contourograph, but the trigger, indicated by the intensity pips, was obtained from the electrical signal from the heart. Withdrawing the transducer results in an increase in the propagation delay between the pulse wave and electrical signal representing the onset of contraction.

Combining two or more signals in the contourogram shows correlations that would otherwise be missed or seen less clearly. Although it is possible to present several variables on multitrace recorders, and from them determine correlations, the near superposition of the contourogram may show them more easily (Fig. 10).

Fig. 8. Two eight-hour EKG contourograms (A–B and C–D), obtained by means of a portable tape recorder from subjects engaged in normal daily activities. Note the episodes of drastic heart-rate and waveform change lasting for about an hour in the early part of record A. Waveform and rate changes continue throughout the eight-hour period of record C–D. (These records were obtained through the courtesy of Dr. B. Tabatznik, Sinai Hospital, Baltimore, Md.)

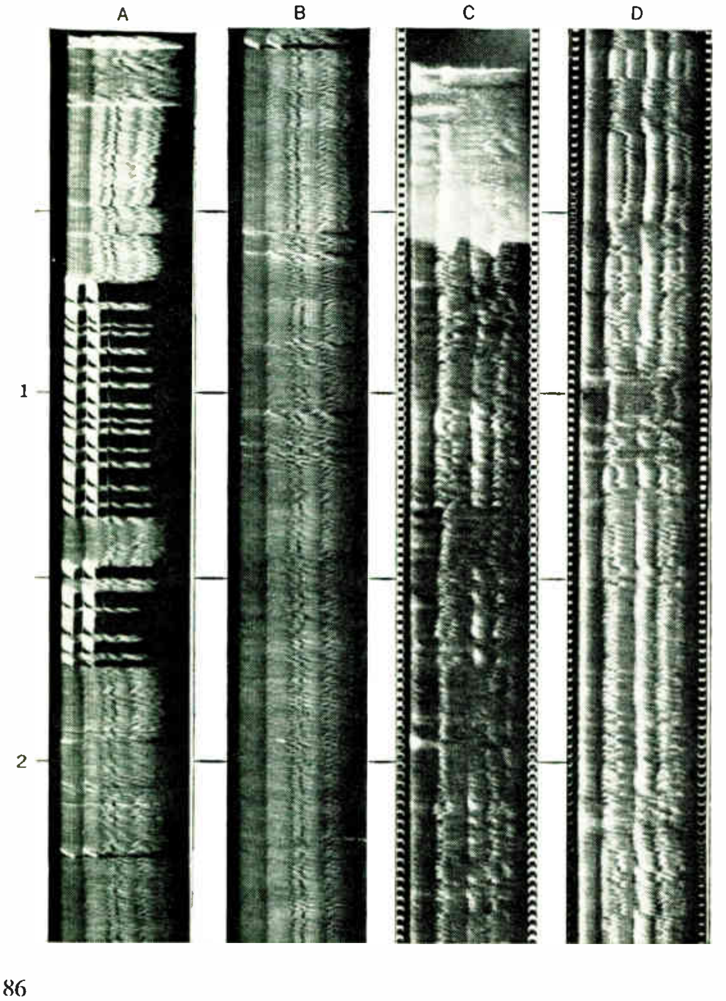
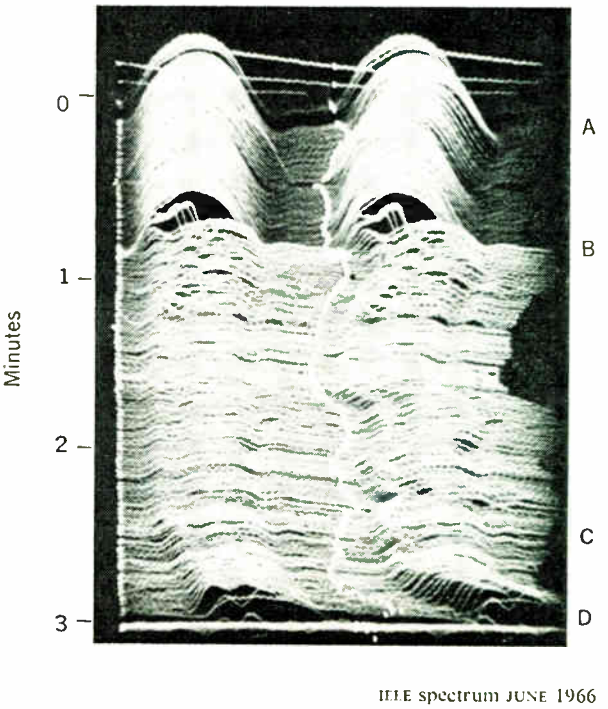


Fig. 9. Intracardiac pressure contourogram, in which the transducer is initially in the left ventricle (A–B), is withdrawn into the aorta (B), and then moved along the aorta away from the heart (C–D). Each sweep is triggered from an R wave of the electrocardiograph. The R-wave trigger pulses are shown as the bright marks at the left of the picture as well as down the center. Note pulse propagation delay and gradual increase in pulse amplitude as transducer moves along aorta away from the heart (C–D).



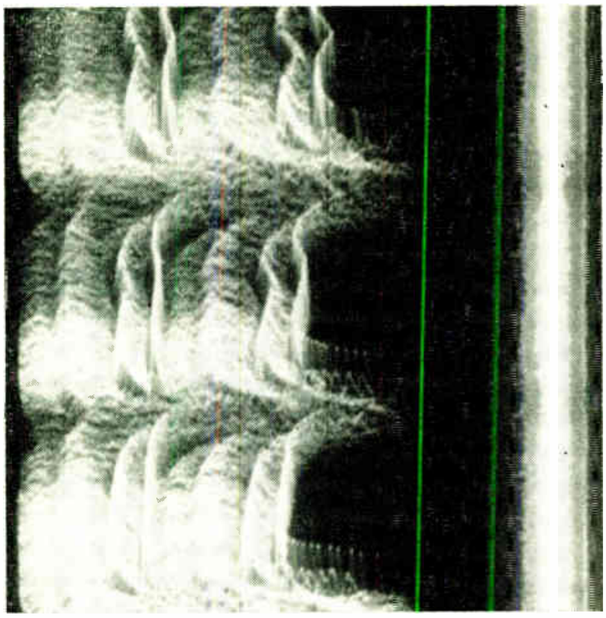
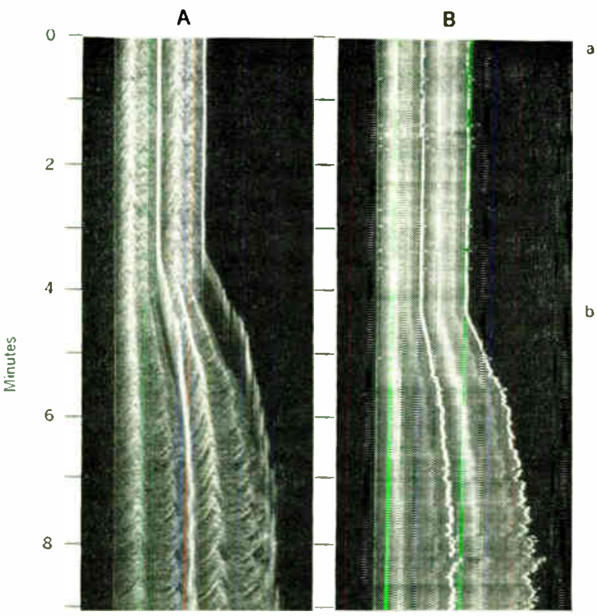


Fig. 10. Fetal EKG with the intra-uterine pressure signal superimposed on the intensity axis. Drastic slowing of the fetal rate is associated with each uterine contraction (bright areas). Trace on right is EKG presented as an amplitude (right to left) vs. time display and illustrates the minimal amount of information shown when the same packing density is used for the contourogram. (The data for the contourogram were supplied by Dr. Edward Hon, Yale Medical School, New Haven, Conn.)

Fig. 11. Contourogram modified for recording systems that produce only intensity changes, linear horizontal sweep, and linear vertical movement, such as facsimile recorders for wire photo transmission. A—Conventional contourogram. B—Duplication of A except that no vertical signal deflection was used. Note very uniform heart rate during exercise (a–b) and gradual slowing of the rate following exercise. The contourogram on the right contains more information than a record of heart rate or R-to-R interval time but not as much information as the full contourogram does. Primary advantage of format in B is that it can be produced on instruments that provide immediate copies without the need for photographic processing. Triggering artifacts are evident in B between a and b.



Other contourograph forms. For some cases of monitoring physiological data in real time, immediate records in contourogram form are needed; the situation cannot wait for a full photographic development. A number of instruments can scan a horizontal trace linearly showing intensity pattern without vertical deflection. Figure 11(B) illustrates this type of record as recorded from the oscilloscope. The only thing that has been eliminated here is the vertical deflection. It can be seen that although some information that is contained in Fig. 11(A) is lacking in this display, there is still much more information than a simple rate meter tracing would show. Modifications in waveform detail can be seen, although they cannot be accurately read.

Data handling for the contourograph. In general, the signals we have applied to the contourograph have come from magnetic tape recordings, but we have also made contourographs in real time.

The data from which most of the illustrations were made came from FM analog recordings made at speeds of $^{15}/_{32}$, 1⅞, and 7½ inches per second. These tapes were played back at speeds from 7½ to 60 inches per second, thus giving a maximum speed-up of 128 times. Figure 10 was made from magnetic tape directly recorded at 7½ inches per minute and played back at 3¾ inches per second.

The sweeps on the contourograph have been made to accept speed-up time steps of 1, 4, 8, 16, 32, and 128, giving the same rate calibration for any contourograph. When other speed-change ratios are used, the closest available range is employed, resulting in a noncalibrated picture.

The optimum packing density for our signals, created by the moving-film rate, seems to be of the order of 8 inches of film per hour of data. Since it is large enough to view directly in negative form or to make contact positives, 70-mm film has been found to be most satisfactory.

Conclusions

The development of the contourograph has provided a new tool for the compression of masses of nonidentical, quasi-periodic data into a compact form for easy visual review, organization of data into meaningful patterns, and illustration of complex relationships. This display method has already found a place in presenting physiological data gathered over extended periods of time. Other areas in which nonidentical periodic signals might well be presented in contourographic form include electrical phase-shift studies; mechanical reactions to period impact stress, either linear or rotational; combustion-engine pressures during various load conditions; weather conditions over long periods; and spectral-photometric analysis of chemical reactions occurring over long periods.

Modifications of modern oscilloscopes for this type of presentation are relatively minor and could become one of the standard plug-in options. The full capacity of X, Y, and Z modalities of the oscilloscope for scientific work has not been exploited.

The work described in this article was supported by grants HE-02923 and GM-10895 from the National Institutes of Health.

REFERENCES

1. Webb, G. N., "The contourogram," *Bull. Johns Hopkins Hospital*, vol. 116, pp. 221–228, Apr. 1965.

2. Holter, N. J., "New method for heart studies," *Science*, vol. 134, pp. 1214–1220, Oct. 21, 1961.

Progress in ionized-argon lasers

Many gas lasers have been discovered during the five-year span between 1961 and 1966, but it would appear that the argon laser holds the brightest possibilities for practical applications in many diverse and interdisciplinary fields

Roy A. Paananen *Raytheon Company*

During the past five years, the development of gas lasers, as measured in maximum available, continuously operating optical output power, has increased at the rate of about ten times per year. This growth, however, has not occurred in only one type of gas laser—four different wavelengths, in three gaseous media, have been responsible for maintaining this growth envelope. The author discusses the discovery of ionized noble gas lasers and their modus operandi, and presents the standard operation of ~ one-watt units to exemplify most of the salient characteristics of this kind of laser. The nature of the problems involved, and their tentative solutions in the quest for higher argon laser power (5–100 watts), is the subject of the main section of this article. This is followed by a few remarks on present and predicted applications for the interesting—and important—argon laser.

Within a particular species of gas lasers, the chief growth factors, as shown in Fig. 1, have been volume enhancement (note the substantial lengths indicated at the top of each curve) and technical refinement. In the two He–Ne wavelengths, this latter factor was predominantly responsible for the hundredfold growth advantage of the important 6328 Å line over its longer wavelength counterpart. Further advancement in the CW output power of either of the He–Ne lasers is not expected.

In contrast, both the ionized-argon (Ar^{II}) and the CO_2 lasers are the subjects of considerable contemporary interest with regard to their ultimate power potentialities. Although the initial CW laser phenomenon in argon occurred at the relatively high level of 160 mW, it apparently "peaked out" after only two years of growth. It is felt, however, that the ceiling is only temporary; the immediate hiatus is associated with the rapidly increasing difficulty of additional advances. Also, full advantage has not yet been taken of volume enhancement in ionized argon.

Lasers operating between excited levels in ionized—rather than neutral—gases are characterized by their extremely high current density at laser thresholds, accompanied by the need for maintaining some fractional ionization of the total gas population. This energy expenditure is not inconsiderable and it requires a new technology to cope with the unusual thermal and erosive discharge properties thereby engendered. Without compensatory advantages, it was thought that this drawback would relegate CW ion lasers to a minor role in the overall gas laser evolutionary cycle. But fortunately, this is not so: ionized noble gases—and argon, in particular—open up new and previously unavailable portions of the visible spectrum to continuously operating laser coverage. The ten ionized-argon laser wavelengths fall into the violet, blue, and green portions of the chromatic stimulus, with complementary coverage being afforded in the red and yellow-green by means of ionized krypton. Extensions of CW laser coverage farther into the actinic spectral bands by means of Ne^{II} are now a reality. In at least five of these wavelengths, one can readily generate optical powers of the order of one watt, so the spectral "blanket" is energetic as well as diverse.

Although four of the five nonradioactive noble gases (Ne, Ar, Kr, and Xe) show ionized laser action, argon remains the most efficient.

Discovery and prelude

More than any other single factor, the increased utilization of heavy-current-pulsed gas-discharge techniques by laser researchers led to the first observation of laser action between excited states in gaseous ions. This was reported by Bell,[1] using singly ionized mercury (Hg^{II}), early in 1964.

Prior to 1964, many investigations of possible gaseous laser media were conducted with continuously operating discharges in which excitation for maintaining the column came from an RF or dc source. Generally, as each new laser transition was observed, it was comfortably within energetic and thermal capabilities of both the source and laser for the determination of the power saturation point, without much incentive to explore beyond this. Essentially then, the CW searchers tended to uncover lines with which they could easily cope.

The work of Mathias and Parker,[2] and Boot, Clunie, and Thorn,[3] as reported about mid-1963, gave us our initial realization of the utility of heavy-current-pulsed discharges in establishing laser action in a previously intrac-

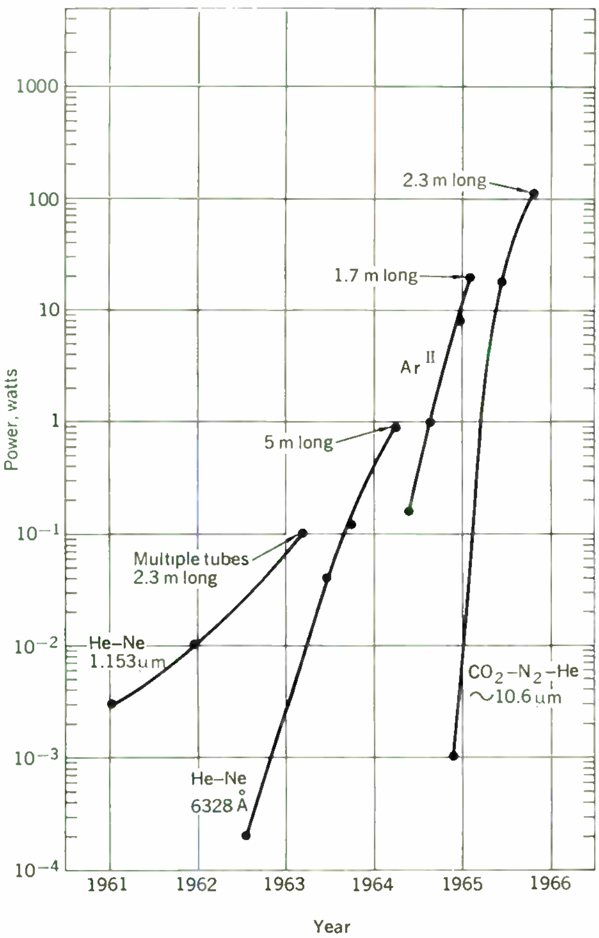


Fig. 1. Gas laser power growth during past five years.

table gas (N_2) species; or, in exploring altogether new pressure regimes in a known (He–Ne) laser plasma. Although these efforts certainly did not produce the first pulsed laser discharges (this had been done frequently for diagnostic purposes), the work of these five scientists was unique in the peak power employed—0.75 to 1.0 MW.

An appraisal of the work of Mathias and Boot, and others who followed, reveals three distinctive attributes of a high-current, pulsed-discharge system for gaseous media:

The discharge currents and power inputs applicable to laser tubes and cathodes in low-duty-cycle pulse systems can be orders of magnitude higher than those allowable in continuous service. For example, peak currents up to 2000 amperes have been reported. Simple structures such as air-cooled quartz laser tubes and oxide-coated cathodes lend themselves readily to all but the very extreme pulsed service. Cold cathodes can be employed for gases that are not compatible with the poisoning characteristics of oxide-coated cathodes.

A *principle of exclusiveness* applies; the pulse method will demonstrate laser action on some transitions that are inherently transient in character (i.e., they are permanently excluded from operating in a continuous manner). Bennett[4] gives a simple example in this class, whereby short-lived population inversions can occur in certain two-level systems—characteristically those bearing metastable lower levels. The inversions can be maintained for the short time interval T so that $T \lesssim \frac{1}{2} A_{21}$, where A_{21} is the Einstein A coefficient for the laser transition in question. Typical T times are 10^{-7} to 10^{-8} second, and the excitation pulse rise time must be even shorter than this to show any inversion.

In other instances, laser action will be achieved transiently at pulse turnoff, in the so-called "afterglow" region. This is particularly true for binary gas mixtures, in which selective upper-level excitation can continue for 10^{-5} to 10^{-4} second through collisions of the second kind with metastable levels of the buffer gas. The more rapid decay of the electron population then deactivates the lower level and allows a short-lived inversion to be established.

A *principle of inclusiveness* applies, since it is likely that the pulse technique will pick up all lines subject to laser action whether or not further investigation shows them to be inherently transient or continuously operable. This is particularly true if the pulse set is flexible enough to operate both with short pulses (10^{-8} to 10^{-7} second) and long pulses (10^{-4} to 10^{-3} second), in which case a preliminary sorting of the lines can be made. Use of millisecond-long pulses is sometimes termed *quasi-CW operation*, since the interval is greater than the time constants of any reasonable atomic process.

Within months after the first success in ionic laser action, announcements came from three independent groups that reported pulsed-laser action in singly ionized argon. The coincidence in the laser species probably occurred from the extensive employment of argon (as well as He) as a buffer gas in mercury studies. The Ar^{II} gain coefficients are now known to be large enough to facilitate laser action even under quite unfavorable conditions; He^{II} laser action, however, has not yet been achieved.

In strict terms of publication priority, some recognition must be given to the work of Convert, Armand, and Martinot-Lagarde,[5] who observed one blue laser transition in a mercury vapor–argon gas mixture, but they failed to distinguish the source. It remained for Bridges,[6] with a publication date only days later, to provide a comprehensive list of ten visible-region laser lines that arose from discharges in pure and buffered argon, along with their tentative energy-level assignments and some estimates of gain. Subsequently, Bennett *et al.*[7] published data on pulsed-gain measurements for seven of the argon lines, as well as power output information in the quasi-CW regime. For two of the transitions, the long-pulse data seemed particularly attractive; Bennett achieved one watt at 4880 Å, and approximately 10 watts at 5145 Å.

Ordinarily, one might expect follow-up work in the CW direction to proceed cautiously in extending the pulse lengths, uncovering thermal problem areas, etc. But in the spring of 1964, this was not the case. True CW operation at the approximate power of 0.1 watt was reported in a paper contemporary with that of Bennett. Gordon, Labuda, and Bridges[8] had also recognized the continuously operating possibilities of the ionized-argon transitions, and, using a new and appropriate technology, they were able to operate all ten of them in a CW fashion. Indeed, Gordon's paper announced CW laser action (but gave no details) at many of the 20 additional visible wavelengths observed in krypton and xenon (for which wavelength listings had not then been published).

Aside from the obvious implications of achieving energetic-pulsed and continuous laser action over many wavelengths throughout the visible spectrum (especially into the blue), we must appreciate the fact that the work of Bridges, Bennett, and Gordon was carried out within the family of inert gases. The virtues of the four members of this group (excluding He and radioactive radon) as ionized laser media are not always given full cognizance.

A short summary of their useful properties reveals that

1. They are fairly inexpensive and readily available in rather pure form.
2. Ne, Ar, and Kr are easily trapped for impurities by use of liquid N_2; Xe can be trapped with solid CO_2.
3. They do not react with oxide-coated cathodes, nor with getters.
4. They are nonpoisonous.
5. They all show ionized laser action by themselves, although mixtures are often employed for certain reasons.
6. No laser wall temperature adjustment is necessary, although experimentation in this variable is sometimes found to be useful.

Ionic laser action in the inert gases—and argon in particular—is important, not only because of the intrinsically favorable factors such as continuous operation, multiplicity of strong lines in the visible spectrum, etc., but also because the medium itself is ideally tractable within our immediate technology.

Pulsed, ionized-gas laser investigations continued almost unabated even after the CW breakthrough. The field was so open that some ionic laser transitions were discovered even with CW lasers[9] in spite (or because) of their discharge current handicap.

The short wavelength spectrum limit, at present, is well into the ultraviolet at 2358 Å, caused by excited levels in triple-ionized neon.[10] The laser threshold here, in a 4-mm-diameter tube, is 1200 amperes.

Presumably, pulsed-ion laser researchers are presently giving some thought to the enormously difficult task of helium lasers and to the question of its theoretical possibility. Their efforts may also encompass the field of metal vapors (other than mercury) wherein, for example, Fowles and Silfvast[11] achieved pulsed laser action in neutral lead vapor.

The advancement and partial development of continuously operating ionized-argon lasers remained within the framework established by Gordon *et al.* for almost a year. The one-watt milestone, in argon, was attained by the end of August 1964, using simple laser volume enlargement as well as benefits that accrued from the subsequently discovered magnetic field effect. While the refinement of the one-watt lasers continued until they became an ordinary article of commerce, similar and larger structures for the achievement of 5–20-watt continuous operation revealed operating deficiencies in several important aspects. These problem areas and their tentative solutions are treated in considerable detail in a later section of this article.

Inversion mechanism in the argon-ion laser

The excitation of ionic levels in noble gas capillary discharges is caused basically by electron impact. Although the question—impact on what?—has been a subject of controversy, recent unpublished data from measurements made at Bell Telephone Laboratories have con-

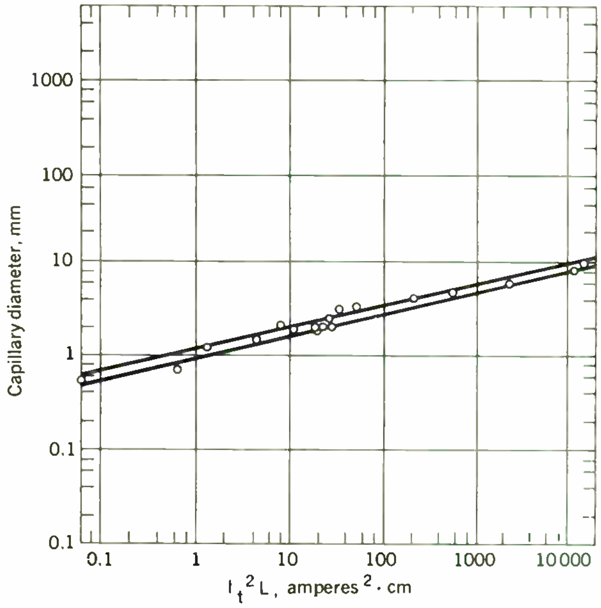
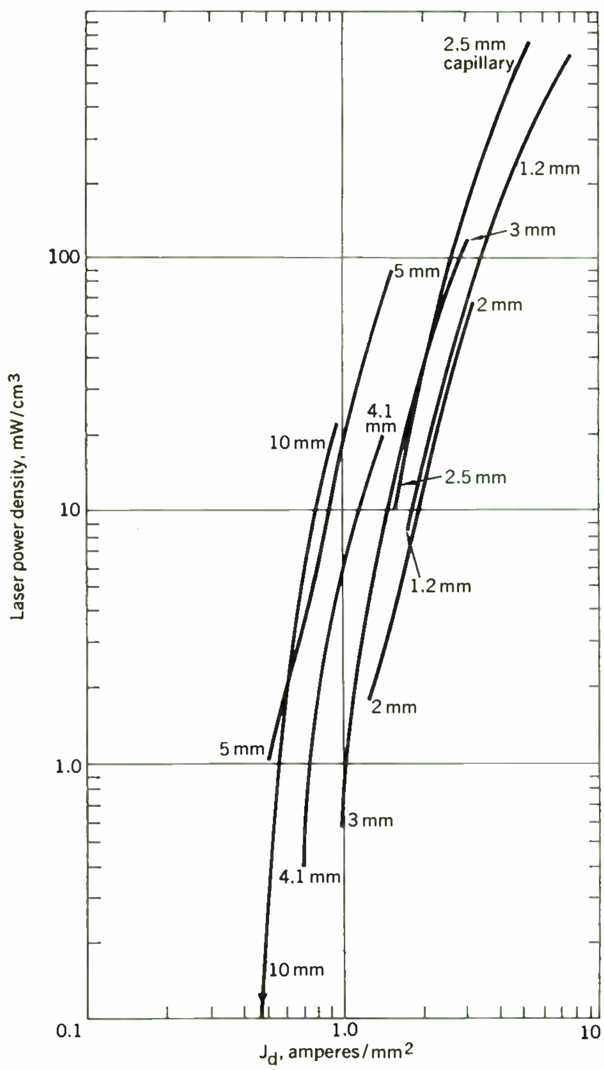


Fig. 2. Graphic plot of dominant line diametrical threshold characteristic for ionized-argon laser.

Fig. 3. Graphic demonstration of quasi-invariance of generation density vs. current density.



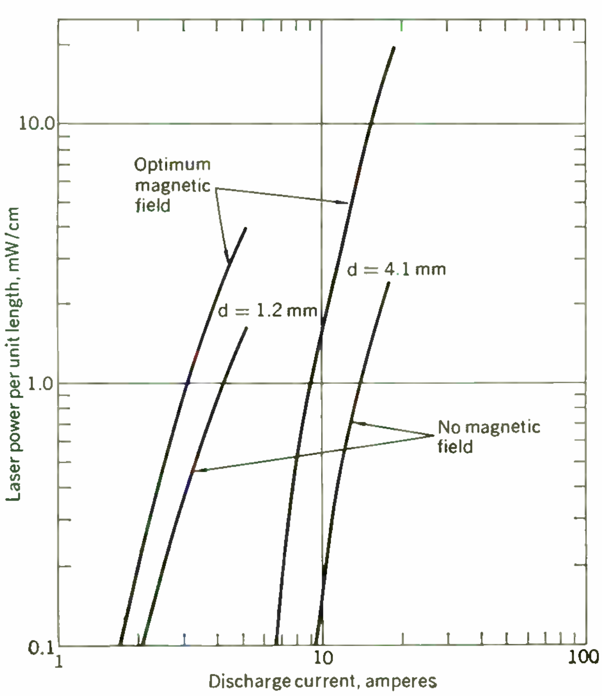
firmed the following stepwise nature of the excitation mechanism:

1. Neutral-atom-excited states and the ion ground state are formed by electron impact on the neutral-atom ground state and the "1s" metastable states.
2. The ion-excited states are formed by electron impact on the ion ground state and ion metastable states.
3. The contribution of cascade radiation to the population of the upper laser levels is negligible compared to that of electron impact.
4. Population inversion occurs because the lower laser levels are kept nearly empty by short-lived vacuum ultraviolet transitions to the ion ground state. If the excitation rates to the upper and lower laser levels are comparable, then the longer lifetime of the upper laser level automatically produces inversion. At sufficiently high current densities, this inverted population may become large enough to provide optical gains that are greater than the cavity losses.

Under conditions in which the electron drift velocity and average thermal energy are independent of current density, the stepwise excitation process predicts these variations (with current density J_d for various level populations): all neutral-atom-excited levels—population independent of J_d; ion ground and metastable levels—population proportional to J_d; ion radiating levels—population proportional to J_d^2.

The capillary discharge used to obtain ion laser action is in between the weakly ionized glow discharge and the strongly ionized arc. Over most of the pressure and current-density range for which continuous-duty laser action occurs, the electron drift velocity and thermal energy are observed to be independent of J_d, and the predicted dependences of level populations on J_d are observed to occur.

Fig. 4. Magnetic field enhancement for two capillary sizes, plotting laser power per unit length vs. discharge current.



The measured parameters of the capillary discharge are in reasonable quantitative agreement with ambipolar diffusion theory at large pressure-diameter (pd) products, and with the Tonks–Langmuir ion free-fall theory at small values of pd.

The favorable lifetime ratio exists for the neutral and ionic species of nearly all the elements. Therefore, it is not surprising that ion laser action has been found to occur widely throughout the periodic table. Inversion in *neutral* atoms, however, by the process just described, is not observed because the effective lifetime of potential lower laser levels is usually much longer than their *natural* lifetime. This increase in lifetime occurs because the ultraviolet resonance radiation is strongly reabsorbed by the atomic ground states, thereby repopulating the lower laser level. Such resonant reabsorption also occurs in ions and sets an upper limit to the current density that may be employed and to the laser power that can be obtained. This limit has not yet been reached in continuous-duty ion lasers, but it has been reached in the higher-current-density pulsed lasers.

The line width of ionic transitions in a capillary discharge is determined by Doppler broadening produced by the high ion temperatures (>2000°K, typically). The natural line width is negligible compared to the Doppler width. Usually, many $c/2l$ modes of the Fabry–Perot cavity (c = speed of light, l = length of cavity) will oscillate within the Doppler-broadened line (typically, $c/2l$ is of the order of a few hundred megacycles per second, and the ion Doppler width is in the order of a few thousand megacycles per second). At oscillation threshold, however, only the mode nearest to line center will oscillate, and the actual gain is then essentially equal to k_0, the zero signal gain at line center.

In the stepwise excitation process, k_0 is proportional to J_d^2L, in which L is the discharge length, and is a rather insensitive function of pd.

If threshold currents for capillaries of different length and diameter are compared in cavities of equal optical loss, then the quantity I_t^2L/d^4 should be constant (I_t is the current required for oscillation threshold).

Figure 2 illustrates unpublished data from measurements at Bell Telephone Laboratories, covering five orders of magnitude in the variation of I_t^2L vs. capillary diameter d for the 4880 Å transition in Ar^{II}. The measured dependence lies between d^{-4} and d^{-5}, which is regarded as a reasonable agreement with the predicted d^{-4} dependence in view of the experimental uncertainty in assuring the same losses from cavity to cavity. These data were obtained at an approximately constant value of pd.

The output power per mode of a multimode laser is a function of k_0 and is proportional to the modal volume. The functional dependence of power on k_0 is faster than linear at low levels of oscillation, and it approaches linearity at very high levels of oscillation. Therefore, if the oscillating modes completely fill the discharge volume, then the output power per unit disharge volume should depend only on J_d, varying more rapidly than J_d^2 at low J_d, and approaching J_d^2 at high J_d.

Further data from the unpublished measurements just cited are shown in Fig. 3. The predicted trend of the J_d dependence of laser power per unit volume is observed for capillaries of widely different diameters. If there were no pd dependence of k_0, if the modes completely filled the discharge volume, and if the ratio of line center gain to cavity loss were the same for all of the data, then the curves in Fig. 3 should coincide. The differences existing between these curves can be ascribed to the failure to achieve these conditions and also, for the larger diameters, to the possibility of a discharge self-constriction, which causes the true current density to be higher than that estimated from the capillary dimensions and the current.

Influence of an axial magnetic field

Application of an axial magnetic field to a capillary discharge causes the electrons to spiral about the magnetic lines of force, thereby reducing the radial diffusion of the electrons. Although the ions are negligibly affected by the magnetic forces, their radial diffusion is also reduced because the condition of space charge neutrality requires them to have the same radial velocity as the electrons. Thus, since the ion loss rate decreases, the electric field and the electron temperature and drift velocity also decrease (i.e., for a given discharge current density, the magnetic field produces a decrease in the excitation rate coefficients, but an increase in the electron density). Therefore, it is not logically obvious whether the magnetic field will enhance or reduce the ionic radiation. Experimentally, however, the ion laser power and spontaneous emission can be increased by the presence of a suitable longitudinal magnetic field.

The laser power enhancement obtained by using a magnetic field is illustrated in Fig. 4 for two different capillary diameters.[12] The results are in qualitative agreement with a study of the positive column in a longitudinal magnetic field by Bickerton and von Engel.[13] The optimum value of the magnetic field decreases with increasing capillary diameter, and for a given diameter the optimum is rather broad. At currents sufficiently far above the threshold, the degree of enhancement, for diameters between 1.2 and 4.1 mm, is found to be almost proportional to the capillary diameter. Since the power dissipation per unit length of capillary and the erosional damage to the capillary both decrease as the capillary diameter increases, it is obvious that the use of a magnetic field improves both the efficiency and the life of the laser.

General characteristics of ionized-argon lasers

In this section, a rather complete description of the physical and operational features for the archetype of a medium-power ionized-argon laser will be given. Figure 5, representing the heart of an argon laser system, shows several features that immediately distinguish it from the older and better known He–Ne low-power gas lasers, although similarities such as dc discharges, Brewster-angle windows, etc., also do occur.

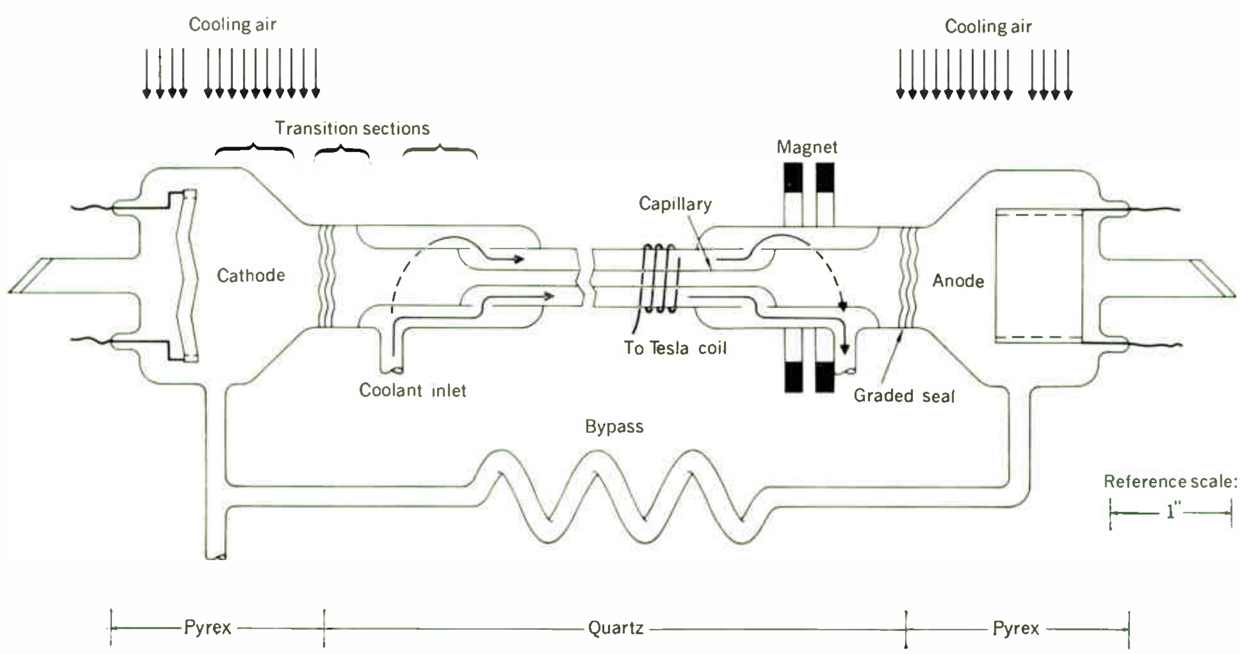
Let us assume the rated power output is one watt over all the lines for the purpose of seeing the essence of the laser, stripped of the problems that arise when higher powers are attempted.

Fused quartz is ideally suited for the construction of the critical capillary and transition regions by virtue of its low thermal expansion and its ready plasticity. The three stepwise transition sections connect smoothly the large-diameter (about 50 mm) cathode and anode to the capillary bore (about 2.2 mm in diameter).

Oxide-coated cathodes are efficient emitters for arc discharges, an active area of 20–40 cm^2 being provided for the 14 amperes (or less) needed to generate one watt in typical lasers. The anode may be made of molybdenum or stainless steel, and it is adequately cooled by radiation.

The need for a gas return path, or bypass, as shown in Fig. 5, has been mentioned briefly in the original CW reference[8] and in more detail in a later communication.[14] The helical form for the bypass provides compact storage for the tube length needed to avoid striking the arc through this path. Somewhat greater gas conductance can be achieved in variations of this idea by the use of porous glass frits to inhibit the discharge, a strong magnetic field for the same purpose, a glass stopcock, and the incorporation of a separate gas pump in the return line.

Fig. 5. Schematic diagram of an argon laser apparatus.



Ordinary tap water is the usual liquid coolant used, although other fluids may have advantages in special circumstances. From the coolant temperature rise standpoint, a 1 gal/min flow rate is more than adequate, but larger flows are desirable to maximize the film heat-transfer coefficient at the quartz–coolant interface.

The laser apparatus configuration may include any of several modifications:

1. The cathode and anode bulbs swung over at right angles to the optical axis.
2. Heat shielding on the cathode.
3. Normal incidence windows.
4. Linearly tapered instead of stepwise transitions.
5. An anode small enough, taken together with a coaxial helical bypass, to permit single-ended loading into the solenoid. The slender appearance of the laser, resulting from this last variation, intensified by the strong chromatic impressions in the pink, blue, and aquamarine areas, has caused one viewer to coin the phrase, "photon cannon."

In addition to the essentials shown in Fig. 5, a complete ionized-argon laser system includes (1) power sources: magnet supply, cathode-heating supply, main discharge supply; (2) monitoring equipment: gas pressure gauges, discharge voltmeters and ammeters, power output thermopile; (3) protective equipment: water-flow interlocks, ballast resistors, overcurrent relays; and (4) miscellaneous equipment: starting circuit, gas reservoir, getters. A set of mirrors is also included.

The operating current threshold for the onset of ionized-argon laser action in dimensions (2.2-mm bore by 650-mm length) suitable for the generation of one watt continuously, will be about 0.67 ampere. High-reflectance ($R > 99.6$ percent) mirrors, effective over the 4545–5287 Å wavelength interval, appropriate to argon, are assumed. The discharge voltage in these circumstances will be approximately 480 volts, so a power input of a substantial fraction of a kilowatt is necessary to see anything.

The initial laser action always occurs on the dominant blue 4880 Å transition. All threshold data such as those shown in Fig. 2, as well as the results of other experiments reported in the literature, will refer to this important wavelength.

As the operating current level is raised above the threshold value, one is conscious of the very rapid initial increase in power output—of the order of an I^6 dependence. This rate of growth diminishes to I^4, and eventually to an I^2 dependence before the rated power is attained. Well in advance of this point, however, one is aware of some color changes occurring in the laser beam that are not attributable to intensity effects in the human retina. These signal the passage of one current threshold after another for laser action on additional ionized-argon visible transitions. Reference to Table I makes this clear. These are data taken on a one-watt-type laser, with fully reflecting mirrors and a rather weak magnetic field.

It is interesting to observe, in Table I, how the two strong lines at 4880 and 5145 Å dominate the 4545, 4658, and 4727 Å lines through the accident of sharing the same lower level. Only through the use of moderately selective elements such as a prism do we gain easy access to these latter transitions.

The number of laser lines at 15 amperes (shown as eight "checks" in Table I) will ordinarily drop to six or seven in practical lasers. Because it would be prohibitively difficult to construct a "dumping" mirror for optimum transmission in each and every transition, some loss in output is occasioned by the necessity of using commercially available mirrors. These have essentially flat bandwidths over a span approximately commensurate with the argon spectral range. The average transmission needed for any laser, at a chosen operating current, can be determined empirically.

The distribution of output power among the six or seven laser lines at "full on" current levels will generally have the ranking—but not the exact proportions—of $(I_t)^{-1}$. The preponderance of output power will be in the 4880 Å and 5145 Å wavelengths, in which the percentages might be 45 and 37 respectively. Behind these, either the 4965 Å line or the 4765 Å makes a poor third, with about 6 percent of the total.

Logically, one could hope to obtain monochromatic outputs of about 70–82 percent of the polychromatic power by the conversion of at least part of the 5145 Å inversion energy to the 4880 Å transition when only the latter is allowed to trigger laser action through selective means, and coupling between the lines might be expected through a common lower level. Experimentally, this is found to be untrue, however, and only negligible cross-coupling effects are observed in a three-mirror experiment for any but quasi-threshold laser experiments.

In a three-mirror test, a dispersive prism at one end of the laser allows angular displacement of the two wavelengths in question, each being reflected by adjacent independent mirrors. Occultation of either beam permits

I. Argon laser characteristics

Color Range	Å	Reaction	Threshold Current with Prism, amperes	Threshold I without Prism, amperes	Laser Lines without Prism 5 amperes	7.5	10	15
Violet	4545	$4p\ ^2P^0_{3/2} \rightarrow 4s\ ^2P_{3/2}$	10.5	>15				
	4579	$4p\ ^2S^0_{1/2} \rightarrow 4s\ ^2P_{1/2}$	5.2		√	√	√	√
Blue	4658	$4p\ ^2P^0_{1/2} \rightarrow 4s\ ^2P_{3/2}$	6.9	12.5				√
	4727	$4p\ ^2D^0_{1/2} \rightarrow 4s\ ^2P_{3/2}$	6.7	>15				
	4765	$4p\ ^2P^0_{3/2} \rightarrow 4s\ ^2P_{1/2}$	3.8		√	√	√	√
	4880	$4p\ ^2D^0_{5/2} \rightarrow 4s\ ^2P_{3/2}$	1.45		√	√	√	√
	4965	$4p\ ^2D^0_{3/2} \rightarrow 4s\ ^2P_{1/2}$	4.0		√	√	√	√
	5017	$4p'\ ^2F^0_{5/2} \rightarrow 3d\ ^2D_{3/2}$	6.0			√	√	√
Green	5145	$4p\ ^4D^0_{5/2} \rightarrow 4s\ ^2P_{3/2}$	3.6		√	√	√	√
	5287	$4p\ ^4D^0_{3/2} \rightarrow 4s\ ^2P_{1/2}$	10.5	10.0			√	√

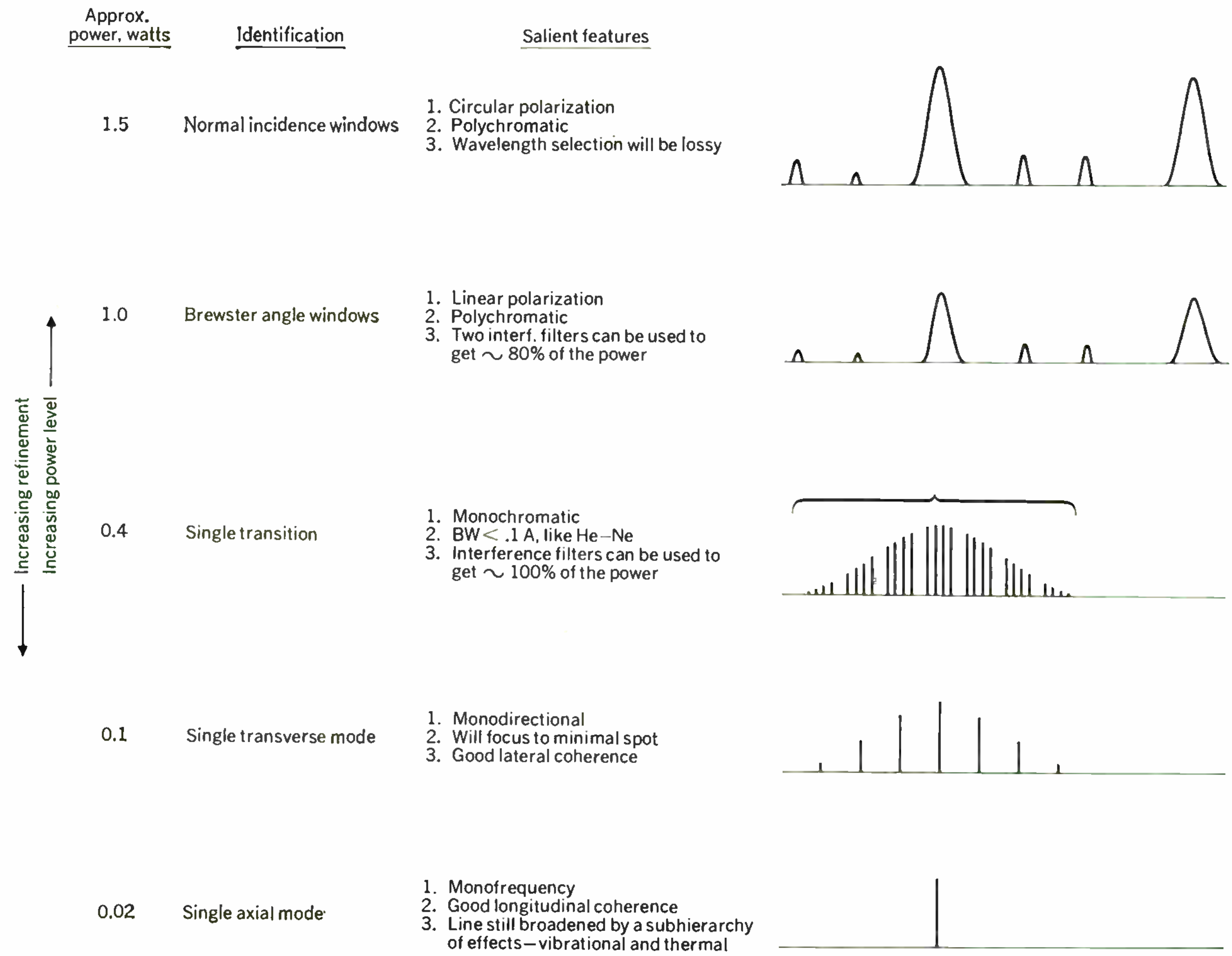


Fig. 6. Spectral diagrams of a set of distinguishable argon lasers.

the examination of the occultation effect on the other beam. The programming of mirror reflectivities allows examination of the phenomena over a wide ratio of cavity power levels.

It may be interesting at this point to make a short digression into the matter of monochromaticity, particularly as it applies to ionized-argon lasers. As Heavens[15] indicates, "the great interest of the maser [laser] as an oscillator is that it generates highly coherent radiation *at substantial power levels.*"

By the use of a pinhole and a high-quality lens, a plane wave with excellent spatial (lateral) coherence properties, can be achieved from any source. The plane wave can also be made as monochromatic as desired by passing it through a succession of increasingly selective filters (Fabry–Perot etalons). Of course, at the end there is no energy content since our prelaser sources are not bright enough.

The laser possesses a hierarchy of coherences as shown in Fig. 6, but here the intermediate steps tend to be discrete rather than continuous. There is nothing "sacrosanct" about the size of the power steps chosen between hierarchical levels; they are only meant to be representative values. An attempt has been made to correlate the size and number of the elements in the spectrum drawings with the assigned step values. There is also an arbitrary point in ranking the "single transition" above the "single transverse mode," as indicated in Fig. 6. For some lasers, they are reversed, but, in any case, the intent is to illustrate the power penalty (about 75 to 1) in going from the top to the bottom of the figure.

Casual inspection of an argon laser beam shows it to have many of the familiar attributes of the ubiquitous He–Ne red laser. The beam spread angles look the same, transverse mode patterns appear as usual, and the characteristic "granularity" is readily observed. If we look hard enough, however, we should find some manifestation of the fact that ion laser emission takes place between excited levels in a moving medium—the ions themselves. In neutral gas lasers, the atoms have motion, but not in any directed sense.

Some of these effects have been reported recently by Rigrod and Bridges.[16] They have noted bistable traveling-wave (rather than standing wave) lasers in ring resonators, caused—in part, at least—by the Doppler-shift effect. Certain more subtle manifestations are only discernible with instrumentation suited to mode analysis. As they indicate, ion drift effects may lead to the eventual realization of energetic single-frequency oscillators (thereby superseding part of Fig. 6) and high-power nonreciprocal amplifiers.

Finally, a word about the individual efficiencies of these devices: a physicist working on the analysis of the many processes that occur simultaneously in the arc will, with consummate logic, invent a "discharge" or "column" efficiency, whereby

$$\eta_D = \frac{W_{\text{optical}}}{E \cdot l \cdot I}$$

in which E is in volts/cm, and l measures only the uniform part of the arc column.

Since the majority of the input power is dissipated in the arc column, any substantial improvements in the overall efficiency must originate here. One simple analysis shows a reasonably firm upper limit at 1.7 percent for η_D, but in actuality it is far from this at present. Labuda[12] reports $\eta_D > 0.1$ percent, using large-diameter (8 mm) tubes. For the bore diameters we are considering (2.2 mm), η_D may be about 0.03 percent at the one-watt level, with Brewster-angle optics and overall transitions simultaneously.

The most accessible efficiency arises simply from the product of the laser terminal voltage and current. It will differ from η_D by the losses associated with creating and destroying the ions and electrons at the cathode and anode terminals. For one-watt lasers, this efficiency could be 0.026 percent.

We can appreciate the power required to operate all the laser auxiliaries such as thermionic cathode, magnet, coolant circulators, cooling fans, ballast resistors, etc., and to allow for losses in the main discharge supply, if we examine the overall efficiency. This parameter is directly proportional to the amount of effort expended on items such as cathode heat shielding and conversion to permanent magnet fields. A reasonable value for this efficiency would be 0.022 percent.

Contemporary trends

There is some risk involved in attempting to describe and evaluate current research in the exciting area of argon lasers. We hope our descriptions of recent developments are reasonably accurate, but assessment, added to description, cannot. of course, take refuge under the same claim. Thus we will attempt to keep our evaluations at a modest level.

As in any field sustaining a high degree of research endeavor, major new discoveries can redirect subsequent efforts into totally unforeseen channels. At this writing, the very recent introduction of induction pumping of ion lasers may be an event of such significant magnitude.

As the deficiencies of fused quartz in high-current-density (500 A/cm^2 and above) discharge tube service began to be fully recognized, a good portion of the ion laser effort went into searches for substitute materials and alternate ways of working with quartz. These investigations, which are far from complete, have led to three general ways of dealing with what we may call the "lateral containment" problem. The discussion of these three methods will be in an ascending order of complexity.

The first technique under consideration simply calls for the utilization of large-bore discharge tubes in lieu of the commonly used 1–3-mm size. The term "large bore" designates a 7–8-mm diameter as a representative range.

Consider. in an elementary way, the processes involved —disregarding magnetic field effects for the moment—in a double-diameter laser, 6 mm in diameter, derived from a 3-mm base. With twice the conventional bore size, the operating current must be increased fourfold to maintain the current density. The voltage gradient will be cut approximately in half, so the input power only doubles. Since the cooling perimeter also doubles, the wall thermal loading remains unchanged. The optical gain will be reduced, but since it is already reasonably high in ion-argon lasers, a downward readjustment of the output coupling mirror transmissivity will bring in a new optimum coupling. Thus, as long as the gain remains significantly well above the unavoidable resistive optical losses—say 0.3 percent per pass—this readjustment will be quite effective. Since we are assuming a reasonably long laser at the outset, the output power can be expected to be three to four times greater than that of the base laser, simply on the basis of the four-volume enhancement.

The application of an optimized axially directed magnetic field will be differentially favorable (in output power) to the double-size tube, as shown in Fig. 4. The field-favorable factor may be something on the order of two, to yield a total output power gain of 6 or 7 over the base laser. Since the latter may bear a rating of 1.5–2 watts (for a 3-mm bore), we have escalated to the 10-watt level rather easily.

Moreover, the concomitant reduction in the voltage gradient (by as much as 2 to 3), coupled with the visually observable condensation of the ionized column away from the walls, brings the wall thermal-loading factors down considerably—sufficiently, in one case, for the researchers[12] to attempt ambient cooling only in a multiwatt laser. The discharge efficiency, η_D, approaches its best values, as previously mentioned, under these conditions.

But there are a few disadvantages. First, the overall efficiency will not necessarily reflect the discharge efficiency improvements caused by change in the laser from a device of some tens of ohms dc resistance to one of about one-ohm resistance. Substantial fractions of the total dissipation may unavoidably occur at the conductor–plasma interfaces, and there are problems of meeting cathode emission requirements in structures of modest size.

Second, there is little hope of operating large-bore argon lasers at their usual power outputs in the single transverse-mode regime. Even the relatively docile He–Ne red laser is made to do this poorly (in large bores), and the task is compounded in argon by the high neutral gas temperatures and gradients that prevail. It is interesting to observe the intermittency of laser action in argon caused by gradient fluctuations when the optical cavity is bounded by two flat mirrors.

The third drawback relates to the necessity of providing a magnetic field of about 1000 gauss over a substantial volume. Only a small amount of large-bore operation will be attempted without the magnetic field. Labuda[12] points out the manner in which the incorporation of an odd number of symmetrically disposed magnetic field reversals not only will avoid the loss of power caused by Brewster-angle polarization effects but also will allow significant weight reductions to be made in the magnet. The concept has probably not yet been validated by experiment.

As an example of the practicality of the large-bore idea, we may cite experience at the Raytheon Research Division with a 6.75- by 1670-mm "house" laser, operated intermittently for the last six months. A large variety of experiments (chiefly biological, but including microcircuit trimming, Stark-effect broadening, and other digressions into the physical sciences) have been completed at power levels up to—and beyond—10 watts, dependent upon the state of the mirrors. The laser operates normally to the supply limit, 75 amperes. Quartz darkening is scarcely observable at the half-year operation point.

The second general solution to the lateral containment problem involves the substitution of a more refractory and more thermally transmissive—but still electrically nonconductive—material for the fused quartz. A casual examination of the tabulated properties of refractory ceramics reveals a marked superiority of certain of these over quartz in the important categories of maximum working temperature and thermal conductivity. Such ceramic materials that have served, or are potentially adaptive for argon discharge service, include beryllia and zirconia and the following forms of alumina: the dense, opaque bulk variety; thin, anodic films; sapphire; and translucent Lucalox®. Boron nitride is also in this category. The high-temperature capabilities of graphite—although it is not a ceramic—are under consideration.

Each of these materials will possess, in one degree or another, additional attributes favorable to the success of lasers. A short list of these properties would include a low thermal expansion rate, transparency or translucency to significant bands of radiation from the arc column, vacuum tightness, maintenance of thermal conductivity values to high temperatures, and availability in forms and at prices appropriate to this task. The relative resistance to ionic attack (sputtering) among these refractory materials is also important, but information on this point may be incomplete. The susceptibility to thermal shock of the otherwise very suitable ceramics may be the most serious problem to overcome in bulk ceramic investigations.

Efforts at Raytheon, over the past six months, on thin, anodized alumina films allows us to present a short, first-hand progress report on this particular variation of ceramic lasers. In the anodization of aluminum, a thin aluminum oxide film is built up on the anode of the work that is immersed with a cathode in a suitable electrolyte. Most of the laser tube anodization has been carried out with a sulfuric-oxalic acid and water electrolyte, operated at 15°C, to optimize a certain combination of properties such as durability, pore size, and crack resistance.

Operating experience with these coatings on monolithic 3- by 570-mm aluminum laser tubes has been limited because of arc breakdown at the tube ends. Generally, this is traceable to cracks in the anodic film on curved surfaces. Much more satisfactory operation, up to and above 20 amperes, has been achieved in coarsely segmented structures and in 300-mm single pieces. Medicus[17] has demonstrated significant life in a 400-mm discharge body, using oxalic-acid-derived films that have less sensitivity to cracking.

The proprietary Sanford process of anodizing has also been evaluated in more finely segmented 4-mm high-current (75-ampere) discharge lasers, with excellent results except at the energetic cathode and plasma transition region. Apparently, single-piece anodized lasers can be successfully operated, provided that sufficient consideration is given to minimizing the voltage stress at corners and in keeping the total length within bounds, and in setting the discharge current to values commensurate with the desired lifetime.

In all, one might think that ceramic-derived lasers would fulfill an intermediate role between fused quartz structures and the segmented metal ones. They have many advantages in ruggedness, cleanliness of the discharge, and simplicity—provided certain engineering problems in their fabrication can be overcome.

The idea of metal discharge tubes for argon-ion laser service was conceived more than a year ago by at least two completely independent groups active in this general area. The utilization of metal for arc confinement, while immediately attractive because of its excellent thermal conductivity and formability properties, is possible only through transverse segmentation or fractionation of the desired metal column length into shorter, separate, and insulated pieces. Thus the operation—and more particularly, the starting—of an arc discharge through continuous metal envelopes of length to diameter ratios: $L/D >$ 100, as typical of gas lasers, is essentially impossible. The metal block will merely serve as a combined cold cathode–anode to the discharge, with no gaseous conduction in the bore. For L/D ratios of 5—or even 10—the usual voltage span across the segment is then insufficient to support significant cold-cathode emission, and the discharge negotiates the bore instead as the easier path. A collection of such segments placed end to end, each segment of which is electrically "floating," constitutes the metal discharge body.

Although the thought processes that led to successful segmented metal discharge tubes can be regarded as a classical case of invention to meet a need, subsequent literature reference reveals an ideological ancestry that predates the laser. As examples, Maecker[18] in 1956, Shumaker[19] in 1961, and, more recently, Emmons[20] have all employed liquid-cooled segmented copper discharge channels at arc currents, or current densities, substantially in advance of those later attempted in laser usage.

Two topologically distinct forms of metal-tube lasers have evolved over the past year. In the simpler of these, the stack of separate metal segments is completely internal to a transparent envelope (quartz may be assumed), thereby avoiding any additional vacuum seals required by the metal construction. The segment cooling is by radiation only, through and to the envelope, so that materials for the latter must be chosen for their transparency to 1.5–2.0-μm wavelength radiation, which corresponds to the radiation maxima at 1500–2200°K segment temperatures. In such cases, ambient cooling would suffice for the envelope, but it would be necessary to intercept the transmitted radiation eventually (usually on the inner wall of the solenoid structure) for the comfort of operating personnel. The method is somewhat limited, since only the refractory metals such as molybdenum and tungsten can be selected for segment materials.

A liquid coolant can be introduced into or around each segment in the other practical form of metal-body lasers. The attendant excellent cooling properties may be obtained only by the added complexity of one vacuum seal per segment. The segments themselves can be conveniently made of copper, aluminum, or alloys of either metal. Composite structures—for example, a molybdenum sleeve pressed into an aluminum matrix (see Fig. 7)—offer significant advantages in cost and weight over any single material segment. Various means of separating and insulating the segments have been employed, including anodized films, ceramic spacers, and plastic annuli.

Both vacuum-grade epoxy and Viton "O" rings have provided good service as the vacuum sealant, although the latest research work is on the more sophisticated ceramic–

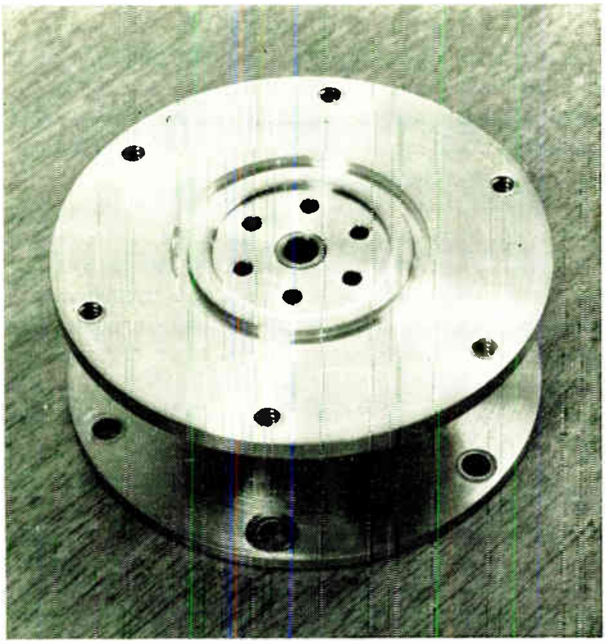


Fig. 7. Close-up photo showing a composite aluminum-molybdenum segment in which a molybdenum sleeve is pressed into an aluminum matrix.

metal brazing techniques that are commonly used in microwave tube construction.

While gross melting of the discharge channel in metal lasers is not experienced, there are subtle changes in the segment bore, particularly at the cylinder edges or rims. The phenomenon responsible for this is sputtering—the ejection of metal atoms into the discharge space produced by ionized-argon atomic impact. We note that the gas species, ion density, and pressure ranges that prevail in ionized-argon lasers are, unfortunately, a combination of properties well suited to promote high sputtering rates.

Effects of sputtering in segmented structures are aggravated over those that occur in uniform discharge channels. First, the discrete intersegment voltage drop creates strong local electric fields to accelerate positive argon ions into the adjacent bore rim, and the resulting metal ejecta will, in great part, deposit directly on the opposite rim, thereby assuring the probability for ultimate short-circuiting of segments in this critical region. The intensification of sputtering is the penalty for the benefits that segmentation otherwise affords.

The axial segment length, therefore, together with the selection of molybdenum or tungsten as a sputter-resistant bore material, becomes a dominant factor in determining the sputtering rate. More subtle factors certainly are operative, however, in the sputter problem: the choice of gap and paraxial form for the segments, the role played by the magnetic field, the current level for which the influence of doubly charged ions becomes important, etc. Sputtering effects may limit the long-term life of a laser, but for short terms the maximum laser operating current will find other limits.

One way of easily applying more load to the laser structure is to substitute neon for argon. The arc column-voltage gradient is about one third larger in this case. In such work, the first CW ultraviolet laser action, with an anodized-bore aluminum segment laser, was recently observed at Raytheon. Considering that the laser beam threshold was 65 amperes in a 4-mm-bore tube, the practicality of metal-derived laser structures is self-evident.

When the course of argon laser development (in quartz) surpassed the one-watt level, research workers began to observe certain erratic tendencies in the intensity and form of the optical beam. Over a period of hours, for example, decay of the output power was noted. Another novel manifestation was the asymmetry of the beam, which usually appeared narrow at the anode end and wider than normal at the cathode end (exit) in approximate proportion to the discharge current. Generally, the narrowing degenerated into serious distortion of the beam from its usual round cross section, and, in such cases, the output power failed by a wide margin to move upward as a quadratic function of the current. These are some of the more evident phenomena associated with the "longitudinal containment," or mirror, problem.

The results of some simple tests, taken collectively, yield a reasonably firm proof that mirror heating effects are the principal source of the phenomena. Time constants of seconds are often observed in the beam narrowing when the optical cavity is occulted and released. These phenomena are evident in proportion to the optical cavity power as this parameter is varied at a fixed discharge current by mirror substitution. Bleach spots have been seen in several dielectric-coated mirrors used on energetic lasers. Quite marked differences between mirrors of different dielectric makeup are readily noted, particularly in their response to full illumination.

Although multiple-layer dielectric-coated mirrors have been the backbone of gas-laser technology, we must recognize their possible limitations for high-power (5–100-watt) argon laser service. All such mirrors are at least slightly lossy; in one theory,[21] supported by calculations, the resulting heat generated in the thin films is readily coupled to the mirror substrate and deforms the latter locally. Such a concept would provide a satisfactory explanation of the observed beam asymmetry, since a pair of initially identical mirrors would be transformed into a more hemispherical geometry. The situation is stable because the dense heating, caused by the narrow beam, would ideally maintain the narrowness. Ionized-argon lasers have been shown to act as their own gas "lens," and it may be the "sign" of this lens that initially determines which end will narrow down.

If the theory is correct, a number of remedial measures are apparent, but none are as refined, inexpensive, or lossless as the original dielectric mirror technology. One of these is shown in Fig. 8, in which an air-driven mirror mount exposes the greater area of an annulus to the optical beam. Dielectric-coated metal mirrors can also be utilized with output coupling achieved by means of a small hole. In addition, work is proceeding with totally internal reflective optics.

The question is sometimes raised: Can argon-ion lasers be efficiently RF-excited as is done successfully in most other gas lasers? The answer is "no," if the usual *modus operandi* is employed; that is, with discrete metal RF electrodes encircling the discharge tube exterior. Huge displacement currents through the dielectric wall would be needed in the argon case to match the very low arc impedance. Moreover, the dielectric wall should include an intervening liquid coolant, with its attendant RF losses, if any significant power magnitudes are attempted.

By using the dual of the *E*-field coupling just described, W. E. Bell[22] has solved the RF-drive problem for arc lasers. Reference to Fig. 9 makes the method clear; the closed-circuit laser tube becomes a one-turn secondary for the RF transformer, which is conceptually ideal for this application.

The absence of a cathode and anode is immediately indicative of important applications in reactive gas investigations and, indeed, this ring-discharge laser is being so used. Confining the discussion to argon, however, Bell[23] reports the attainment of CW operation in such rings, with a tentative goal of one watt. Magnetic fields effect large power enhancements (of the order of 10 to 1) and bore erosion in quartz discharge channels is conspicuously absent. Presumably, this latter benefit stems partially from the inability of the ions to execute large motions in the high-frequency fields. Much higher radio frequencies (10 MHz) than originally reported are now being used to good effect.

Other useful features of RF pumping, observed by Bell, are:

1. An absence of gas absorption with time.
2. A noiseless optical beam (dc-arc discharge beams will often bear evidence of a variety of discharge phenomena).
3. Thirty percent narrower atomic line widths.

Bell believes the discharge (plasma) efficiency to be about twice that of conventional argon lasers, while the overall efficiency may be about the same for both. This differential is the penalty for "preprocessing" the excitation into a more elegant form.

Projection of the role to be played by ring-discharge lasers in the high-power quest is probably conditioned by two factors: first, that the anomalously favorable discharge efficiency and magnetic field enhancements will hold up for high, continuous, circulating currents; and, second, that ring discharges will operate normally in structures other than quartz. The growing level of interest in ring discharge lasers should provide us with the answers within the next six months.

Uses and prospects

An active program to build and assess two kinds of discharge structures, with projected ratings of 100-watt output in ionized argon, is under way at Raytheon. The first of these structures will be a 4-mm by 1.7-meter segmented aluminum–molybdenum body, with provision for substitution if tungsten is needed. The 100-watt level should be reached at 110–120-ampere discharge current, based on extrapolation from available 20-watt CW results and the informative pulsed data of Cheo and Cooper.[10] A large-bore (8-mm by 8-meter) water-cooled quartz laser is also planned, in which 125–135 amperes will be required to attain this goal. Each laser will need 500–625 volts across it, so that the expected accessible efficiency (based on voltage × current) will be about 0.15 percent.

The optical containment problem will probably not soon be solved in any completely satisfactory way, so that the efficiencies just postulated may not be met for this reason. A genuine breakthrough in this area is almost mandatory if full advantage is to be taken of the discharge tube development pace.

The outlook for optical powers much greater than 100 watts is not very good because of expected saturation effects in Ar[11]. When and if this is proved to be true, some of the argon effort will undoubtedly devolve on the full

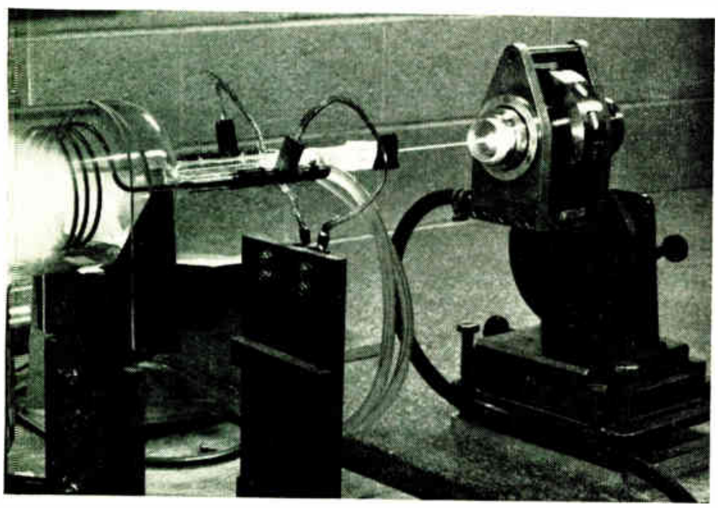
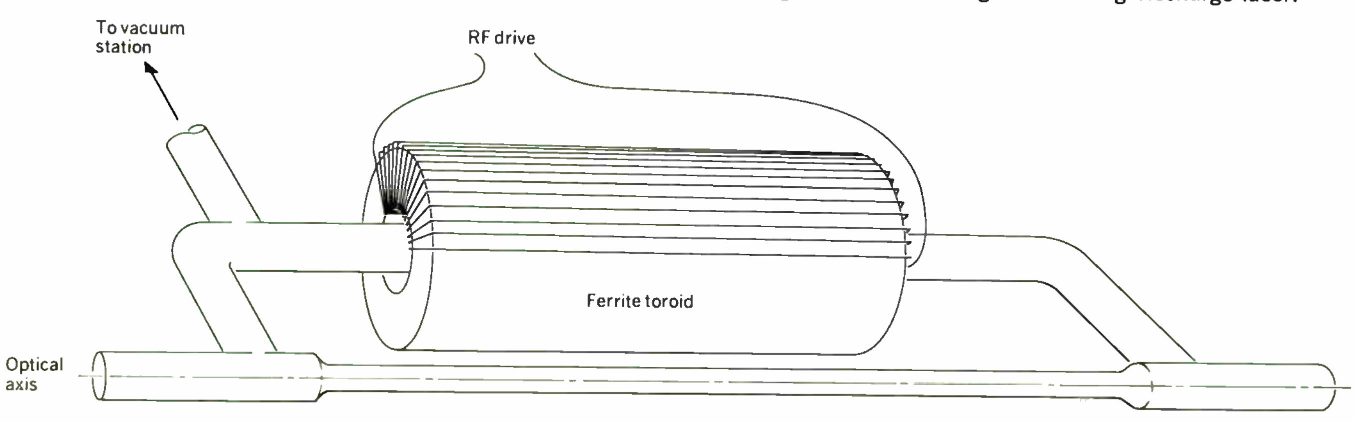


Fig. 8. Rotary mirror mount in stopped position. The air-driven mirror mount exposes the increased area of an annulus to the optical beam.

Fig. 9. Isometric diagram of a ring discharge laser.



exploration of the CW potentialities of neon in the ultraviolet area. Here, of course, the mirror problem will be greatly accentuated because of the inherently lossy nature of most useful coating materials in this spectral region.

Inasmuch as it is probably the first energetic CW laser, the optical beam from ionized-argon devices has been focused on a variety of materials for a number of expected effects. We might label one such effect, in which high-energy densities over a small area are required, an "optical torch." For example, argon lasers may possibly be used in the trimming of microcircuit resistors[24]; for making self-cauterizing incisions in laboratory animals[25]; for welding, brazing, and severing small wires; for grooving plastics, etc. In the optical torch applications, the user is essentially indifferent to the wavelength supplied. Hence, some of these tasks may be more effectively accomplished with the very efficient CO_2 laser. The scattered light from the focused argon beam impingement on a target is as much a hindrance as a help because of its extreme brightness. It is important to note that for highly focused work, the advantage will favor the argon laser because of its 20-to-1 wavelength superiority over CO_2. The power levels used in this work have been in the readily available 1–5-watt range; but more power is clearly indicated for surgical incision effectiveness.

The argon laser should find its intrinsic applications predicated upon the new and strategic spectral region it has introduced. Ignoring the obvious esthetic appeal of visible range lasers, one must appreciate the more tangible advantages of the 4500–5200-Å band. Along with argon laser beam visibility and all that it implies in the ease of preparing experiments, the wealth of well-developed optical hardware available is a prime asset.

We note that photoemissive detectors and many film emulsions "peak up" in the blue and violet spectral ranges, with 25–30 percent quantum efficiencies possible for the detectors. Further, these wavelengths are short enough to meet more than the need of biologists for CW radiation in new and differently absorbing spectral regions. The selection of wavelengths affords some opportunity for testing chromatic effectiveness.

The best known utilization of argon lasers emanating from one or more of these attributes is in connection with the Gemini 7 space flight. Three ground-based, geographically separated argon lasers were installed to serve as aiming beacons or visual targets for the astronauts. Although the beacons were not all seen, the method itself should find increased application, inasmuch as a laser is by far the brightest of controllable, man-made light sources.

We would be surprised if argon lasers are not found well adapted to at least a few other long-distance tasks, such as the tracking of satellites equipped with retroreflectors, the probing of the atmospheric microstructure by backscatter, and the verification of the prediction of King and Kainer[26] as to the practicality of reasonably wide-spaced (100 miles) terrestrial scatter links.

In the laboratory, the blue-green radiation from argon lasers is being intensively evaluated for applications in data storage, hologram illumination, as a source in scattering experiments, and in bright displays (the last-mentioned sometimes in conjunction with spectral neighbors in the yellow-green and red, as obtained from ionized krypton lasers).

The list is almost endless, but let us observe, in conclusion, that even with its present important applications, the definitive task for this laser may be still to come.

The author acknowledges with thanks the efforts of Dr. R. C. Miller, Bell Telephone Laboratories, who supplied the material condensed in the sections "Inversion mechanism in the argon-ion laser" and "Influence of an axial magnetic field."

REFERENCES

1. Bell, W. E., "Visible laser transitions in Hg+," *Appl. Phys. Letters*, vol. 4, p. 34, Jan. 15, 1964.

2. Mathias, L. E. S., and Parker, J., "Stimulated emission in the band spectrum of nitrogen," *Appl. Phys. Letters*, vol. 3, pp. 16–18, July 1, 1963.

3. Boot, H. A. H., Clunie, D. M., and Thorn, R. S. A., "Pulsed laser operation in a high-pressure helium–neon mixture," *Nature*, vol. 198, p. 773, May 25, 1963.

4. Bennett, W. R., Jr., "Inversion mechanisms in gas lasers," *Appl. Opt.: Supplement on Chemical Lasers*, pp. 3–33, 1965.

5. Convert, G., Armand, M., and Martinot-Lagarde, P., "Effet laser dans des mélanges mercure–gaz rares," *Compt. Rend.*, vol. 258, p. 3259, Mar. 23, 1964.

6. Bridges, W. B., "Laser oscillation in singly ionized argon in the visible spectrum," *Appl. Phys. Letters*, vol. 4, pp. 128–130, Apr. 1, 1964.

7. Bennett, W. R., Jr., Knutson, J. W., Jr., Mercer, G. N., and Detch, J. L., "Super-radiance, excitation mechanisms, and quasi-CW oscillation in the visible Ar+ laser," *Appl. Phys. Letters*, vol. 4, pp. 180–182, May 15, 1964.

8. Gordon, E. I., Labuda, E. F., and Bridges, W. B., "Continuous visible laser action in singly ionized argon, krypton, and xenon," *Appl. Phys. Letters*, vol. 4, pp. 178–180, May 15, 1964.

9. Horrigan, F. A., Koozekanani, S. H., and Paananen, R. A., "Infrared laser action and lifetimes in argon II," *Appl. Phys. Letters*, vol. 6, pp. 41–43, Feb. 1, 1965.

10. Cheo, P. K., and Cooper, H. G., "Ultraviolet ion laser transitions between 2300 and 4000 Å," *J. Appl. Phys.* vol. 36, pp. 1862–1865, June 1965.

11. Fowles, G. R., and Silfvast, W. T., "High-gain laser transition in lead vapor," *Appl. Phys. Letters*, vol. 6, p. 236, June 15, 1965.

12. Labuda, E. F., Gordon, E. I., and Miller, R. C., "Continuous-duty argon ion lasers," *IEEE J. of Quantum Electronics*, vol. QE-1, pp. 273–279, Sept. 1965.

13. Bickerton, R. J., and von Engel, A., "The positive column in a longitudinal magnetic field," *Proc. Phys. Soc.* vol. 69B, pp. 468–481, 1956.

14. Gordon, E. I., and Labuda, E. F., "Gas pumping in continuously operated ion lasers," *Bell System Tech. J.*, vol. 43, p. 1827, July 1964.

15. Heavens, O. S., "Optical masers," *Appl. Opt. (Supplement on Optical Masers)*, pp. 1–23, 1962.

16. Rigrod, W. W. and Bridges, T. J., "Bistable traveling-wave oscillations of ion ring laser," *IEEE J. of Quantum Electronics*, vol. QE-1, pp. 298–303, Oct. 1965.

17. Private communication.

18. Maecker, H., "Ein zylindrischer Bagen Für hohe Leistungen," *Z. Naturforschung*, vol. 11A, pp. 457–479, June 1956.

19. Shumaker, J. B., Jr., "Arc source for high-temperature gas studies," *Rev. Sci. Instr.*, vol. 32, pp. 65–67, Jan. 1961.

20. Emmons, H. W., "Recent developments in plasma heat transfer," in *Modern Developments in Heat Transfer*, W. Ibele, ed. New York: Academic Press, 1963.

21. Hsieh, H., Private communication.

22. Bell, W. E., "Ring discharge excitation of gas ion lasers," *Appl. Phys. Letters*, vol. 7, p. 190, Oct. 1, 1965.

23. Private communication.

24. Segré, J. P., "Use of lasers for microcircuit resistance trimming," *Proc. Nat'l Electron. Conf.*, pp. 48–51, 1965.

25. Hoye, R., and Minton, J., "Noble gas ion laser as a light-knife," *Surg. Forum*, vol. 16, pp. 93–95, 1965.

26. King, M., and Kainer, S., "Some parameters of a laser-type beyond-the-horizon communication link," *Proc. IEEE*, vol. 53, pp. 137–141, Feb. 1965.

BIBLIOGRAPHY

Bridges, W. B., and Chester, A. N., "Spectroscopy of ion lasers," *IEEE J. of Quantum Electronics*, vol. QE-1, pp. 66–83, May 1965.

IEEE Reports for 1965

Introduction

William G. Shepherd *President IEEE*

Each year at this time, it is the privilege and the duty of the President of the IEEE to report to the members on the state of our society. As in years before, the reports of the Secretary and the Treasurer, submitted here, provide the facts and statistics about the society. However, in addition to these facts and statistics, I should like to sketch out what I believe were the main features of last year's activities, to assess their significance, and to highlight a few of the unifying and guiding principles underlying the many-faceted aspects of the Institute's activities.

In general, the major events may be described on a number of levels—those changes occurring in our organizational structure, those changes in the content of our program of publications, and those changes occurring in the technical fields in which IEEE members exercise their keenest professional interest. The changes occurring on all of these levels are mutually interdependent.

Organization after the merger

We can at last say that the act of merger is complete. There have been problems of membership, mergers of Committees, Groups, Sections, and so on. These, in turn, have raised questions of communication, coordination, and cooperation on an unprecedented scale. Some members have had at least a slight brush with some of these problems, if only at the level of discovering that they had not been included on new mailing lists. However, our preoccupation with the massive mechanics of merger is now over. The machinery is now effectively installed, and the way has been cleared to bring our major energies to other than procedural matters.

Although many of the organizational details are cited in section C of the Secretary's report, I should like to point up one that I think is particularly important: namely, the emergence during 1965 of the Operations Committee (OpCom) of the Technical Activities Board (TAB) as a strong administrative and operational organism. Although this role for OpCom was not originally intended in its formation, the discovery that it could carry out this function is quite important, for inasmuch as OpCom is open to the Chairmen of the Groups and General Committees through the larger Technical Activities Board, it provides a clear mechanism for communication between these entities and the Executive Committee. In particular, Group members have an effective channel for bringing problems unique to the Groups up to the highest executive level of the Institute.

The publications program

Chief among our present concerns is the increased emphasis on improving the services for members. Thus, in the past year, we have begun to make an objective reappraisal of the *total* publications program of the Institute. Our publications goal has always been to publish the most advanced, the most authoritative, and the best references on developments in which IEEE members have a vital professional interest. Clearly, along with symposia in which the most up-to-the-minute developments are reported, the IEEE publications constitute one of the most obvious benefits of membership. Some highlights of our publications activities are as follows:

During the past year, new editorial boards were set up for SPECTRUM, for the PROCEEDINGS, and for the STUDENT JOURNAL, through which members could be given a stronger voice in the editorial content and direction of these core publications.

The Group memberships and Group publications have come under a closer scrutiny, and studies of their effectiveness continue unabated into this present year.

A Publications Board was created.

A new type of publication, the JOURNAL OF QUANTUM ELECTRONICS, sponsored by two Groups, but not coming directly under either Group structure, has been launched experimentally.

More technical meetings were sponsored than in previous years, and, most significant, these meetings sheltered a greater number of interdisciplinary panel discussions than ever before.

Other direct manifestations of our increased concern over the achievement of a most effective and flexible publications program appear in the appointments last year of Ralph H. Flynn as Director of Publishing Services, of Alexander A. McKenzie as Assistant to the Director of Editorial Services, and of J. M. Kinn as Director of Educational Services, all of which are newly created posts within the Institute structure.

Furthermore, in 1965, the policies were initiated that were to lead to the creation in January 1966 of a new official position, Vice President of Publication Activities, and of a new Editor for SPECTRUM.

These are but a few indications of the recognition that emerged in the past year and that now prevails throughout the Institute, namely, that the Institute publications can no longer be viewed as isolated bits, with policies and programs formulated by nearly autonomous Groups. The editorial content of many of the publications has been found to be uneven, and not always up to the best standards of our profession. But, more important, with the tide of interdisciplinary research and engineering growing stronger with each passing year, there arises continually now the danger that the separated editorial policies will find some Groups overlapping in the areas of their interest, and, worse, will leave certain new areas of development not served adequately by any existing Group. The JOURNAL OF QUANTUM ELECTRONICS, previously mentioned, was formed for precisely these reasons. It represented a way in which the Institute could act quickly and flexibly in response to a new need. If this new type of experimental Institute publication continues to develop, and proves to be a successful instrument, it may well serve as a model for other such publications as the need arises.

In light of present developments, the publication policies, practices, and standards of the existing 31 Groups cannot be regarded in effect as wholly autonomous activities. There has come about the realization that TAB, especially through OpCom, and the Publications Board must begin to act jointly to resolve the significant publication problems as they emerge and crystallize, and to resolve such questions as speedily and flexibly as possible in order to serve the needs of the entire membership. To this end, TAB and the Publications Board have begun an intensive review of all publication policies, as outlined in the report of the Secretary.

Better services for members

Tightly interlocked with our concern for the most effective total publications program is the question of the *true needs* of all of our membership. We need to know the needs of our members, as these needs change and evolve, in order to provide them with the most effective and appropriate services *at the time their needs are most pressing.* Ideally, Headquarters should have a technical interest profile of every single member of the Institute from students on up; and indeed, in the future, as our computer facility gains in its effectiveness, such a profile may be possible. Even our preliminary studies this past year into our membership has brought the recognition that there are essentially "two populations" of engineers within the IEEE—those who are active in Groups and those who are not. It has also been recognized that the Group members are also the most active in the Institute; they write more papers, they go to more meetings, and so on. In effect, they are not only making the greatest contributions to IEEE but they also *seem* to be deriving the greatest benefits from their IEEE membership.

We have before us a serious question, then, of whether or not we are doing enough for the non-Group members. Non-Group members have, as the major vehicles serving them, SPECTRUM, PROCEEDINGS, the STUDENT JOURNAL, the large, yearly regional meetings, and the more local Section programs. We want to be sure we are providing programs that meet their needs. The activities of many engineers are general rather than special. Therefore, we also need a broad publication such as SPECTRUM.

But we are not satisfied with our own assumptions about how well we are serving the non-Group members, and we have therefore initiated in this past year a number of studies that relate to the achievement of a more effective, more flexible, and a more responsive total editorial program.

Other continuing activities

There are, of course, many other special activities in which we have been engaged in the past year. Many of these are by necessity continuing into the present year—the self-appraisals undertaken by the Groups, the merger of the Groups and Technical Committees (Standards), the reappraisal of our relations with other professional societies, our reviews of the patterns of our technical meetings, our policies with regard to the publication of conference papers, our cultural and scientific exchanges with the U.S.S.R. and other Eastern European nations, and many others.

Not least among our deepest concerns are our programs and relations with Student members, our progressive involvement in problems of continuing education, and our careful examination of our proper roles in the questions relating to the mechanization of information retrieval. This last sits as a problem of the profoundest dimensions across the whole spectrum of IEEE activities. Both creators and victims of the information revolution, all professional individuals and societies must now bring their best resources to the resolution of the multitudinous questions involved, questions that we have all been debating and shall continue to work at in the years to come. The evolution of a consistently effective and high-quality publications program is *one* avenue that, in the long run, can go some way toward alleviating the intensity of these problems in other areas. But it is only one, and its interrelationship to the others must be kept in view.

The membership will recall that in transmitting the Secretary's report for 1964 President Oliver commented on two problems with major financial implications for the Institute. One concerned the position taken by the Internal Revenue Service that the IEEE is retroactively subject to taxation on the net income realized from the exhibition that forms an integral part of the International Convention. This matter is still pending.

The second matter was concerned with the reduction in advertising income. This reduction in income has continued and now appears to be a fact of life resulting from a change in U.S. Governmental policies with respect to allowable costs on contracts. We can expect a continued growth in these revenues but from a reduced base. As a consequence, the committee charged with a study of the long-range financial trends of Institute operations has made recommendations to the Board of Directors, which are detailed elsewhere in these pages.

Conclusion

As I said earlier, all of the problems facing IEEE are interlocked, so that it is not possible to discuss one problem for long before bringing up its connection to another. The test of a viable organization is its capacity to respond appropriately to the real core problems and needs of the day, not merely to shadowbox with superficial symptoms and complaints. I believe we have a viable organization, but it is now a very large one; and inasmuch as it is, each of us must ask himself whether or not he is doing all that he can for it, as well as for himself. He cannot, in fact, serve one without serving the other.

The reports of the Secretary and Treasurer and the audit report are submitted herewith. I hope that the few themes I have tried to sound here, and that these reports reflect, will give each member a little more appreciation of the direction that the IEEE is taking.

Report of the Secretary—1965

Haraden Pratt

To the Board of Directors
The Institute of Electrical and
Electronics Engineers, Inc.

Gentlemen:

The Report of the Secretary for the year 1965, the third year of operations of the merged IEEE, is presented herewith.

Membership has grown slightly during the year. There was a decline in the number of Associates, following a trend during the past few years. The number of Students remains about the same, though considerably lower than in 1963. A 10 percent increase in Group members is noted, but Group Affiliate memberships dropped to about one half the number at the end of year 1964.

Your attention is called to the increase of about 3 percent in the total number of editorial pages published by the Institute during 1965.

Attendance at the 1965 International Convention was about 8 percent less than in 1964, the number of exhibitors also was down 8 percent, but the number of technical papers presented increased about 14 percent.

The number of IEEE- and jointly sponsored technical meetings went up 10 percent to a total of 95.

These data indicate a general growth trend with increased services to members. Organizationally, the Headquarters management appears to have achieved an increased effectiveness.

Respectfully submitted,
Haraden Pratt
Secretary

Section A—Membership

Table I gives a two-year comparison of the distribution of the membership by grade and by percentage and Table II gives the geographical distribution of the membership.

Section C—Report of technical activities (Groups and related activities)

1. Planning initiated early in 1964 led to the establishment of a Technical Activities Board (TAB), effective January 1, 1965. TAB assumed the responsibilities and functions formerly residing with the Groups Committee and Technical Operations Committee. TAB membership consists of its Chairman and Vice Chairman, eight members-at-large, the Chairman of each IEEE General Committee and Group, liaison or *ad hoc* appointments by the TAB Chairman, and a Staff Secretary. Because of the large size of TAB, an Operating Committee (OpCom) was established consisting of only the officers and members-at-large to carry forward TAB work in the intervals between TAB meetings. Accordingly, in 1965 TAB met four times: February 2, May 5, August 23–24, and November 16. OpCom met on February 3, March 12, April 9, May 4, June 10, July 8, August 23, September 29, November 15, and December 15.

2. In the absence of precedents, TAB had many problems in 1965, the answers to which will apply with minor modifications in future years. For each major problem a task force was established by the TAB Chairman as follows:

65-A. To draft rules and procedures for TAB and OpCom, including a plan for the nomination and appointment of the nonelective memberships on TAB.

65-B. To respond to a request by the Executive Committee for a review and evaluation of the roles, missions, accomplishments, and plans of the General Committees and Groups.

65-C. To examine the problems, real or imagined, that seemed to hinder the merger of Groups and Technical Committees (Standards).

65-D. To resolve differences appearing among several Groups having interests in systems.

65-E. To determine a basis for selection of IEEE representatives to an outside activity when one Group (or Committee) has a primary interest but others have secondary interests.

65-F. A case parallel to that responsible for 65-E.

65-G. To review objectives and trends of the Technical Program for the IEEE Convention and to recommend how the program might be improved.

65-H. To review the financial structure of the Groups.

65-I. To examine the Group member/nonmember ratio for all IEEE, and related recruitment problems, and to recommend means of improvement.

65-J. To examine the quality, timeliness, and support for the Group Transactions.

65-K. To examine the existing awards sponsored or administered by Groups and to recommend how to improve the Group award structure.

65-L. To examine the present pattern of technical meetings and to establish criteria either for the continuation of an existing series of meetings or for the approval of a new meeting.

65-M. To review IEEE activities in the technical area usually identified as radio science and to propose ways to improve the IEEE programs.

65-N. To review policies and practices pertaining to financial assistance for overseas participants in conferences held overseas.

65-O. To respond to an Executive Committee request to provide unified bylaws for all the General Committees.

3. In addition to the above undertakings, TAB has joined with the Publications Board in a study of conference publication practices. TAB has worked with the Intersociety Relations Committee in developing a list of representatives or appointments to outside organizations and in clarifying the TAB responsibilities with respect to these positions. An *ad hoc* committee, established by the Executive Committee to review the IEEE position with respect to new technologies, submitted a report that bears directly on the activities of several TAB entities, particularly the Group on Education, the Group on Engineering in Medicine and Biology, and the General Committee on New Technical Activities.

4. The Joint Technical Advisory Committee (JTAC). The Joint Technical Advisory Committee, sponsored by IEEE and EIA (Electronics Industries Association), held five meetings during 1965. Five JTAC Subcommittees were active during the year, the Subcommittee on Future Needs and Uses of the Spectrum (65.1) having been established in September 1965.

Radio Spectrum Utilization—Subcommittee 62.1. Under the chairmanship of Philip F. Siling, the JTAC volume, *Radio Spectrum Utilization*, was completed in late 1964 and released in early 1965.

Electromagnetic Compatibility—Subcommittee 63.1. Since 1961 the Joint Technical Advisory Committee has urged the stimulation of interest of engineers associated with nonmilitary technology in the increasing technical and economic problems arising from electromagnetic compatibility of equipment. In November 1963, the JTAC established a Subcommittee on Electromagnetic Compatibility, under the chairmanship of Richard P. Gifford, to undertake a study of needed technical areas and to formulate objectives for dealing with this subject.

In addition to the foregoing subcommittees, there exist the following: Microwave Radio Relay System Reliability—Subcommittee 63.2; the JTAC Subcommittee on Microwave Radio Relay System Reliability, under the chairmanship of Dr. William H. Radford; Mobile Radio Services—Subcommittee 64.1; and Future Needs and Uses of the Spectrum—Subcommittee 65.1.

5. Intersociety Relations Committee. The Intersociety Relations Committee met five times during 1965. On recommendation by the ISRC, the IEEE appointed more than 50 representatives to 18 organizations during 1965.

One of ISRC's major efforts is the continuing evaluation of IEEE participation in the activities of outside organizations. A great deal of progress has been made. A study aimed at developing meaningful guidelines for such evaluation was started in the latter part of 1964 and has continued to the present. The study is expected to be

I. IEEE membership by grade, by percentage (two-year comparison)

Grade	December 31, 1964 No.	December 31, 1964 Percentage	December 31, 1965 No.	December 31, 1965 Percentage
Honorary (H)	8		6	
Fellow (F)	2 595	1	2 681	2
Senior Member (SM)	26 182	17	26 387	17
Member (M)	87 435	57	87 916	57
Associate (A)	13 682	9	12 949	8
Student (S)	24 076	16	24 259	16
Total	153 978	100	154 198	100

II. IEEE membership by region

Region	Total	Percentage
1	38 522	25
2	25 854	17
3	13 999	9
4	17 941	12
5	12 982	8
6	30 235	20
Subtotal	139 533	91
U.S. overseas military	736	0.5
U.S. possessions	12	
Subtotal	748	0.5
7	6 825	4.5
8	3 513	2
9	3 579	2
Subtotal	13 917	8.5
Grand total	154 198	100

completed during 1966. As an outgrowth of this continuing study, a method of appointing IEEE representatives to outside organizations was formulated and adopted during 1965. Included in this newly adopted method is a delineation of TAB responsibility for appointment of representatives.

Also during 1965, ISRC established a Subcommittee on Cultural and Scientific Exchanges between the U.S. and the Soviet and Eastern European nations. This new subcommittee will attempt to ensure that proper and effective channels of communication are maintained between the delegates of IEEE and scientific and technical societies existing with the Eastern and Soviet bloc nations. A specific task, in particular, involves making appropriate arrangements with the Popov Society.

Section D—Report of publication activities, 1965

General. The year 1965 was one of reorganization as well as considerable editorial activity. During the year a record number of pages were published, new journals were added, some existing journals were merged, and several steps were taken to improve the publications management structure of the Institute. The state of the

IEEE publications program at the end of the year may be summarized as vigorous and expanding.

The total 1965 IEEE editorial program, excluding translated journals, resulted in the publication of 2888 papers and 1276 letters making up a total of 29 187 editorial pages. During 1964, the total editorial pages published numbered 25 696.

In August, the Board of Directors made several changes in the IEEE Bylaws that affected the publications operations of the Institute. The IEEE Editorial Board was replaced by the Publications Board in recognition that the functions required of this Board had expanded into several areas other than those of a strictly editorial nature. As well as editorial policies, the Publications Board is concerned with problems such as the implications of automated procedures for handling information, the IEEE translated journals program, the improvement of abstracting and indexing procedures, the appropriate role of advertising in IEEE publications, and the overlap between publications. To make it representative of all publications, the new Board includes two Group Editors, the Editors of the STUDENT JOURNAL and the PROCEEDINGS, and will include the Editor of SPECTRUM. Besides the Chairman and Vice Chairman, three to six members other than the above will serve on the Publications Board. Late in the year the voting members of the Institute approved a change to the IEEE Constitution giving the Chairman of the Publications Board the title of Vice President, Publication Activities, instead of Editor of the IEEE.

In a further change, the PROCEEDINGS and SPECTRUM are now to be organized with an appointed Editor and an Editorial Board in order to strengthen these publications and make them more responsive to members' needs. This change has not yet been implemented for SPECTRUM.

Another modification to the Bylaws formalized the Panel of Group Editors as a standing committee under the Publications Board. This Panel, whose membership consists of the Editors of all of the Group publications, is a mechanism for providing advice on and for implementing overall editorial policies. It began operation by forming subcommittees to examine TRANSACTIONS schedules, paper-processing procedures, and information retrieval matters.

The new Bylaws also provided for the establishment of a new committee, the Information Processing Committee, to operate as a policy and technical advisory body to the Publications Board in this area of rapidly growing importance. The present Information Retrieval Committee, meanwhile, remains as an *ad hoc* working committee in this field to develop and propose new plans and procedures.

The Publications Board and the Technical Activities Board took significant steps toward collaborating on matters of mutual interest through the establishment of two joint task forces, one to explore broad questions related to the Group TRANSACTIONS and the other to examine problems associated with conference publications.

Late in the year Ralph H. Flynn joined the Headquarters staff as Director of Publishing Services. In this newly created position, Mr. Flynn is responsible for directing all business and advertising aspects of the Institute's publications. Alexander A. McKenzie was appointed in April to the new post of Assistant to the Director of Editorial Services. His duties consist of supervising the translated journals program, acting as staff liaison for IEEE information processing activities, and assisting with various other editorial operations. Shortly after receiving his electrical engineering degree in June, Robert R. Beck joined the editorial staff as Assistant Editor of the STUDENT JOURNAL.

IEEE Spectrum. With the close of 1965, IEEE SPECTRUM successfully completed its second year of publication. A total of 81 articles was published during 1965, as compared with 69 in 1964. Although most articles were invited or contributed, 14 were staff written. In addition to the appearance of articles and the regular departments, a program was initiated of publishing technical reports on important conferences in the feature section of SPECTRUM in order to give the results of this important IEEE activity much broader dissemination.

All told, 2204 pages were published in SPECTRUM during the year, of which 1395 were devoted to technical and editorial material and 809 consisted of advertising and related matter.

Proceedings of the IEEE. With SPECTRUM firmly established as the "core" publication of the IEEE, the PROCEEDINGS OF THE IEEE became more clearly defined as the appropriate medium for papers of broad interest written at technical levels as high as the subjects warrant. With the role of the PROCEEDINGS thus clarified, the IEEE Publications Board took action during the year to form and implement a new program for this research-oriented journal. The most important step was the appointment of Prof. M. E. Van Valkenburg of the University of Illinois as the Editor of the PROCEEDINGS; a distinguished PROCEEDINGS Editorial Board was appointed late in 1965 to assist and advise Professor Van Valkenburg. Other innovations planned and partly effected during the year were an increase in the number of special issues, the inclusion of high-level invited papers from leading specialists, and a decrease in the time between receipt and publication of correspondence.

The PROCEEDINGS year was highlighted by the appearance of three large special issues devoted to the subjects of radar reflectivity, ultrasonics, and nuclear test detection.

The number of papers submitted for publication consideration for regular issues during 1965 and their disposition were as follows: of the 167 manuscripts received, 33 percent were accepted, 33 percent were referred to TRANSACTIONS Editors for consideration, and 34 percent were rejected. The number of letters published increased to 692 (as compared with 513 for 1964), and 592 pages were devoted to this quick method of announcing research results. During the year, the PROCEEDINGS carried a total of 2239 editorial pages and 421 pages of advertising and other noneditorial material, making a grand total of 2660 pages published.

Group Publications. As in previous years, the Group TRANSACTIONS encompassed the major share of the IEEE publication output. During the year 144 issues were published, totaling 15 356 pages.

At the beginning of the year there were 32 TRANSACTIONS being published. During 1965 three new Groups began publishing TRANSACTIONS, viz., the Electrical Insulation Group, the Magnetics Group, and the Systems Science and Cybernetics Group. As a result of Group mergers, the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS replaced the TRANSACTIONS of four Groups (Aeronautical and Navigational Electronics, Aerospace, Military Electronics, and Space Electronics and Telemetry) and the IEEE TRANSACTIONS ON PARTS, MATERIALS AND PACKAGING replaced the TRANSACTIONS of two Groups (Component Parts, and Product Engineering and Production). With the encouragement of the Editorial Board, the IEEE JOURNAL OF QUANTUM ELECTRONICS began publication in April under the joint sponsorship of the Electron Devices Group and the Microwave Theory and Techniques Group.

The net result of the mergers and the addition of new publications was that the Group publications at the end of the year numbered 32, the same as at the beginning of the year.

Three Group publications are now issued monthly: the TRANSACTIONS of the Power and the Electron Devices Groups and the IEEE JOURNAL OF QUANTUM ELECTRONICS.

IEEE Student Journal. The bimonthly student magazine of the IEEE continued to serve the Student membership during 1965. In its six issues it provided its nearly 25 000 readers with 312 pages of technical and other career information, including 44 articles. An additional 20 000 copies of the September issue were sent to IEEE Counselors in colleges and technical institutes for distribution to incoming students who were potentially interested in electrical and electronics engineering and the IEEE.

Translated Journals. The IEEE maintained its program of translating and publishing papers from four Russian and two Japanese technical journals. This program is carried out with the support of the National Science Foundation. The year 1965 saw the appearance of 5678 pages translated from the Russian and 3654 pages from the Japanese. Advanced tables of contents of issues to be translated were carried in SPECTRUM, as were signed critical reviews of selected papers that had been published.

Special publications and preprints. The papers presented at the 1965 International Convention held in New York in March resulted in publication of a 13-part *Convention Record*, totaling 2658 pages.

In addition, 11 special Conference Records were issued, comprising 222 papers and 2276 pages. Finally, a total of 294 papers, amounting to 4952 pages, were individually preprinted by photo-offset means for five meetings.

Report of the Treasurer—1965

S. W. Herwald

The following statements of our financial condition have been certified by Price Waterhouse & Co., based on their audit of our operations in 1965. They are presented in a revised format that is designed to display our activities on a functional basis and to establish clearly the source and utilization of funds. This new approach has been carried forward in the establishment of our 1966 budget as an aid in maintaining control over operations in relation to available income.

The Comparative Statement of Financial Position shows the disposition of the 79th Street property and the addition of the proceeds to our investment portfolio. This sale has permitted us to consolidate all our operations at the United Engineering Center. During 1965, the IEEE Computer Center, planned and tested in prior years, became operational. While there were, as is usual, some initial operating difficulties, we have achieved more comprehensive statistics, faster processing of data, and other expected benefits. In particular, we have been able to exercise sharper control of print orders and publication inventories. Future computer operations will generate additional cost savings.

The Comparative Statement of Income and Operating Fund for 1964 and 1965 shows a substantial decline in income, due almost entirely to a sharp reduction in advertising revenue. Operations thus far in 1966 have shown improvement in advertising income, but as pointed out elsewhere in this issue by President Shepherd, it is doubtful that this source can be relied upon to the extent that it has in the past. The other items of income were relatively stable. The expenses show variations that can be primarily traced to (1) reduction of advertising printing costs and (2) shifting of certain publications and conference items to Group operation. Costs for administration have been controlled to the extent that 1965 showed a slight reduction in this area. However, an offsetting effect due to continuing inflation must be anticipated.

Since investment of the proceeds from the sale of the 79th Street property represented a major addition to our investment portfolio, a complete review of our holdings was undertaken. Investment counsel was instructed not only to select investments for the newly available funds, but also to review all former holdings. Additions and changes made as a result of this analysis have created a balanced fund that provides both income and growth potential. During 1965 the income and appreciation of the fund exceeded the national indices. Since the investment market is subject to constant change, we have inaugurated a system of continuing review of our holdings to provide maximum return to the Institute and its members.

Price Waterhouse & Co. audit report

Price Waterhouse & Co.

60 Broad Street
New York 10004
March 1, 1966

To the Board of Directors of
The Institute of Electrical and Electronics
Engineers (Incorporated)

In our opinion, subject to the final determination of the Institute's income tax liability, if any, as referred to in Note 3, the accompanying statement of financial position, the related statement of income and operating fund and the statement of changes in restricted funds present fairly the financial position of The Institute of Electrical and Electronics Engineers (Incorporated) at December 31, 1965 and the result of its operations for the year, in conformity with generally accepted accounting principles applied on a basis consistent with that of the preceding year. Our examination of these statements was made in accordance with generally accepted auditing standards and accordingly included such tests of the accounting records and such other auditing procedures as we considered necessary in the circumstances.

Price Waterhouse & Co.

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS (INCORPORATED)

COMPARATIVE STATEMENT OF INCOME AND OPERATING FUND

	For the year ended December 31, 1965	1964*
Income:		
Membership, entrance fees and dues	$1,902,669	$1,912,743
Advertising	997,188	1,392,744
Periodicals subscriptions	721,618	878,181
Other publications and sales items	325,666	271,729
Convention and technical conferences	1,111,347	1,127,145
Investments, including gain on sale of securities	237,846	210,090
Miscellaneous other	17,685	31,766
Total income	5,314,019	5,824,398
Expenses:		
Headquarters services to members	815,776	676,088
Support of sections and branches	401,798	388,899
Support of groups	812,435	734,357
Periodicals publication	1,541,897	1,732,029
Other publications and sales items	329,703	288,827
Convention and technical conferences	824,519	941,231
General administration	831,825	839,267
Total expenses	5,557,953	5,600,698
Excess of income over (under) expenses for the year	(243,934)	223,700
Less—Special charges:		
Provision for computer installation and record conversion		196,000
Provision for relocation and disposal costs		167,000
Contribution to United Engineering Trustees, Inc.		175,000
		538,000
Income (expenses) for the year, and in 1964, special charges transferred to operating fund	(243,934)	(314,300)
Operating fund balance, January 1	3,061,399	3,375,699
Operating fund balance, December 31	$2,817,465	$3,061,399

*Certain amounts have been reclassified for comparative purposes including salaries of $1,589,579. Salaries in 1965 were $1,642,880.

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS (INCORPORATED)

COMPARATIVE STATEMENT OF FINANCIAL POSITION

	December 31, 1965	December 31, 1964
Operating Fund		
Current assets:		
Cash including $600,000 in savings accounts (1964—$725,369)	$1,066,575	$1,228,256
Marketable securities, at cost, market value $4,617,600 (1964—$3,485,000)	3,467,152	2,469,762
Notes and accounts receivable, less doubtful accounts	338,429	211,673
Prepaid expenses, inventory, etc.	254,292	279,668
Total current assets	5,126,448	4,189,359
Less—Current liabilities:		
Accounts and accrued expenses payable	663,102	704,448
Funds held for the use of professional groups	530,887	407,114
Deposit received on sale of real property (Note 4)		215,000
	1,193,989	1,326,562
Deferred income:		
Dues	794,371	1,068,204
Subscriptions	255,489	408,819
Convention	663,766	620,575
	1,713,626	2,097,598
Total current liabilities	2,907,615	3,424,160
Working capital	2,218,833	765,199
Note receivable, 6%, installments due after 1966	116,656	121,656
Fixed assets:		
Land and buildings, at cost (Note 4)		1,641,455
Office equipment and leasehold improvements, at cost, less accumulated depreciation and amortization $221,661 (1964—$336,958)	481,976	533,089
	481,976	2,174,544
Operating fund balance (accompanying statement)	2,817,465	3,061,399
Property Fund		
Advance to United Engineering Trustees, Inc. (Note 2)	265,000	265,000
Restricted Funds		
Cash	74,257	63,797
Marketable securities, at cost, market value $154,395 (1964—$148,300)	98,231	98,231
Restricted funds balance (accompanying statement)	172,488	162,028
Total funds	$3,254,953	$3,488,427

STATEMENT OF CHANGES IN RESTRICTED FUNDS FOR THE YEAR ENDED DECEMBER 31, 1965

Restricted fund	Fund balance January 1, 1965	Receipts from contributions and marketable securities	Disbursements for awards and related costs	Fund balance December 31, 1965
Life Member Fund	$ 64,395	$13,374	$2,518	$ 75,251
International Electrical Congress—St. Louis Library Fund	6,686	138	138	6,686
Edison Medal Fund	12,912	1,302	553	13,661
Edison Endowment Fund	8,521			8,521
Lamme Medal Fund	9,361	345	315	9,391
Mailloux Fund	1,091	47	100	1,038
Volta Memorial Fund	19,200	728	2,725	17,203
Kettering Award Fund	2,316	60		2,376
Browder J. Thompson Memorial Prize Award Fund	5,620	125		5,745
Harry Diamond Memorial Prize Award Fund	1,110	40		1,150
Vladimir K. Zworykin Television Award Fund	5,345	112	500	4,957
W. R. G. Baker Award Fund	8,978	400		9,378
William J. Morlock Award Fund	5,229	213		5,442
W. W. McDowell Award Fund	10,459	425		10,884
William D. George Memorial Fund	805			805
Total	$162,028	$17,309	$6,849	$172,488

NOTES TO FINANCIAL STATEMENTS

NOTE 1: The unfunded past service liability at December 31, 1965 for pensions of employees is estimated to be approximately $200,000.

NOTE 2: In accordance with a Founder's agreement between the Institute and the United Engineering Trustees, Inc. the Institute has agreed to permanently maintain its principal offices in the United Engineering Center, which in 1966 will involve a lease payment of approximately $180,000. The $265,000 advanced to United Engineering Trustees, Inc. is repayable only out of available reserve funds on dissolution of United Engineering Trustees, Inc. and carries interest at an annual rate of 4%.

NOTE 3: In 1964 the Institute was notified by the Internal Revenue Service that the Institute's annual convention involved the operation by the Institute of an unrelated trade or business and that the net income therefrom, if any, is subject to federal tax. If the Service's position, which should be applicable to 1954 and subsequent years, were ultimately sustained, the amount of potential liability for taxes and interest would be material.

The Institute has filed both a tax return with respect to the test year 1954 and claims for refund in the amount of the taxes shown to be due thereon. Most of the claims for refund have been denied and the Institute presently intends to sue to recover the amounts in controversy.

NOTE 4: The sale of the land and buildings was completed in 1965 and the funds received from the sale are now included in working capital.