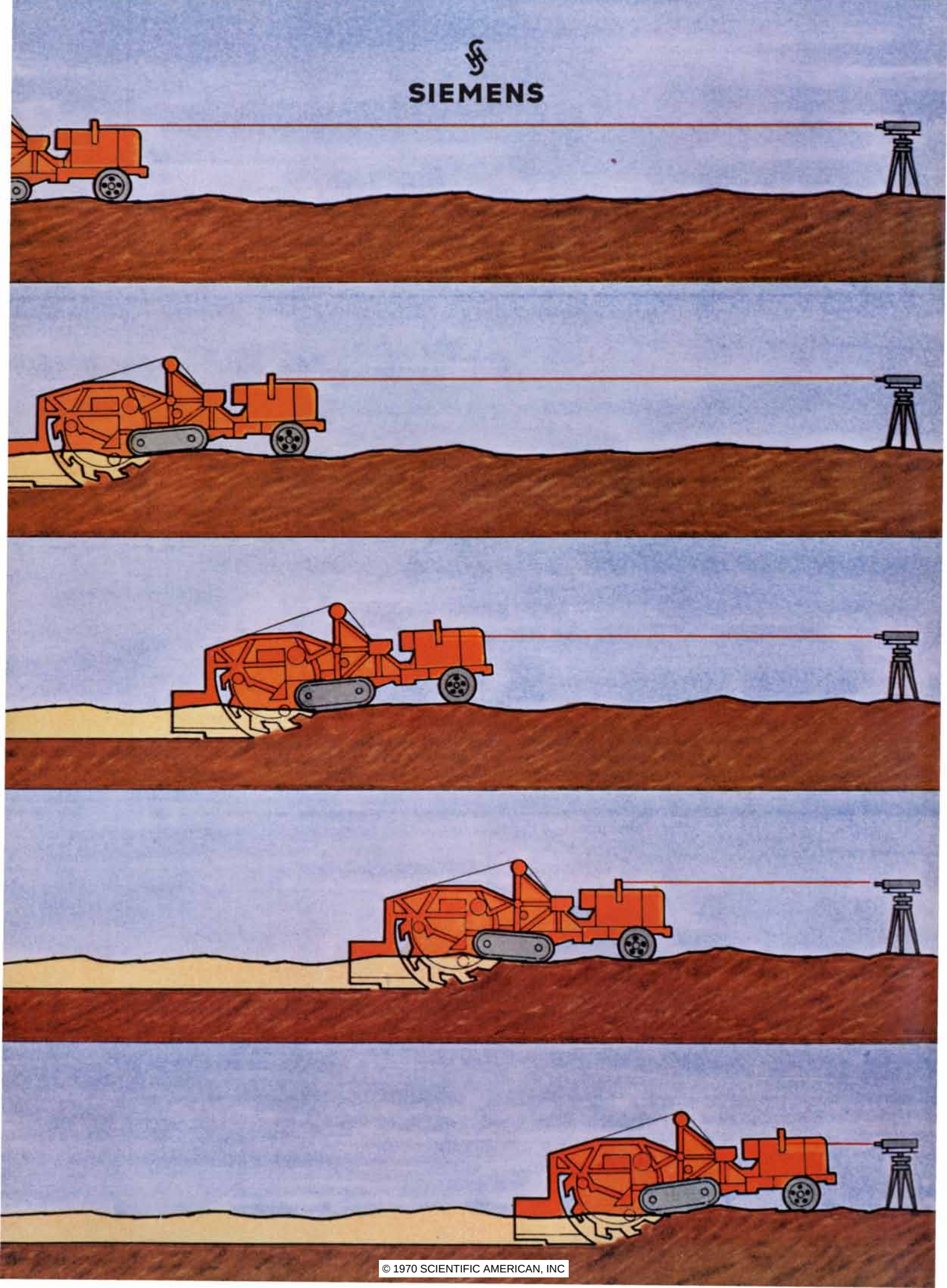
SCIENTIFIC AMERICAN

GLOBULAR-CLUSTER STARS

ONE DOLLAR

July 1970



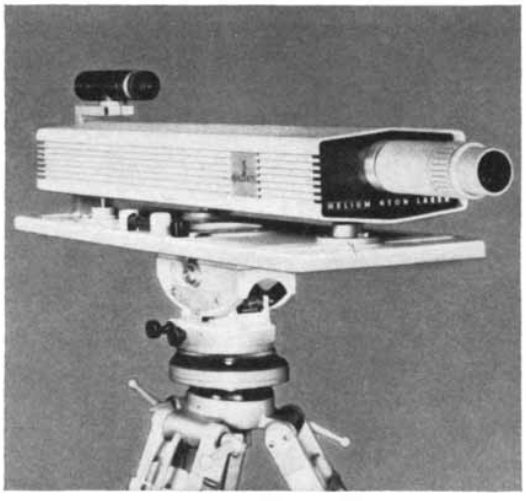
Siemens

A three billion dollar name in quality products. We are changing the laser from a laboratory curiosity to a practical industrial tool.

After a decade of publicity and problems, lasers are beginning to emerge from the laboratory as useful tools.

One of the reasons commercial success has been so elusive is that up till now, lasers have been designed for the controlled conditions of laboratory situations. Not for the unpredictable realities of the industrial world.

Siemens has overcome this problem by designing laser systems for the specific environments in which they are to be used. One of these happens to be one of the most rugged environments imaginable. Heavy construction.



Our LG-661 helium-neon laser system is now being used to provide continuous, highly stable operation for automatic alignment of the trench digging equipment illustrated at left. Similar systems are used in dredging, grading, tunneling, pipe laying, bridge building, asphalt paving and mining.

This completely portable system is weather-proof, shock-proof and dust-proof. And for mining applications, Siemens supplies an explosion proof housing.

As unique as this equipment may be, it's just one of a whole line of Siemens laser systems designed to perform a variety of useful jobs in many different situations.

Another Siemens helium-neon laser system, for example, has been developed for use as a secondary length standard approaching the primary length standard of Krypton 86. But it's also being used for alignment and measurement applications in numerical control of metalworking equipment.

Again, for the metalworking field, Siemens has developed a high-efficiency CO_2 laser system. This versatile system can be easily focused to produce a concentrated beam for micromachining, cutting, welding and drilling.

Siemens has also introduced compact YAG (yttrium-aluminum garnet) lasers of both continuous and pulse designs. These systems have medical applications including retina coagulation, and special surgical procedures.

Looking to the future, a number of our lasers are being tested for data transmission and communications applications where some of the biggest breakthroughs are probably yet to come. In fact, Siemens has in development state-of-the-art frequency multiplying lasers that may open a new range of telecommunications applications.

For more information, please write to M. A. Stamberger at Siemens Corporation, 186 Wood Avenue So., Iselin, N.J. 08830.

Siemens. A worldwide company in telecommunications, instrumentation, medical equipment, power engineering, electronic components and data systems.

Infonet is as close as the nearest telephone

A freeway in Los Angeles, a stockyard in Chicago, a boutique on Park Avenue: unlikely spots for a data processing terminal? Unlikely, yes. But very possible.

Because if you can get to a telephone, you can get INFONET computer timesharing service — whether your problem is one of engineering design, statistical treatment of experimental data, or simulation of a business enterprise.

Why use INFONET? Because our product is a full product — with an interactive BASIC language and high speed Remote Job Entry capability. And there is our customer support — responsive, comprehensive, broad — designed to meet all your data processing needs. Standing behind INFONET is Computer Sciences Corporation — 5,000 people world-wide specializing in the design, development and management of computer-based information systems.

Whatever your data processing problem, isn't it good to know that INFONET's products, service and support for your needs are the finest available? The total resources of the unquestioned leader in the field are as close as your nearest telephone.

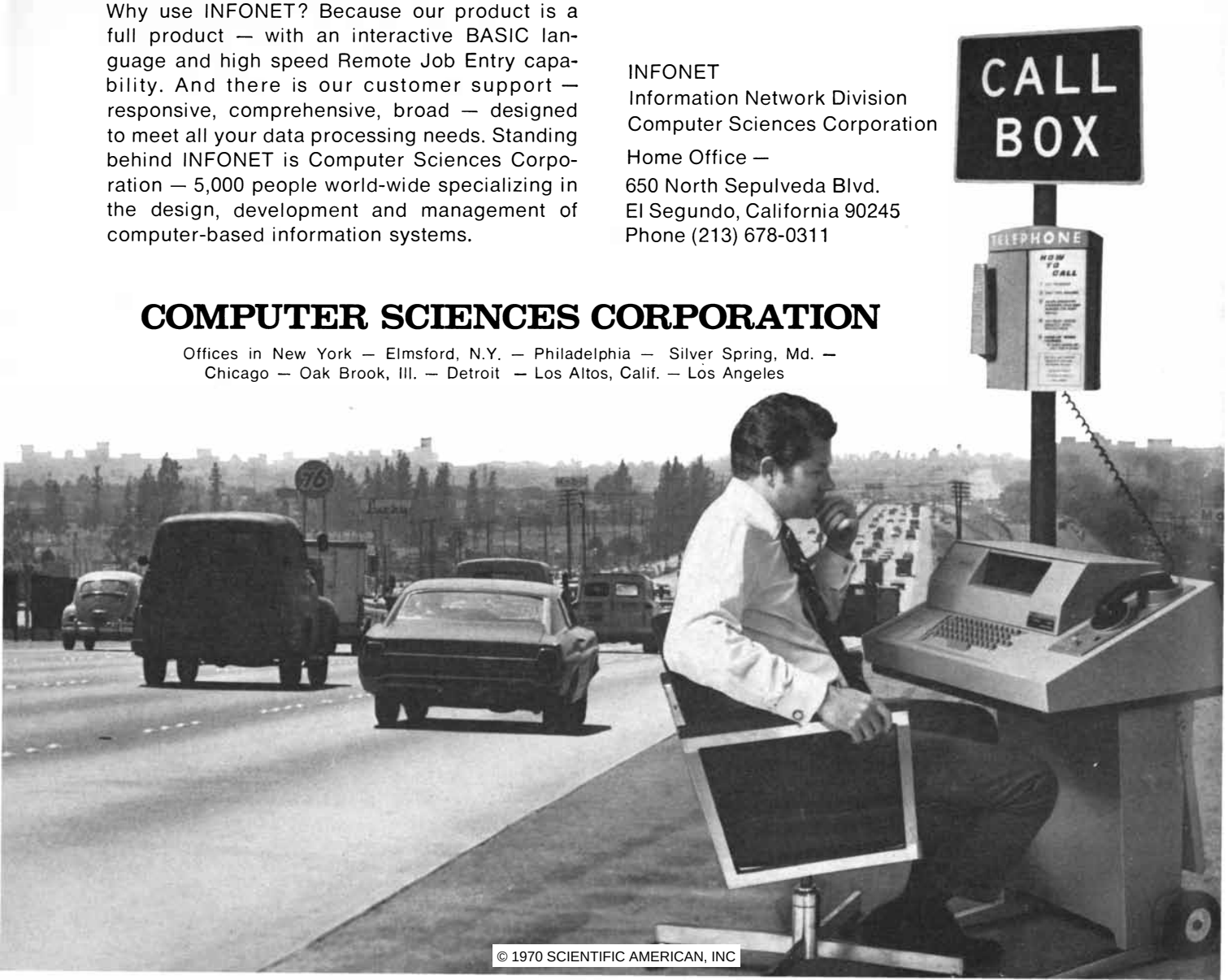
Give us a ring.

INFONET
Information Network Division
Computer Sciences Corporation

Home Office —
650 North Sepulveda Blvd.
El Segundo, California 90245
Phone (213) 678-0311

COMPUTER SCIENCES CORPORATION

Offices in New York — Elmsford, N.Y. — Philadelphia — Silver Spring, Md. — Chicago — Oak Brook, Ill. — Detroit — Los Altos, Calif. — Los Angeles



Established 1845 **SCIENTIFIC AMERICAN** July 1970 Volume 223 Number 1

ARTICLES

DEPARTMENTS

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $10 PER YEAR.

THE COVER

The color photograph on the cover depicts astronomical object No. 13 in the catalogue of the brightest stellar clusters and nebulas compiled in the 1780's by the French astronomer Charles Messier. The photograph was taken through a blue filter at the U.S. Naval Observatory at Flagstaff, Ariz. M 13 is located in the constellation Hercules and can be seen overhead with the unaided eye on summer evenings. The brightest globular cluster visible in the Northern Hemisphere, it is a compact collection of perhaps 500,000 stars. Such stars are discussed by Icko Iben, Jr., in "Globular-Cluster Stars" (*page 26*). In the center of such a cluster the density of stars is probably 50,000 times as great as the density of stars in the vicinity of the sun. What that would look like on a clear summer night is fascinating to imagine.

THE ILLUSTRATIONS

Cover photograph courtesy of U.S. Naval Observatory

Page	Source
17	William Vandivert
18–19	Charles B. Slackman
20–25	Jerome Kuhl
26	Hale Observatories
28–29	Allen Beechel
30	Hale Observatories
31–39	Allen Beechel
41–42	Albert Einstein College of Medicine
43–45	Lorelle A. Raboni
46	Murray B. Bornstein, Albert Einstein College of Medicine
58–67	Bunji Tagawa
73	Dave Fultz, University of Chicago
74–80	Graphic Presentation Services, Inc.
84	H. C. Bennet-Clark, University of Edinburgh
86	Tom Prentiss
87	A. W. Ewing, University of Edinburgh
88–89	H. C. Bennet-Clark, University of Edinburgh (*top*); Tom Prentiss (*bottom*)
90	Tom Prentiss (*top*); H. C. Bennet-Clark, University of Edinburgh (*middle*); A. W. Ewing, University of Edinburgh (*bottom*)
92	H. C. Bennet-Clark, University of Edinburgh
95	Jim Thomas
96–102	Jim Egleson
107	Carl E. Taylor, Johns Hopkins University
108–112	Enid Kotschnig
114	Carl E. Taylor, Johns Hopkins University
118–119	Alan D. Iselin
120–128	Roger Hayward

INTERNATIONAL PAPER ANNOUNCES A $101 MILLION, FOUR-YEAR PLAN TO COMBAT POLLUTION.

International Paper Company believes that the aspirations of our society for a better life can be met, that the pollution of our environment can be controlled, and that the vital quality of the basic resources we all share can be maintained within the framework of our economy. International Paper is dedicated to do its part as an industrial citizen to achieve these goals.

I can now report to you that the Company has adopted a four-year plan, to be completed by 1974. This plan places International Paper in the forefront of those taking positive, constructive measures to solve the problem of environmental quality.

We estimate the total cost of this program will be $101 million.

When this program is completed every one of our U.S. pulp and paper mills will be equipped with primary and secondary waste water treatment systems.

Water so treated does not adversely affect the complicated life chain in natural waters—from bacteria to plankton, to plants and fish life.

In terms of air quality this program will utilize the latest technology, which will permit us to remove over 99% of particulate matter from mill emission points. Presently installed equipment has an efficiency factor of about 90%. It will also include adaptation of new technical developments that will control the odors of a kraft paper mill.

International Paper pledges to apply its technology, its resources, and the efforts of its people to this end.

EDWARD B. HINMAN
President, International Paper Company

LETTERS

Sirs:

I have read with the greatest interest Sidney R. Garfield's account of the Kaiser-Permanente medical-care program and his argument that regular health testing is an ideal regulator of entry into medical care ["The Delivery of Medical Care," by Sidney R. Garfield; SCIENTIFIC AMERICAN, April]. Undoubtedly it is very much more efficient than the traditional haphazard "on demand" utilization of medical care. I do not think, however, that Dr. Garfield has faced up to what I believe to be a more fundamental issue, namely that whatever system is adopted for the delivery of medical care, the potential demand appears to be unlimited. The Kaiser-Permanente experiment itself has confirmed the unsurprising truth that the more systematically one looks for abnormalities, the more one finds. As the population ages, more disease will exist waiting to be detected by regular screening. At the same time our technical ability to treat disorders is rapidly extending, thus increasing the scope for medical care. Finally, our threshold of "acceptable" disorders is being continually lowered, so that people now expect care for conditions they would have been prepared to suffer in silence a few decades ago. Because of this, I believe we should recognize the possibility that regular health testing will in practice bring in more and more patients for medical care as time progresses. We in Britain mistakenly believed the creation of our National Health Service in 1948 would catch up with a backlog of untreated sickness and the demand for medical care would subsequently diminish. As we now know for the reasons described above, the reverse happened, and we believe the situation would be further aggravated if health checks were freely available on the health service. I think it would be foolish for the U.S. to repeat our own error by deluding itself that a system of prepaid medical care, relying on health testing as a regulator to entry, would necessarily contain the demand for treatment to manageable proportions. Some other more effective regulator of demand will, I think, be needed in addition.

G. TEELING-SMITH

Director
Office of Health Economics
London

Sirs:

Mr. G. Teeling-Smith's letter is an important and thoughtful document that warrants reply. He errs when he states "Garfield has [not] faced up to . . . a more fundamental issue, namely that whatever system is adopted for the delivery of medical care, the potential demand appears to be unlimited." It was exactly that issue, namely our inability to satisfactorily meet accelerating demand and our efforts to find a solution, that led to the new concept my article was all about.

Since it appears this was not clear, it is appropriate that I reiterate. I did not propose health testing alone as a solution. Adding health testing to today's medical-care delivery system cannot materially increase its capacity to meet demand. What I proposed was a broader concept based on a simple fact we have uncovered that has somehow passed unnoticed by us and everyone else. This fact is that the elimination of personally paid fees, as occurs in our health plan and Britain's, significantly changes the input into the delivery system. It not only increases the entry of the sick but also inundates the system with a large component of well and worried-well people. This altered entry mix no longer matches today's medical-care delivery system, which has evolved over the years as a sick-care delivery system, with the doctor at the point of entry. The result is a drastic waste of doctors' time and impaired services to the sick and well. What I did propose was a new medical-care delivery system that realistically matches that new input. This is to consist of (1) health testing, to separate the entry mix into its three basic components, the well, the asymptomatic sick and the sick, and (2) three respective services to receive these components: a new health-care service, a new preventive-maintenance service and an unadulterated sick-care service.

It is obvious that health testing would be of little benefit if all these components of the entry mix were guided into the existing sick-care service, as they are today. It is the addition of the new health care and the new preventive maintenance that gives this new system its real value and thrust. These new divisions provide care for the well, the worried-well, the asymptomatic sick and much of high-incidence chronic illness follow-up with paramedical, automated and relatively inexpensive services. Relieved of all that load, the capacity of sick care becomes amplified manyfold to meet the demands of today and the future. This is the great promise of the new delivery system: augmenting the capacity of the physician by relieving him of work that is not his primary concern, much of which he was not trained for and that can be performed simply by other relatively easily trained personnel. It is health testing and its filtering function that makes all of this possible. I too am concerned about the U.S.'s repeating Britain's error. I am also concerned about the U.S.'s repeating its own errors of Medicare and Medicaid. That is why we need health testing, health care and preventive maintenance before embarking on free medical care as a right. These are essential, in Mr. Teeling-Smith's words, to make the resulting demand manageable.

There is nothing mystical about this concept. It is elementary that proper work-system design must match and adjust to changes in input, to create manpower effectiveness and best possible costs. The existing medical-care delivery systems of this country and Britain have unwittingly never adjusted to this basic systems rule. Correcting this maladjustment now becomes our opportunity.

I believe a rational system can now

Scientific American, July, 1970; Vol. 223, No. 1. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and treasurer.

Editorial correspondence should be addressed to The Editors, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Manuscripts are submitted at the author's risk and will not be returned unless accompanied by postage.

Advertising correspondence should be addressed to Allan Wittman, Advertising Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 20 cents.

Change of address (or other subscription correspondence) should be addressed to Circulation Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Please notify us four weeks in advance of change. If possible, kindly furnish an address imprint from a recent issue. Send both old and new addresses or complete and mail Post Office Form 357B.

Please enter change of address ☐
Please enter subscription ☐
☐ 1 year, $10 ☐ 2 years, $18 ☐ 3 years, $25

NAME

NEW ADDRESS

OLD ADDRESS

be developed, through this systems approach, that will take care of all Mr. Teeling-Smith's concerns and provide excellent service. The impressive medical service provided today in the U.S. (and doubtless also in Britain) with a grossly mismatched and defective systems design strongly argues for a superb service when and if rational systematization occurs. Once we accomplish this revision we need not fear unlimited demand; we can welcome it and we certainly will not need silent suffering.

Mr. Teeling-Smith may be interested to know that the Federal Government is helping us to develop and evaluate this new medical-care delivery system. It could be an important contribution to world medicine if Britain would initiate a similar experiment, at some strategic location, so that we could trade ideas and experience and test each other's developments. I would welcome that and would cooperate in any way possible.

SIDNEY R. GARFIELD

Kaiser Foundation Hospitals
Oakland, Calif.

Sirs:

Matthew S. Meselson's article "Chemical and Biological Weapons" [SCIENTIFIC AMERICAN, May] was a timely and almost prophetic piece, at least for me. His statement "The massive use of CS in war... will generate serious opposition to its domestic use" is ironic, as CS was used quite liberally on the Madison campus of the University of Wisconsin during recent antiwar demonstrations. I emphasize "quite liberally" because a canister of CS was tossed into my apartment and has thoroughly contaminated it. I can offer myself as living proof that "exposure to either form of CS causes intense pain in the eyes and upper respiratory tract, progressing to the deep recesses of the lungs and giving rise to feelings of suffocation and acute anxiety," particularly when my roommates and I were made the targets of this "humane" gas. (I must admit that it is better to be gassed than to be shot at.) Meselson suggests that no long-term aftereffects of moderate exposure to CS have been demonstrated, although these possibilities are now being studied. If he or any other biologist is interested and is in search of subjects, he need look no further than our Madison campus.

SANDRA BASS

Madison, Wis.

CYBER eliminates the great computer


MARC-I

Until now you've had a choice.

Either you got a computer large enough to handle your peak loads (and absorbed the costs for idle time), or you got one for average through-put (and sweated out the peak periods).

Some choice. It gave you the option of upsetting your financial manager or the people with the hot data processing problems.

Today CDC has eliminated the need for the great computer compromise. You can get the best of both size systems—plus some other advantages you may not have considered—with the CYBERNET System.

Control Data's CYBERNET System is a network of computer centers located in 33 cities across the United States and worldwide. With CDC 6600 and 3300 computer systems, these centers provide a perfect combination of computing power and high capacity data processing capability. Problems that would take hours or days on medium-sized computers are solved in seconds or minutes with the 6600 supercomputer—a factor that means substantial savings in both time and costs.


MARC-III

And, there's no minimum charge for computer time with the CYBERNET System—you pay just for the resource you use for the time which you actually use it. All the computer capacity you need for peak loads is in this network, and it's only as far away from you as your telephone outlet.

CDC's MARC terminals can be used to augment your present computer system, for testing and debugging programs, or as a primary data processing source. There's a complete line of systems and peripherals to choose from—with input/output capacities ranging from 10 to 6000 bytes-per-second.

Pick out the terminal that fits your needs, select the peripherals you want, and install the system *without* the usual problems of environmental control, false flooring, etc. Most systems are ready to run the day they're installed. Your terminal will be tied into the CYBERNET center nearest you—either by voice-grade or wideband telephone lines. Software includes the most complete versions of FORTRAN and COBOL, as well as a comprehensive library of special programs


MARC-II

NET™ system compromise.

—ALGOL, SIMSCRIPT, OPHELIE II, and systems for assigning zip codes to street addresses, for structural analysis, for analysis of random data, for tabulating and analyzing survey data, to name a few.

And now we have two new software systems that help you use the full potential of the CYBERNET System—SHADOW and SHADE. SHADOW is a system that allows you to select any 3300 or 6600 in the CYBERNET System, or a combination of both, to solve your problem with optimum efficiency and economy. In addition, the powerful command and control features of SHADOW allow you to direct the scheduling and routing of your work flow. SHADE is a system that provides you, from your own site, the capability to store data files at the CYBERNET center so all you have to transmit for any problem is new input or changes. You can alter, update, search or copy files from your own terminal.

MARC-IV

MARC-V

Take a few minutes to find out exactly what the CYBERNET System and MARC terminals can offer your firm. Call your nearest CDC Data Center, or contact us directly.

INTERNATIONAL DATA CENTERS

STOCKHOLM	ROME
FRANKFURT	MILAN
STUTTGART	MELBOURNE
THE HAGUE	SYDNEY

CONTROL DATA CORPORATION

Call us "collect" from anywhere in the U.S. on our Hotline (612) 884-8195.

For further information, call or write:

In the United States:
Data Services Division
Dept. 411, Room 346
Control Data Corporation
4550 West 77th Street
Minneapolis, Minnesota 55435
Phone: 612/884-8195

In Europe:
CDC Data Services, Dept. 411
Control Data GmbH
6 Frankfurt am Main 16
Niddastrasse 40
West Germany
Phone: 71-231

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JULY, 1920: "At last we Americans are thoroughly aroused over the seriousness of the threat to our lumber resources. For decades we have been cutting down our forests with little thought of reforestation; for the past 10 years a mighty host of periodicals and newspapers have been clamoring for thousands upon thousands of tons of paper each week, which has meant the cutting down of thousands of acres of splendid timber. With the advent of the great war more forests were cut down to build the expansive cantonments, wooden ships and other things. And now we are confronted with seriously depleted forests. Reformation and conservation are the subjects of the hour, for wood is something without which we can hardly get along."

"The recent 1,400-mile non-stop flight of an all-metal passenger monoplane is notable both because of the distance covered and because of the special design and materials of the machine. The all-metal monoplane has particular interest for this journal, for the reason that as far back as 1910 SCIENTIFIC AMERICAN published its own design for a racing monoplane built entirely of high-grade alloy steel. The suggestion was ahead of its day, but with the coming of the wonderful aluminum alloys that are now available the design took practical shape, and toward the close of the war the Germans built some very successful monoplanes, which dispensed entirely with wood and linen fabric, depended for their wing strength on internal trussing and were covered as to wings and body with thin corrugated aluminum plating. We believed then and we are still of the opinion that the fastest machines of the future will be of the monoplane type."

"That 1920 will be the greatest producing year thus far in the history of the automotive industry can be readily foreseen, with the great demand for motor vehicles in this and other countries. Production in America has reached such a point that the National Automobile Chamber of Commerce announces that this year's business will rank second only to steel. The 1919 production totaled $1,807,594,580, and if bodies and accessories were included, this figure would exceed $2,000,000,000."

"The 17th nova found in the Andromeda nebula is reported by Milton Humason in the *Publications of the Astronomical Society of the Pacific*, No. 185. It appears on plates taken at Mount Wilson last October, but is invisible in earlier and later plates."

"It gives us the greatest pleasure to state that Mr. Eugene Higgins, an American resident of Paris and for many years a close friend of this paper, offers through SCIENTIFIC AMERICAN a prize of $5,000 for the best essay on the Einstein postulates and their consequences, written so that a person of no special mathematical training may read it profitably. The sum of $5,000 exceeds, if we are not mistaken, any award open to a professional man with the single exception of the Nobel prize, for which he cannot specifically compete."

SCIENTIFIC AMERICAN

JULY, 1870: "When crimes or accidents are recorded in newspapers more than formerly, some people fancy that they happen more than formerly. A foreign traveler in England is said to have remarked on the perceptible diminution of crimes committed during the sitting of Parliament as proof of high reverence for that assembly; the fact being that the space occupied in the newspapers by the debates causes the records of many crimes to be omitted. Men are liable to form an over-estimate of the purity of morals in the country as compared with a town. On a given area, it must always be expected that the absolute amount of vice will be greater in a town than in the country; so also will be that of virtue."

"The grounds upon which Christian scientific men can stand secure were admirably stated by Professor Dana in his recent lecture before the seniors, in which the subject of Darwin's theory was considered. In the course of his remarks he stated that belief in a development theory was not atheism, that the facts of science clearly indicate some plan of development, that Darwin's book was a work of great merit and that his theory accounts for the origin of some species. As for genera and higher groups, there will probably be found other laws to account for them. Let no one fear scientific investigation, for its results are only another name for God's truth. Such belief, enunciated by men of science whose position as men of Christian faith is unquestioned, should calm the fears of those who tremble before every new discovery, and show no faith in the strength and majesty of truth."

"The practical application of the condensation of gases for the production of cold is a result that has been attained this year more than in any other period. The fact of the possible compression into liquids was long ago ascertained by Faraday, and feeble attempts were made a few years since to apply it for the production of cold, but it was not until recently that these experiments proved successful. There now appears to be no doubt that the liquefaction of gases is the true method upon which to found the artificial production of ice on a commercial scale."

"The waste of coal-tar products is fast disappearing, and so great has been the progress in the discovery of the new application of the liquid and solid products of the distillation of coal that we may expect to see retorts erected for the purpose of producing them, rather than for the manufacture of gas. Gas will become an incidental product, while the object sought will be the tar from which to make aniline colors, and anthracene from which to manufacture alizarine and artificial madder dyes."

"At a celebrated observatory in the suburbs of London a visitor was desirous of observing a celestial object which was nearly overhead, and having the run of the observatory at the moment, he directed the telescope towards the star, set the clockwork in motion and placed himself on his back in the observing frame attached to the floor. Intent on observing the star, our astronomer failed to notice that the movement of the eyepiece was gradually imprisoning him. His head was fixed by the headrest, and the eye-tube was beginning to press with more and more force against his eye. Fortunately his cries for assistance were heard, and the clockwork was stopped. Otherwise he would have had as good reason to complain as the celebrated astronomer Struve in the case of the Pulkova refractor, which, Struve said, was justly called a refractor since it had twice broken his leg."

In three weeks he may be dead.

In the beautiful woods and valleys of eastern Oklahoma time is running out. Up to fifty percent of the newborn fawns are being lost each year because of ticks.

When large numbers of these crab-like pests attack a healthy young deer, he cannot live for more than a few weeks.

And it isn't just deer that are affected. Ticks will attack virtually any land animal or bird they can get hold of.

But there is a way to control these marauders—kill them on the ground where they breed.

To do this, the Oklahoma Department of Wildlife Conservation together with Oklahoma State University are using Shell's Gardona®, an insecticide. A mere pound to an acre can kill the resident tick population for up to nine weeks.

Yet it will not harm animals, birds, plants or people.

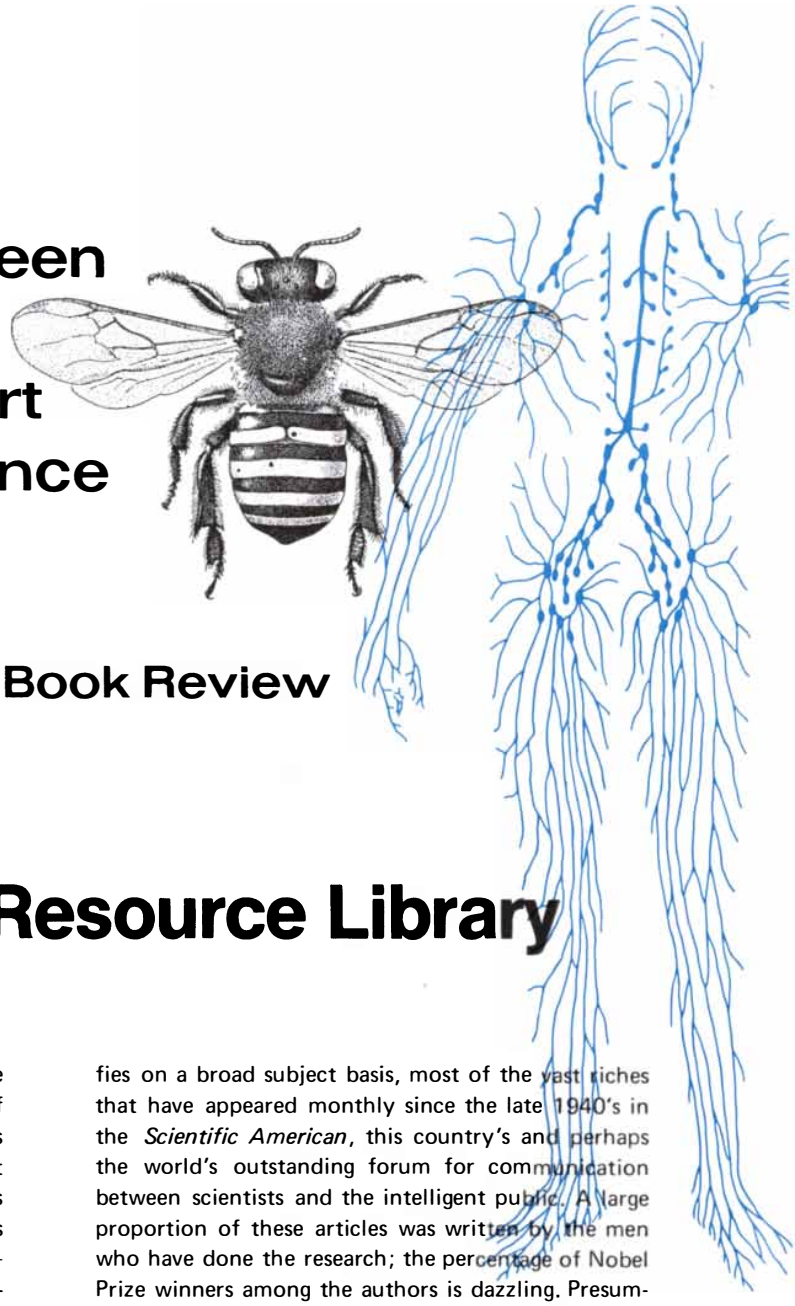
Shell has also funded a grant to Oklahoma State University's Department of Entomology for more intensive study on the control of ticks.

Shell's concern with wildlife is only part of an all-out program to help save our environment. So far we've backed our commitment with millions of dollars a year in the war against pollution.

And we're moving as fast as we can.

Because, like the fawn, we're all running short on time.

SHELL



"Quite simply there has never been anything in the history of the effort to popularize science comparable to these volumes."

—The New York Times Book Review

SCIENTIFIC AMERICAN Resource Library

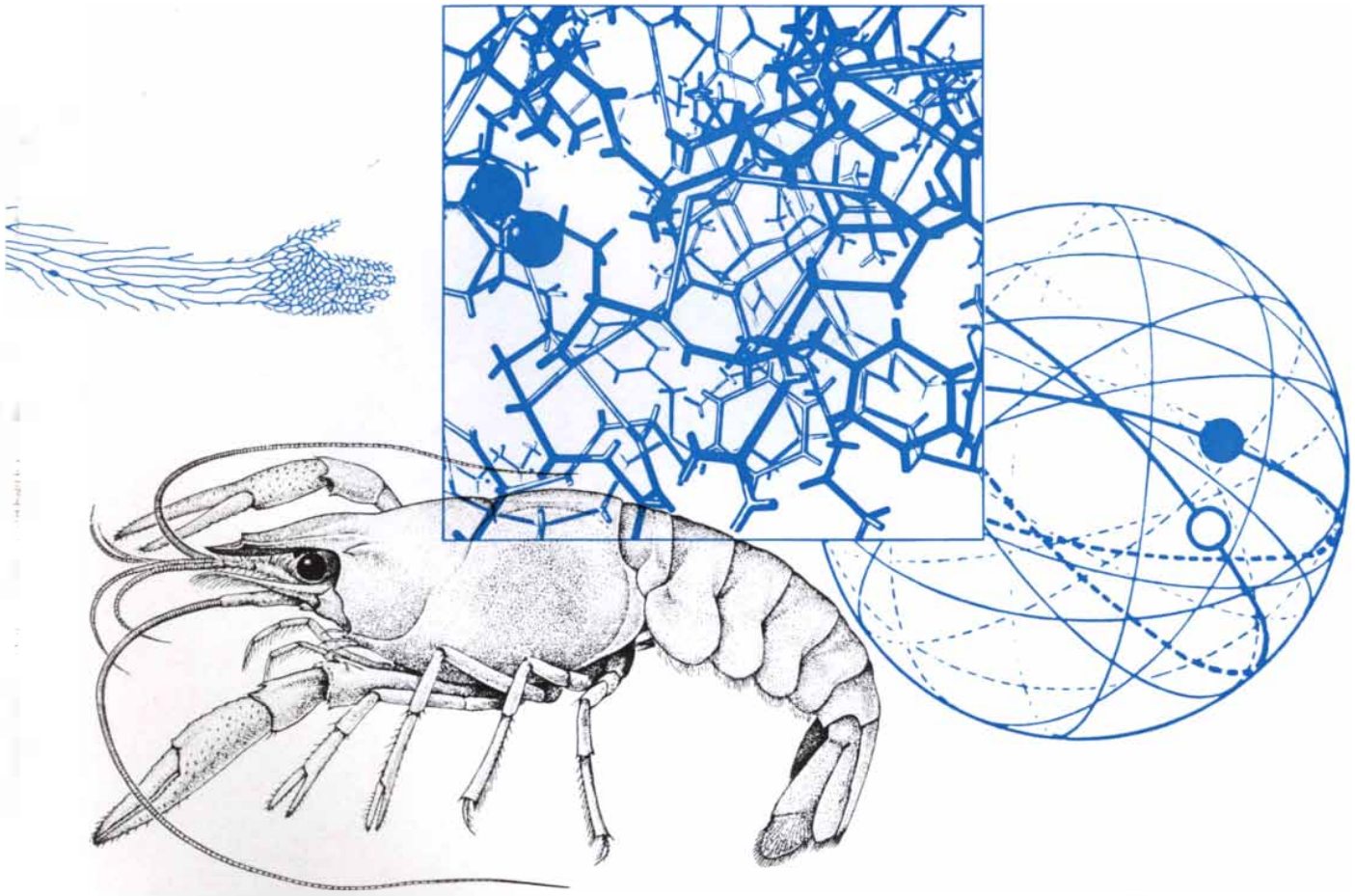
"A better title for these books would have been The Intelligent Layman's Guide to the Golden Age of Science. Their basic contents are hundreds of articles which explain simply, vividly and interestingly what modern science is all about, what its accomplishments have been since World War II, and what new problems have arisen—as in the case of the environment—because of the growth of population and the application and misapplication of modern technology.

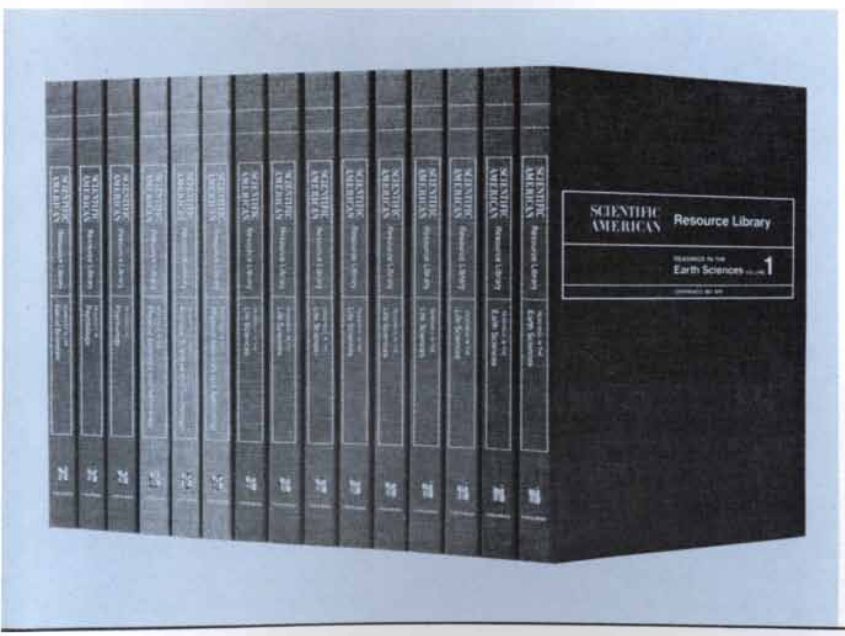
"The collection provides the basic means for any moderately intelligent person to inform himself relatively painlessly and often with great pleasure as to what scientists have been doing in studying the heavens, the earth, the oceans, living organisms, and the mechanism of life, as well as man, his society, his psychology and his economy.

"The popularization of science is, of course, an old and now widespread field. What makes this collection unique, however, is that it brings together and classifies on a broad subject basis, most of the vast riches that have appeared monthly since the late 1940's in the *Scientific American*, this country's and perhaps the world's outstanding forum for communication between scientists and the intelligent public. A large proportion of these articles was written by the men who have done the research; the percentage of Nobel Prize winners among the authors is dazzling. Presumably with the help of the magazine's editors, the scientists have managed to communicate with great clarity their findings and the methods they employed, using both prose and detailed diagrams. . . .

"Quite simply there has never been anything in the history of the effort to popularize science comparable to these volumes, except, of course, for the aging and well-thumbed file copies of the *Scientific American* itself."

—*The New York Times Book Review,* April 26, 1970





READINGS IN THE EARTH SCIENCES
With an Introduction by William H. Matthews III
Lamar State College of Technology
Two volumes, Offprints 801-874, $20.00

READINGS IN THE LIFE SCIENCES
With an Introduction by Marston Bates
The University of Michigan
Seven volumes, Offprints 1-199; 1001-1120, $70.00

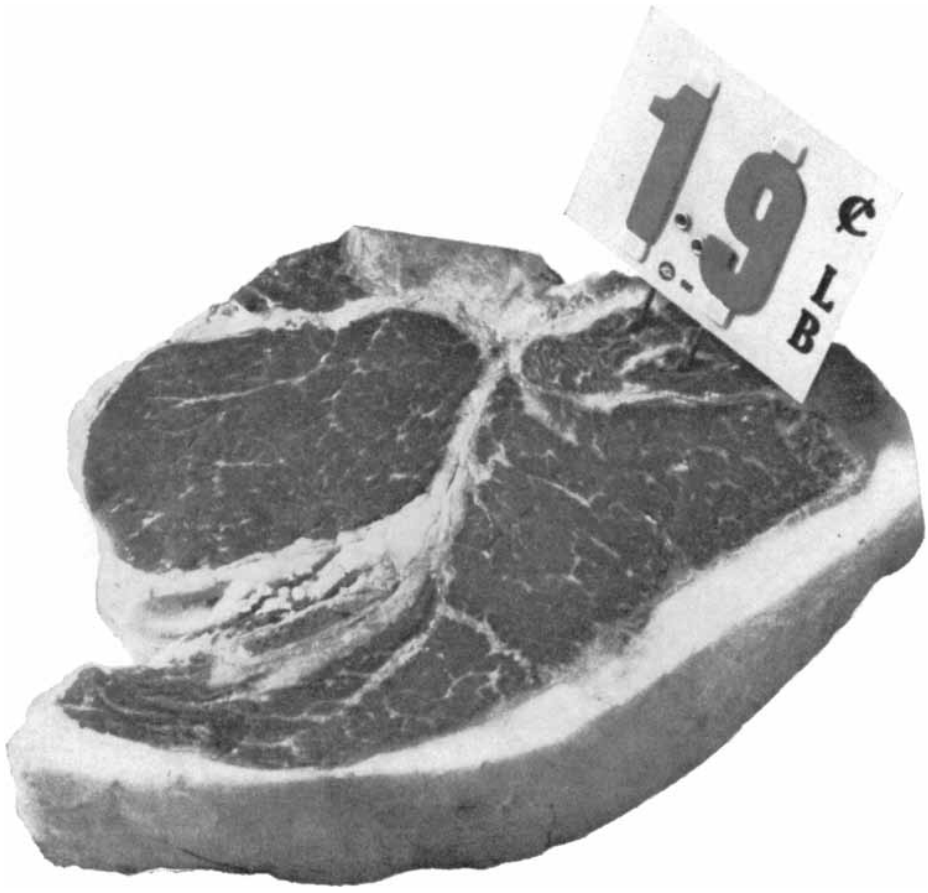
READINGS IN THE PHYSICAL SCIENCES AND TECHNOLOGY With an Introduction by
Isaac Asimov, Boston University School of Medicine
Three volumes, Offprints 201-326, $30.00

READINGS IN PSYCHOLOGY
With an Introduction by Harry F. Harlow
The University of Wisconsin
Two volumes, Offprints 401-515, $20.00

READINGS IN THE SOCIAL SCIENCES
With an Introduction by Harry F. Harlow
The University of Wisconsin
One volume, Offprints 601-637, $10.00

COMPLETE LIBRARY
Fifteen volumes in all, $150.00 the set

W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, California 94104
Warner House, Folkestone, Kent, England



You'd rather pay six times as much.

Admit it.

If we offered you steak for 19 cents a pound you'd be suspicious.

It was the same story when we offered foam manufacturers a urethane catalyst for as little as one-sixth what they'd been paying. They were suspicious.

So instead of having people beating a path to our door, we had to beat one to theirs. It took a lot of selling, but we finally convinced people to test our urethane catalysts for themselves. So that today we have a growing family of urethane and other polymer catalysts.

There's Polycat® 8 which is helping formulators of rigid urethane foams cut catalyst costs by one-third.

Polycat® 12 which makes a flexible foam that won't stain vinyl.

Polycat® 200 which gives can coatings and other cross-linking thermoset resins a much harder finish.

And Polycat® 352 which cures an epoxy with unequaled compressive and flexural strengths.

As the chemical marketing arm of Abbott Laboratories, one of our jobs is finding industrial applications for many of the 1000 new chemicals Abbott synthesizes each year. Polycat is only an example of how Abbott Chemicals is supplying chemicals with applications that make existing products better.

We've also developed one of the first multi-functional acid inhibitors for the chemical treatment of metals.

A better antioxidant/stabilizer for middle distillate fuels.

And now we're putting together a number of multi-functional additive packages containing a complete mix of ingredients that the petroleum industry is looking for.

To do all this, we rely heavily on our people. To find new uses for our existing chemicals. To uncover new applications for never-before-used chemicals. To help find solutions to your chemical problems.

Helping solve your problems is what our product development is all about.

North Chicago, Ill. 60064
Montréal, P. Q., Canada
Zurich, Switzerland

Abbott Chemical Products Division

THE AUTHORS

LORD RITCHIE-CALDER ("Conversion to the Metric System"), Baron of Balmashannar, is chairman of the Metrication Board in Great Britain. Educated in Scotland, he began work as a police reporter for the *Dundee Courier* in 1922 and then worked as a newspaperman in London for many years. During World War II he was with the Foreign Office; after the war he served as science editor of the *London News Chronicle* and member of the editorial board of the *New Statesman.* He was also associated with a number of agencies of the United Nations. From 1961 to 1967 he was professor of international relations at the University of Edinburgh. Author of some 30 books, he is a founder-member and former chairman of the Association of British Science Writers. He was made a life peer in 1966.

ICKO IBEN, JR. ("Globular-Cluster Stars"), is professor of physics at the Massachusetts Institute of Technology. He writes: "Name passed on for at least 10 generations; is perhaps of Hun origin. As an undergraduate at Harvard majored first in political science and linguistics, then in biochemistry and finally in physics. Switch to biochemistry came following loss of faith in the perfectibility of man (Korean war). Switch to physics due to lack of interest in cooking. Spent five years getting Ph.D. in physics at the University of Illinois and wondering why I was in the field. Rekindling of interest in astronomy occurred literally by accident. First day of one summer vacation cracked two vertebrae in collision between myself (on bicycle) and a car. Spent month's vacation in cast while family frolicked on beach. Picked up Fred Hoyle's *Frontiers in Astronomy* and was converted to astrophysics."

GEOFFREY DEAN ("The Multiple Sclerosis Problem") is director of the Medico-Social Research Board of Ireland. Educated in England at Ampleforth College and the University of Liverpool, he moved to South Africa in 1947 and practiced as a consultant physician in Port Elizabeth. He studied porphyria in a number of countries and described a new form, porphyria variegata, which has been inherited by more than 9,000 living white South Africans from one ancestor who married at the Cape of Good Hope in 1688. In South Africa and elsewhere he investigated lung cancer and bronchitis. Since 1947 he has studied multiple sclerosis among South Africans and immigrants of that country. Dean was elected a fellow of the Royal College of Physicians in 1964.

ERIC R. KANDEL ("Nerve Cells and Behavior") is professor in the departments of physiology and psychiatry at the New York University School of Medicine and chief of the newly formed department of neurobiology and behavior at the Public Health Research Institute of the City of New York. Born in Vienna, he came to the U.S. in 1939 and has received most of his education here: he was graduated from Harvard College in 1952 and obtained his M.D. at the New York University School of Medicine in 1956. He was a resident in psychiatry at the Massachusetts Mental Health Center in Boston from 1960 to 1962 and again in 1963–1964. "While in medical school," he writes, "I became interested in research on the nervous system. I was at the National Institute of Mental Health from 1957 to 1960 working on a cellular study of neurons in the hippocampus, a part of the mammalian brain thought to be concerned with memory. On completing this work I began to look for a simple preparation in which it might be possible to study the neural mechanisms of behavior and learning more directly. This search led me to the marine snail *Aplysia.*"

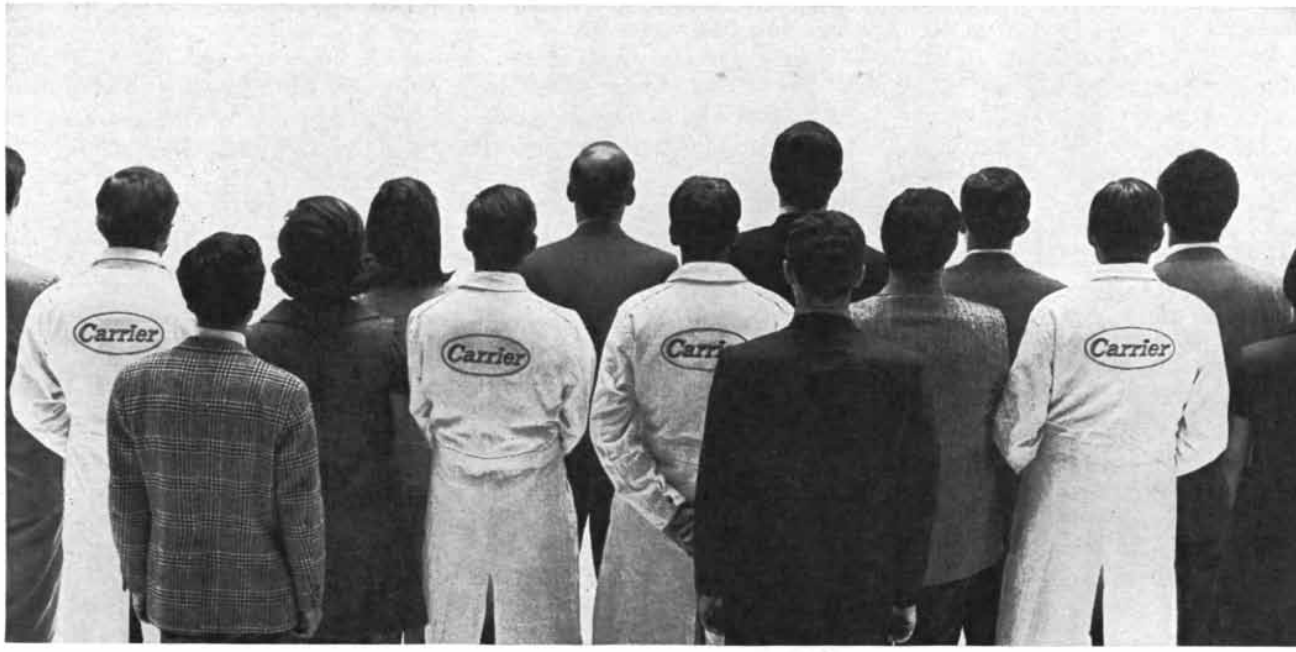
VICTOR P. STARR and NORMAN E. GAUT ("Negative Viscosity") are respectively professor of meteorology and research associate at the Massachusetts Institute of Technology. Starr, who received his Ph.D. from the University of Chicago in 1946, is mainly interested in the dynamics of planetary atmospheres. He is the author of a book, *Physics of Negative Viscosity Phenomena,* that was published recently, and he has also written extensively on meteorology, oceanography and astronomy. Gaut, whose main interests are meteorology, planetary science and the origins of the solar system, obtained his Ph.D. from M.I.T. in 1967. In addition to his work there he is vice-president of Environmental Research and Technology, Inc.

H. C. BENNET-CLARK and A. W. EWING ("The Love Song of the Fruit Fly") are lecturers in zoology at the University of Edinburgh. Bennet-Clark's professional interest is in the application of physical and mechanical principles to biological systems; his recent studies include the energetics of the flea's jump and the acoustics of song production in the mole cricket. His avocations include photography, antiques (in particular Georgian silver and porcelain) and the design of high-fidelity equipment. Bennet-Clark was graduated in zoology from University College London and took his doctorate at the University of Cambridge. Ewing, who was graduated from the University of Edinburgh in 1959 and also took his doctorate there, is interested professionally in comparative aspects of animal behavior. His hobbies include mountaineering (he edits the guides of the Scottish mountaineering club), the restoration of vintage motorcycles and the making of pottery.

HOWARD FRANK and IVAN T. FRISCH ("Network Analysis") are respectively executive vice-president and vice-president of the Network Analysis Corporation. Frank was graduated from the University of Miami in 1962 and received his master's and doctor's degrees at Northwestern University. From 1965 to 1970 he was with the department of electrical engineering and computer sciences at the University of California at Berkeley. During 1968–1969 he was on leave to the Office of Emergency Preparedness of the Executive Office of the President, serving as head of the network-analysis group. Frisch, who was born in Hungary and brought to the U.S. as a child, received a bachelor's degree in physics at Queens College and one in electrical engineering from Columbia University under a combined program; his master's degree and doctorate are from Columbia. He was on the faculty at Berkeley from 1962 to 1970. Frank and Frisch are coauthors of a forthcoming book, *Communication, Transmission and Transportation Networks.*

CARL E. TAYLOR ("Population Trends in an Indian Village") is professor of public health administration and chairman of the department of international health at the School of Hygiene and Public Health of Johns Hopkins University. He was born in India, where his father was a medical missionary doing village health work, but received his higher education in the U.S., obtaining his bachelor's degree at Muskingum College in 1937, his M.D. from the Harvard Medical School in 1941 and degrees of master and doctor of public health from the Harvard School of Public Health. He was on the faculty of that school for a number of years before going to Johns Hopkins in 1962. From 1953 to 1956 he was professor of preventive and social medicine at Christian Medical College in India.

This is The Group that turns down regular work.



They're our independent bunch at Carrier. And there are over 600 of them, including a large staff of engineers and technicians, who work on special air conditioning and refrigeration projects.

The Group came up with a 15-ton centrifugal compressor that weighs under 14 lbs. Pairs of them now cool DC-8 Super 61's and 63's.

And there are scores of other out-of-the-ordinary products they've created. Tested. Proved. And manufactured. Either as prototypes or in production quantities.

Every assignment The Group takes on is backed with the knowledge and giant facilities of Carrier Air Conditioning Company—where we produce more air conditioning equipment than anybody. And average over 200 patents. Every year.

For the unusual in air conditioning and refrigeration, try The Group's therapy. In transportation of people and perishables. The Military. Aerospace. Oceanography. Industry and commerce. Just write Special Products Group, Carrier Air Conditioning Company, Syracuse, N.Y. 13201.

Carrier® **We keep on inventing air conditioning**

Established 1845 SCIENTIFIC AMERICAN July 1970 Volume 223 Number 1

Conversion to the Metric System

Great Britain, which developed the "imperial" system of weights and measures, is now well into a period of transition to the metric system. Before long the U.S. will be the only major nonmetric nation

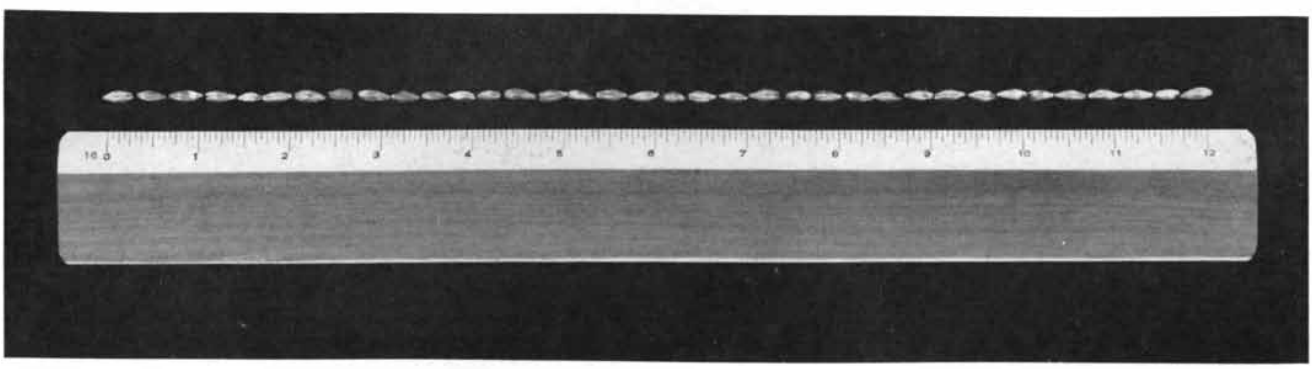
by Lord Ritchie-Calder

In the printed proceedings of the British Parliament for May 24, 1965, the president of the Board of Trade is quoted as follows in reply to a question by a member of the House of Commons: "The Government are impressed with the case which has been put to them by the representatives of industry for the wider use in British industry of the metric system of weights and measures. Countries using that system now take more than one-half of our exports; and the total proportion of world trade conducted in terms of metric units will no doubt continue to increase. Against that background the Government consider it desirable that British industries on a broadening front should adopt metric units, sector by sector, until that system can become in time the primary system of weights and measures for the country as a whole....

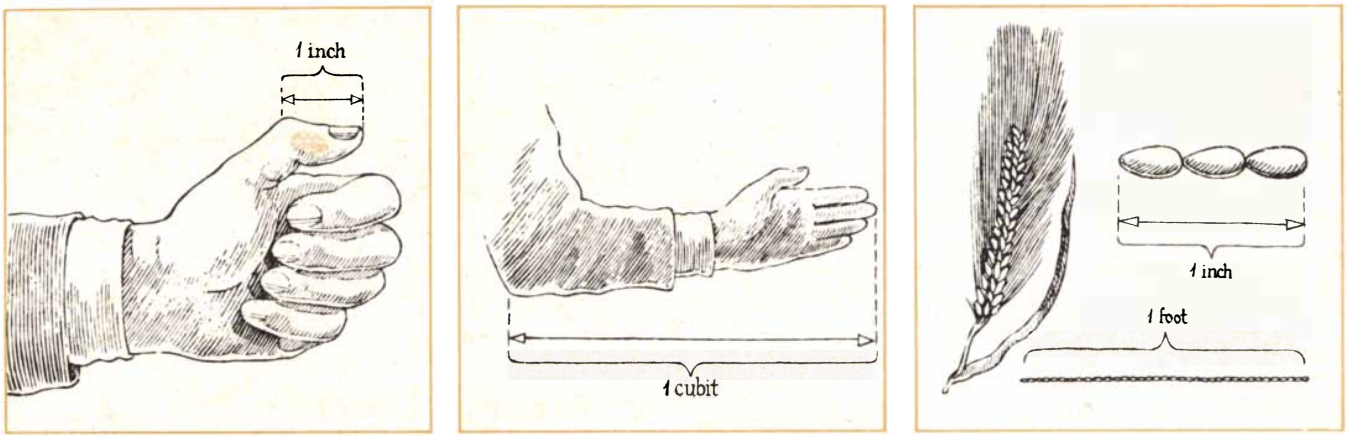
"Practical difficulties attending the change-over will, of course, mean that this process must be gradual; but the Government hope that within ten years the greater part of the country's industry will have effected the change."

Thus the nation that developed the "imperial" system of weights and measures—the system of inches, ounces and quarts familiar throughout the English-speaking world—announced its decision to abandon the system in favor of the decimal scheme first officially launched by France in 1790 and spread widely through Europe by Napoleon. The trend toward "metrication" is such that the U.S. may soon be the only major nation still using the imperial system, and it seems likely that within another decade or two the U.S. will follow the route of the British. Now that Britain is halfway through the period of 10 years established for putting the metric system into effect, the question is how metrication is progressing. The answer, as my colleagues and I on the British Metrication Board see it, is: Quite well.

The folklore of measurement is entrancing. In the Anglo-Saxon measurements that underlie the imperial system the inch was the knuckle of the thumb. The cubit was the distance from the elbow to the end of the middle finger; stated another way it was six palms,



ANCIENT STANDARD FOR THE FOOT was 36 barleycorns, "taken from the middle of the ear." The standard was adopted 1,000 years ago when the earlier definition of four palms or 16 fingers proved to be too variable. Under the continuing trend toward precision the meter, the basic unit of the metric system, is now defined in terms of wavelengths of radiation of the krypton-86 atom.



FOLKLORE MEASUREMENTS that underlie the imperial system derive from Anglo-Saxon times or earlier. They include the inch, which was the knuckle of the thumb; the cubit, which was the distance from the elbow to the end of the middle finger; the 36-

the palm being the width of four fingers. The foot was not what one might think: it was originally four palms or 16 fingers, but since there is much variation in the size of human fingers and human feet, the foot was standardized 1,000 years ago as 36 barleycorns "taken from the middle of the ear." (Barley also measured weight and volume; a bushel was 50 pounds or eight imperial gallons.) A yard was the distance from the tip of King Edgar's nose to the tip of the middle finger of his outstretched arm. A fathom was said to be the length of a Viking's embrace. An acre was the amount of land that could be plowed in a day by a yoke of oxen. The amount could vary depending on whether the soil was heavy or light. In those days the larger unit was the hide; it too was a variable measurement, being based on the amount of land a yoke of oxen could plow in a year. This amount, which averaged about 120 acres, was deemed enough to support a family.

In medieval Germany the length of the *rute* (rod, pole or perch) was established by a random sample of 16 men coming out of church: they stood toe-to-heel, and the overall length was accepted as the standard. The Roman legionaries strode out the mile; their *mille* was 1,000 paces or about 1,618 yards. A pace, however, is a variable unit, and so the mile has varied: the British and American mile was standardized at 1,760 yards, but in a number of European countries the mile ranges from less than a British-American mile to six such miles.

There was much to commend the units based on the human body. The elbow-length cubit was just about right for knotting the cords that measured lengths. The knuckle-inch and the palm-width were perfectly serviceable artisan units when every article was a prototype and did not have to fit precisely into something else. It was fitting that made trouble. For example, in the north of Italy in Leonardo da Vinci's time, lengths were in *braccia* (arm lengths), but Leonardo was a tall man and could outreach most of his contemporaries. If he was using his own *braccio*, he must have run into difficulties when he had to fit something he had designed or made into something measured according to other *braccia*. Even the standard *braccia* varied from city to city; a Florentine door would have been too big for a Milanese doorframe.

Throughout the centuries measurements proliferated in regions, countries and localities and persisted long after their original significance and utility had been forgotten. As a result there were problems not only in fitting but also in making sure of honest measure and in avoiding confusion born of misunderstanding. Statutes and therefore standards had to be imposed. William the Conqueror regularized measurements of area. His son Henry I applied penal sanctions to weights and measures and also legalized King Edgar's arm length as "our yard of Winchester." The Magna Carta spelled out weights and measures for wine, ale, corn and cloth. The framework was thus established for the units of the imperial system.

The first person to propose a decimal system of measures was Simon Stevin, an inspector of dikes in the Low Countries at the beginning of the 17th century. He had little success. In 1670 Gabriel Mouton, a French priest, proposed in the French Academy of Sciences that the length of one minute of arc of the earth's meridian be called a *milliare* and be subdivided into 1,000 *virgas*. He had no success. Sixty-five years later, however, the Academy dispatched three men to equatorial South America with the object of establishing the length of a pendulum with a period of one second at the Equator as a "natural" unit. The measurements were made, but nothing came of the scheme. In the U.S., Thomas Jefferson, who was attracted by the idea of a fully rational system of measurement, proposed to Congress in 1790 that a new unit of length be adopted. The proposal did not carry.

In the same year the French National Assembly appointed a committee to explore the possibilities of a rational system of measurement. The committee's proposals for a new unit of length, the meter, and a new unit of weight, the kilogram, were enacted into law in France a few years later. The meter was defined as one ten-millionth of a quadrant of the earth's meridian as carefully measured over the fraction of the quadrant lying between Dunkirk in France and a point near Barcelona. The unit of weight or mass was established as the kilogram, which was defined as the mass of a cubic decimeter (.1 meter) of water at the temperature of its maximum density; this volume was to be called a liter.

Efforts by France to interest Britain in adopting the system and helping to make it international were unsuccessful. Although the system was installed in other European countries under Napoleon, it took a long time to gain acceptance in France itself. From the outset the system was taught in the schools, but it did not come into general use until 1840.

Meanwhile the metric system was again being considered in the U.S. John Quincy Adams favored the system, but as Secretary of State in 1821 he reluc-

barleycorn foot; the yard, which was measured from the tip of the nose to the tip of the middle finger on the outstretched arm of King Edgar; the fathom, "the spread of a Viking's embrace," and the acre, the amount of land plowed by a yoke of oxen in a day.

tantly advised Congress not to adopt it "because it would be hazardous to deviate from the practice of Great Britain." (A century later the excuse of British industrialists dragging their feet on metrication was that they did not want to deviate from the practice of the U.S., their biggest customer.)

In the 1860's both the U.S. and the United Kingdom acknowledged the virtues of the metric system by authorizing it as a second system, although they retained imperial units as the legal weights and measures. The two nations really had no choice but to acknowledge the metric system, because the traditional units could not cope with the rapid growth of science. Mechanics of the steam age might settle for folklorish units such as horsepower and foot-pound, but units of this kind could not accommodate the requirements of the physicists and telegraphers who were trying to quantify the new phenomena involved in the discoveries of Michael Faraday and Joseph Henry and in the electromagnetic theory of James Clerk Maxwell. The electricians of the 19th century were hard put to devise everyday standards; in expressing electrical resistance, for example, they had to resort to such usages as "25 feet of copper wire weighing 345 grains" and "one German mile of iron wire .1 inch in diameter."

The initiative for a better system was taken by Maxwell and William Thomson (later Lord Kelvin). Under their aegis a committee was set up in 1861 by the British Association for the Advancement of Science. The committee recognized the advantages of having units of electromotive force, current and resistance of convenient size and the desirability of having them directly related to the metric system. The volt, the ampere and the ohm emerged from a series of International Electrical Congresses between 1881 and 1908.

Difficulties in reproducing or comparing metric standards from country to country led France to call in 1870 an international convention to work out a unified metric system. These efforts led to the signing of the Treaty of the Meter in Paris in 1875. The treaty established an International Bureau of Weights and Measures, which would be custodian of the standards, and a General Conference on Weights and Measures, which would meet periodically to adopt new definitions as the need arose.

Under this system the meter was officially defined in 1889 as the distance between two engraved lines on a bar of platinum-iridium alloy at zero degrees centigrade (now degrees Celsius). The bar was stored in a vault near Paris and was available for comparison with the standard of length of any nation; copies of the bar were supplied to the nations party to the treaty. Also in 1889 the kilogram was defined as the mass of a particular platinum-iridium cylinder, and copies of the cylinder were supplied to the treaty nations.

Over the years the General Conference has steadily extended and refined the metric system. In 1960 the Conference adopted the International System of Units (often abbreviated SI for Système International). This is the system, basically metric, now being put into effect in Britain. SI rests on four independent base units for length, mass, time and temperature. The Conference also decided to include as base units the ampere for electric current and the candela for luminous intensity. These units are not independent of the other four; they were classified as base units for reasons of convenience.

The unit of length is still the meter, but the meter is no longer based on the quadrant of the meridian or on the meter bar. In 1960 the General Conference redefined the meter as "equal in length to 1,650,763.73 wavelengths in vacuum corresponding to the transition between the levels $2p_{10}$ and $5d_5$ of the krypton-86 atom."

The unit of mass is still the kilogram, and its definition has not been changed. It seems probable, however, that in time the General Conference will redefine the kilogram in terms of atomic mass.

The unit of time, the second, had been variously defined until in 1960 the General Conference approved the standard of "the fraction 1/31,556,925.9747 of the tropical year 1900 January 0 at 12 hours ephemeris time." In 1967 the Conference turned to the atomic clock and defined the second as "the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom."

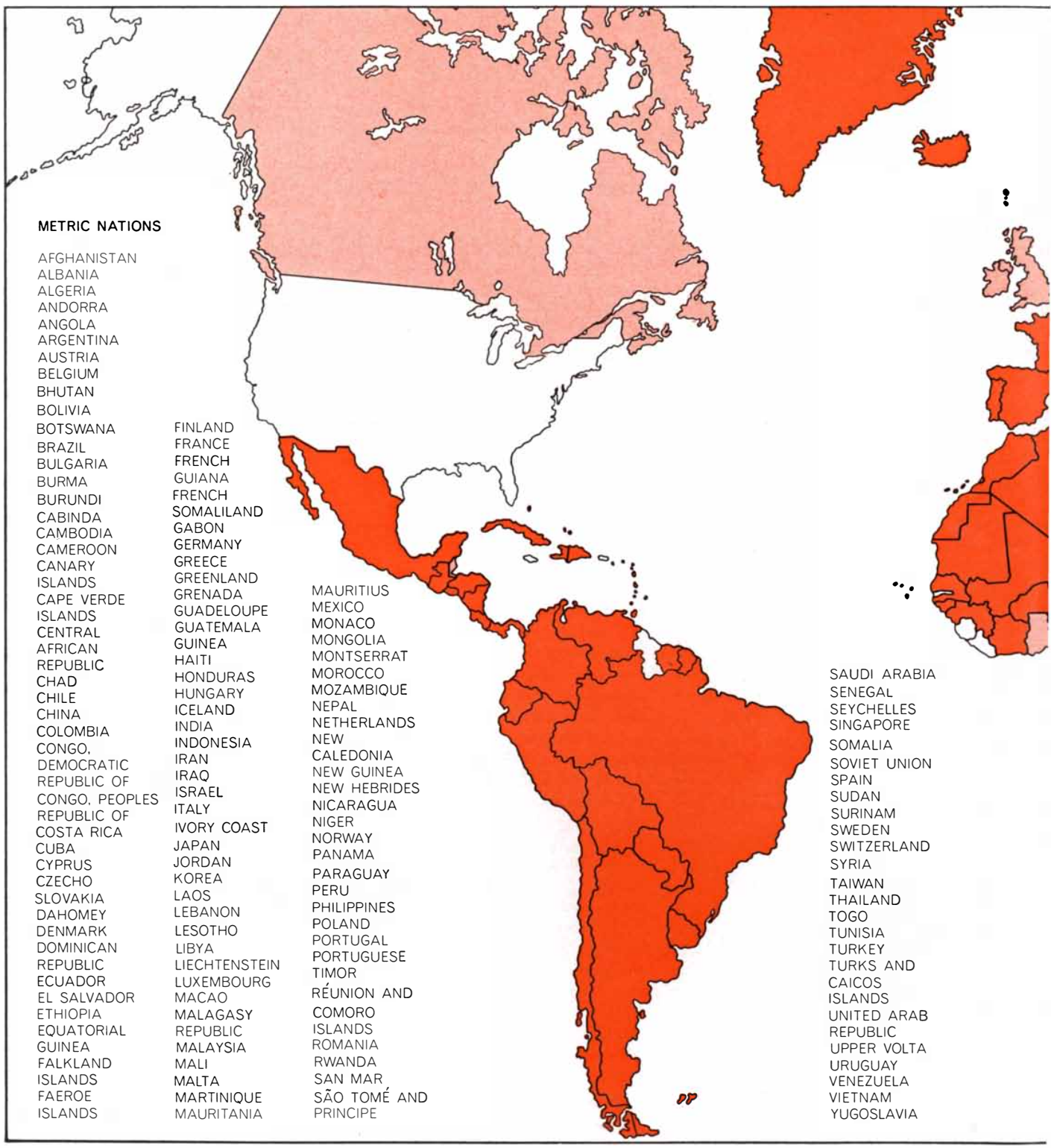
The unit of temperature is the degree Celsius or Kelvin. Originally it was defined as one one-hundredth of the interval between the freezing point of water (0 degrees C.) and the boiling point of water (100 degrees C.). In 1954 the General Conference redefined it as 1/273.16th of the thermodynamic temperature of the triple point of water (.01 degree C.).

The basic electrical unit, the ampere, has been defined since 1948 as "that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross section and placed one meter apart in a

vacuum, would produce between those conductors a force equal to 2×10^{-7} newton per meter of length."

The candela was adopted in 1948. It is defined as the luminous intensity in the perpendicular direction of a surface of 1/600,000th of a square meter of a black body at a temperature of freezing platinum under a pressure of 101,325 newtons per square meter.

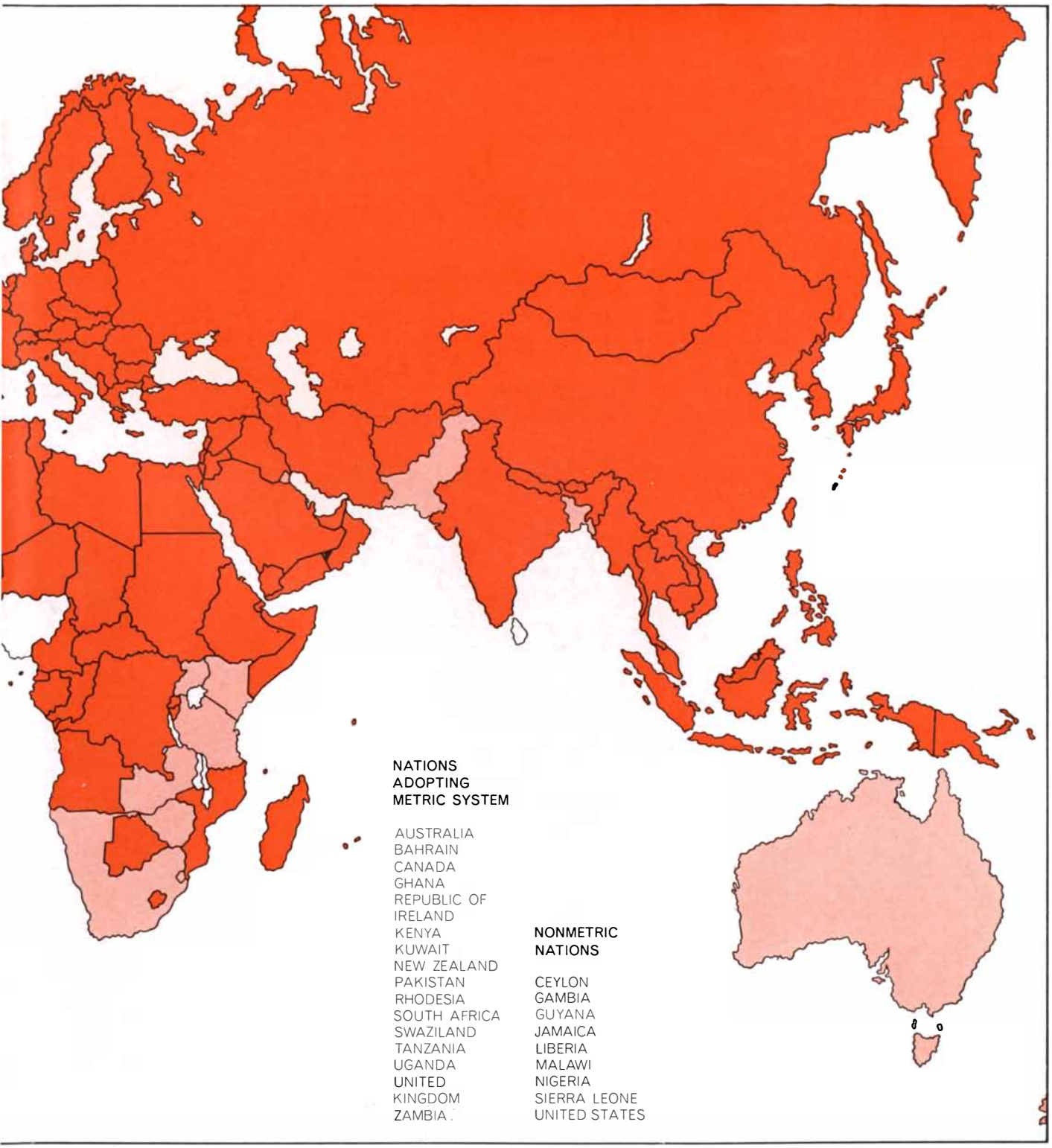
To sum up, SI incorporates and extends the metric system and lays down a basis for deriving from the base units in a coherent way all other units that may be necessary. Less directly, SI points up the fact that the metric but non-SI countries have a system that is not entirely coherent. Coherence imposes a logical relation among units. The product or the quotient of two or more



METRIC WORLD now includes all but a few nations, if one counts the nations that are not formally on the metric system at present but have made a commitment to change to that system. The nations that now use metric standards are depicted in dark color; nations

base units must be the unit of the resulting quantity. For example, the derived unit of velocity must be the meter per second, not the kilometer per hour. The SI unit of force must be the force that, applied to the unit mass, produces unit acceleration. In other words, it is the force that, applied to a mass of one kilogram, gives it an acceleration of one meter per second per second. The unit thus derived is the newton. (I was asked sarcastically at a conference what would happen if I went into a British shop and asked for a newton of potatoes. I said, "Indubitably I would get a ton of new potatoes." In fact, when one buys potatoes or anything else that is sold by "weight," one is really interested in mass, and so the buying will be done by the kilogram.)

It is inevitable that, whatever the size



such as Great Britain that have decided to adopt the metric system or the extended and refined version of it known as the International System of Units, or SI for Système International, are shown in light color. The nations still nonmetric are represented in white.

LENGTH

KILOMETER	(km)	=	621371 MILE
METER	(m)	=	1.09361 YARDS
MILLIMETER	(mm)	=	.0393701 INCH
MICROMETER	(μm)	=	39.3701 MICROINCHES

AREA

SQUARE KILOMETER	(km^2)	=	247 105 ACRES
SQUARE METER	(m^2)	=	1.19599 SQUARE YARDS
SQUARE MILLIMETER	(mm^2)	=	00155000 SQUARE INCH

VOLUME

CUBIC METER	(m^3)	=	1.30795 CUBIC YARDS
CUBIC DECIMETER	(dm^3)	=	.0353147 CUBIC FOOT
CUBIC CENTIMETER	(cm^3)	=	0610237 CUBIC INCH
LITER	(l)	=	.2642 GALLON (U.S.)

VELOCITY

KILOMETER PER HOUR	(km/h)	=	.621371 MILE PER HOUR
METER PER SECOND	(m/s)	=	3.28084 FEET PER SECOND

ACCELERATION

METER PER SECOND PER SECOND	(m/s^2)	=	3.28084 FEET PER SECOND PER SECOND

MASS

KILOGRAM	(kg)	=	2.20462 POUNDS
GRAM	(g)	=	.0352740 OUNCE

DENSITY

KILOGRAM PER CUBIC METER	(kg/m^3)	=	0624280 POUND PER CUBIC FOOT

FORCE

NEWTON	(N)	=	.224809 POUND-FORCE

TORQUE

NEWTON METER	(Nm)	=	737562 POUND-FORCE FOOT

PRESSURE, STRESS

NEWTON PER SQUARE METER	(N/m^2)	=	.000145038 POUND-FORCE PER SQUARE INCH

VISCOSITY (DYNAMIC)

NEWTON SECOND PER SQUARE METER	(Ns/m^2)	=	0208854 POUND-FORCE SECOND PER SQUARE FOOT

VISCOSITY (KINEMATIC)

SQUARE METER PER SECOND	(m^2/s)	=	10.7639 SQUARE FEET PER SECOND

ENERGY

JOULE	(J)	=	737562 FOOT POUND-FORCE
KILOJOULE	(kJ)	=	277778 WATT-HOUR

KEY CONVERSIONS of common SI units to units of the imperial system are shown on the basis of data published by the National Physical Laboratory of the Ministry of Technology in Great Britain. The term "pound-force" appearing under such headings as "Force" and "Pressure" is employed in this scheme in place of such somewhat different concepts as foot-pounds and pounds per square inch.

of the base units and the derived units, they will not be convenient for all applications. Therefore a range of multiplying and submultiplying prefixes will be necessary; associated with the unit, they adjust it to the size required. The General Conference has approved 14 such prefixes, ranging from "atto" (10^{-18}) to "tera" (10^{12}). The most popular prefixes are likely to be "kilo" (10^{3}), "centi" (10^{-2}) and "milli" (10^{-3}). SI purists would point out that the use of such prefixes is a departure from the coherence of multiplying units by units, but the prefixes are at least logical. Moreover, practice demands them. (Although the perfectionists have yielded to the multiple, they reject as heresy such double multiples as "millimicron" for 10^{-9} meter; that short length is a "nanometer.")

The International Organization for Standardization, the authority responsible for the international development of industrial and commercial standards in almost every field of technology, is seeking to discourage the use of prefixes that do not represent 10 raised to a power of three or a multiple of three. The effect of such a practice would be to rule out the centimeter. Engineers may be willing to comply, but to go from the millimeter to the meter would place a severe restriction on many everyday usages. In education, for example, there is a need for submultiples of the meter that are bigger than the millimeter. It does not make sense to tell a five-year-old, who can barely count up to 10, that if a meter is divided into 1,000 parts, each part is a millimeter. It is also to be doubted that a tailor or a dressmaker would work in millimeters, even though the cloth is likely to be measured in meter lengths and millimeter widths.

SI scholars have had to make a further concession by acknowledging that certain widely known non-SI units will continue in service. Examples are the minute and the hour; the degree, the minute and the second as units of angular measurement, and the nautical mile and the knot. In any event it will not be necessary for most people to have a detailed understanding of SI in all its perfection and precision. They will be concerned only with the well-known and long established metric units for length, area, volume and mass.

Britain's decision to adopt the metric system is the recognition of the force of a number of circumstances, including the spread and the speed of the spread of the metric system; changes in the country's export position; closer links between science and technology and industry, and most directly pressure from British industry. It was not always thus. Until not so many years ago Britain dominated the Industrial Revolution and could dictate its specifications. Where Britain's law held, so did its acres, yards, pounds and gallons. There was a certain amount of nostalgia in the adherence to the imperial system. When a member of Parliament, citing the advantages of the metric system that had become apparent elsewhere in Europe, raised the question of going metric, Lloyd George replied: "Do you expect the British workingman to go into a public house and ask for .56825 liter of beer?" That was the end of the debate.

The contemporary phase of British metrication dates from the report in 1950 of the Committee on Weights and Measures Legislation that had been set up by the president of the Board of Trade. The committee concluded that the metric system was "a better system of weights and measures than imperial" and recommended that the government take steps (in concert with the Commonwealth nations and the U.S.) toward complete adoption of the metric system over a period of about 20 years. Industry began to come around, impelled largely by the changing export scene. In 1965 the Federation of British Industries reported to the government that a large majority of its members now supported the metric system. Most of the Commonwealth nations and former British colonies had decided to go metric; the U.S.S.R. was adopting SI, and China used the metric system.

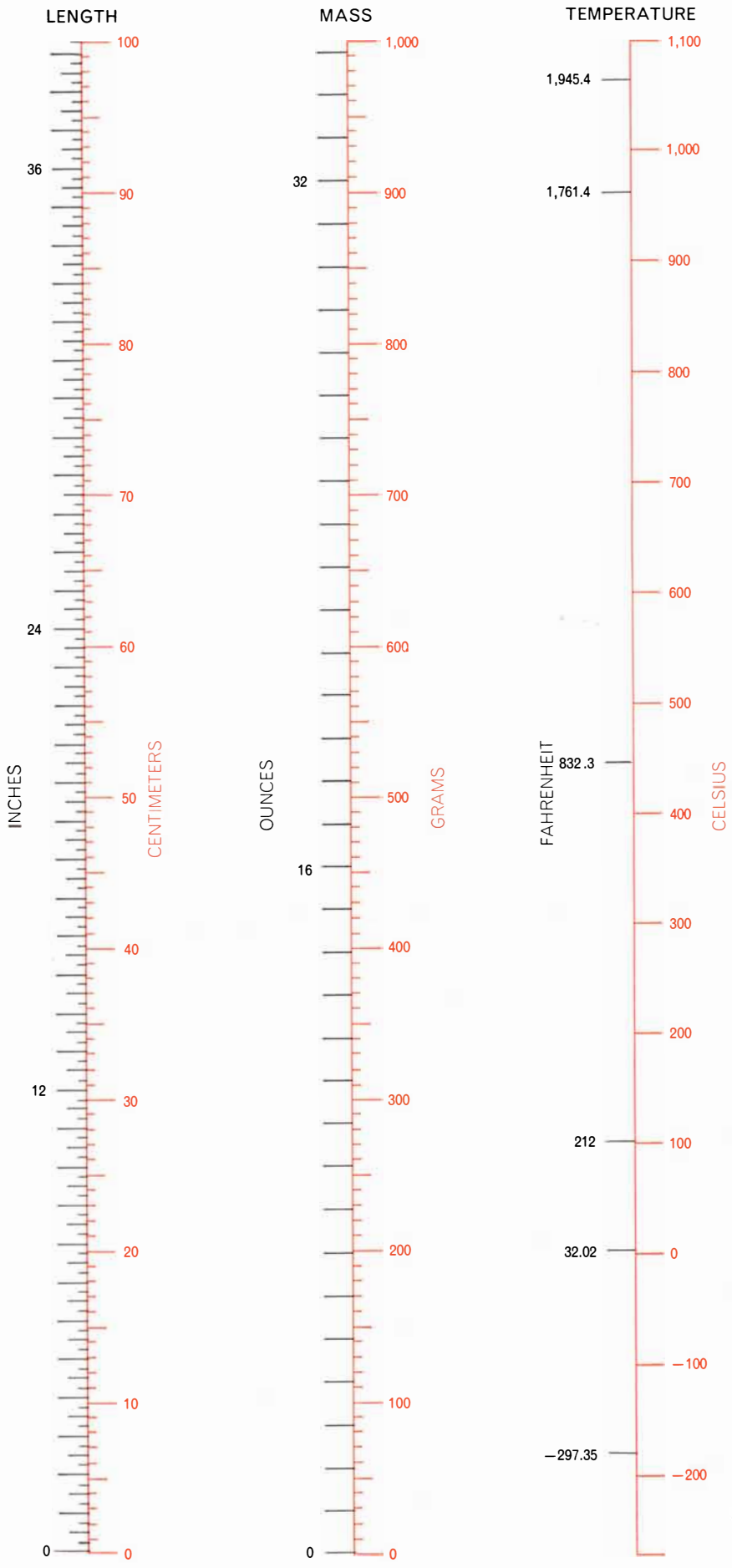
Once the government had made its decision, aiming at 1975 as the year when Britain would be for all practical purposes a metric country, the Minister of Technology set up a Standing Joint Committee on Metrication. Its function was to encourage all sectors of the economy to start planning. The committee also considered the machinery for making the changeover and recommended the creation of a Metrication Board to facilitate the process. The board was duly established; it consists of representatives of journalism, the arts, social research, civil engineering, the central government, local governments, agriculture, mechanical engineering, the universities, the trade unions and women's interests. Under my chairmanship the board held its first meeting in May, 1969.

Everything was to go metric: the construction industry, engineering, transport, the road system, the schools, farming, the land survey, hydrography, the extraction industries, the professions, the whole of distribution and the retail trade and ultimately the kitchen. Nothing—not even that British sanctuary the pub—would remain unaffected.

From our first meeting our board recognized that the task was not nearly as formidable as it seemed. We were taking over a going concern, thanks to initiatives by such groups as the Royal Society, the British Standards Institution, government departments and industry. The Royal Society, in the tradition of British science, which was active in promoting an internationally coherent, scientifically precise system even when imperial units were dominant, held two conferences on metrication for the schools. Initiatives of this kind had determined, before our board was created, that examinations in the higher institutions of learning and in the professions should be conducted under the metric system.

The British Standards Institution, which is a typical British partnership of government and private industry, has a crucial role, since metrication is meaningless unless it is embodied in codes of standards. The Institution, which had actively promoted metrication, began examining and revising more than 4,000 standards so that the pace of change throughout the economy would not be impeded by any lack of metric components. In consultation with the representatives of both large and small firms in its membership, it prepared and published general timetables for the change to the metric system in four key areas: the construction industry, engineering, the electrical industry and the marine industry.

Government departments, some with more energy and enlightenment than others, had given metrication a considerable boost. The departments with big budgets, such as Defence and Technology, can accelerate metrication directly by requiring metric specifications as a condition of purchase, of engineering contracts and of grant-aided projects. Well over 50 percent of building construction is financed from public funds. The Ministry of Public Building and Works set a salutary example not only by specifying the use of metric standards but also by providing guides and publicizing the process. The result is that the construction industry has become a pacemaker for other industries. This does not mean that all difficulties in the industry have been overcome. There are still such problems as ensuring metric standards for raw material and components,



bringing along the many subcontractors and training workmen.

The engineering sector illustrates both problems and opportunities. A major problem is logistics: getting the right parts to the right place at the right time. If a manufacturer adopts the metric system and his suppliers do not, he is going to be in trouble. Conversely, if he waits on his suppliers, he may wait a long time. What is needed is agreed standards and agreed timing. There are also such difficulties as retraining skilled men and the cost of retooling.

These are all valid problems, but from what we know of engineering firms that have adopted the metric system the problems are overstated. With good administration all of them are manageable and the net result is profitable—often quite profitable. Firms have found that the mere exercise of adopting the new system has led to a beneficial reexamination of their habitual methods. With concerted planning and proper warning, suppliers will meet a worthwhile order in metric components. With a metric plan that alerts everyone from the boardroom to the workshop, and one that practices sensible consultation about what is going to happen, the human difficulties are soluble and any retraining is confined to a limited range of operations. When it comes to stocking materials and components, the varieties one needs to carry as a result of working in metric units are so substantially reduced (by 80 percent in some instances) that the saving on inventory far exceeds the cost of carrying imperial-system spares during the transition period. The cost of retooling is misunderstood and sometimes misrepresented by the critics of metrication. Again it is a question of good management. Our case histories from firms that are working on the metric system certainly contradict the pessimists and confirm the optimists.

Having anticipated that we would meet opposition from farmers, we have been agreeably surprised to find them, or at least their organizations, generally cooperative. Before long farmers in Brit-

BASIC UNITS of the metric and imperial systems are compared, with metric units in color and imperial units in black. The reference points shown on the Fahrenheit scale are, reading from the bottom, the boiling point of oxygen, the triple point of water, the boiling point of water, the boiling point of sulfur, the freezing point of silver and the freezing point of gold.

ain will conduct their operations by the hectare instead of the acre. One can only assume that the farmers, having for many years done a great deal of bookkeeping and clerical work in connection with price subsidies, are accustomed to doing arithmetic. The arithmetic will now be much simpler, since it will have in common the factor of 10 for both quantities and prices.

Everyone warns that the trouble will come in the shops, in the bars and in the kitchen, when the public is confronted with kilograms and liters. Perhaps so, but I am optimistic. Custom has already changed drastically. Even the general storekeeper "breaks bulk" less than he once did: he does not often scoop out flour or sugar or put butter on the scales. His supplies, like those of the supermarkets, come already packaged. The success of the metric changeover in the retail trade could lie simply in packaging.

The example of the pharmacies is reassuring. The pharmaceutical industry and the pharmacists began planning to adopt the metric system as long ago as 1945 and changed over in March, 1969, without a tremor. Nobody seemed to notice, not even the older generation of physicians who still prescribed in grains, drams, scruples and minims. The pharmacist deciphered the cryptic prescriptions of the physician as he usually did and then dispensed them in metric units, which the younger physicians were using anyway.

The schools are essential to the changeover. Primary schools were required to adopt the metric system at the beginning of the school year last September. It had been a second numerical language in the other grades for generations. The students now in the primary schools will emerge thinking in metric terms. They should be grateful, because they will have lost a mental rucksack of archaic measuring units. They will have a simple, logical calculating system, which takes far less time than the imperial system to learn and will be the numerical *lingua franca* of the world.

There will inevitably be a period of bilingualism. The difficulty is not in learning the metric system; it is in unlearning the imperial one. It would be unwise to encourage this bilingualism. In Britain we have had a bad example. The weathermen went over to centigrade (which they should be calling Celsius), and radio and television tried to bridge the transition by giving temperatures in both centigrade and Fahrenheit. The result is that everyone still waits for the Fahrenheit figure. With this reminder we do not intend to have road signs give the mile equivalents of distances expressed in kilometers.

During 1970 substantial progress in metrication will be made in many sectors of the British economy. The British Standards Institution expects that all important standards relating to construction, industrial materials, engineering components and equipment will be available. Aluminum, lead, copper, zinc, steel bars, flat steel products, wire mesh and electric cables will be available in metric dimensions this year. Metric bricks, fiberboard, concrete pipe, wood, glass and ready-mixed concrete will be available by the end of the year. The paper, board and printing industries will be metric by then too. The measurement of land in metric units has begun this year and will be changed over completely next year. The engineering industries, including shipbuilding, are in the first stages of the transition. New designs for the armed services are being specified in metric units, and government procurement policy will step up the pace next year. Engineering as a whole should be substantially metric by the end of 1972. Our present thinking is that farming will begin to go metric in 1972 and that the change will be substantially completed in 1973. By the end of 1972 the materials industries will be producing and marketing in metric quantities and dimensions. The freight-transport sector has set the beginning of 1972 as the target date for being metric. In 1973 all posted highway speeds will be in kilometers per hour.

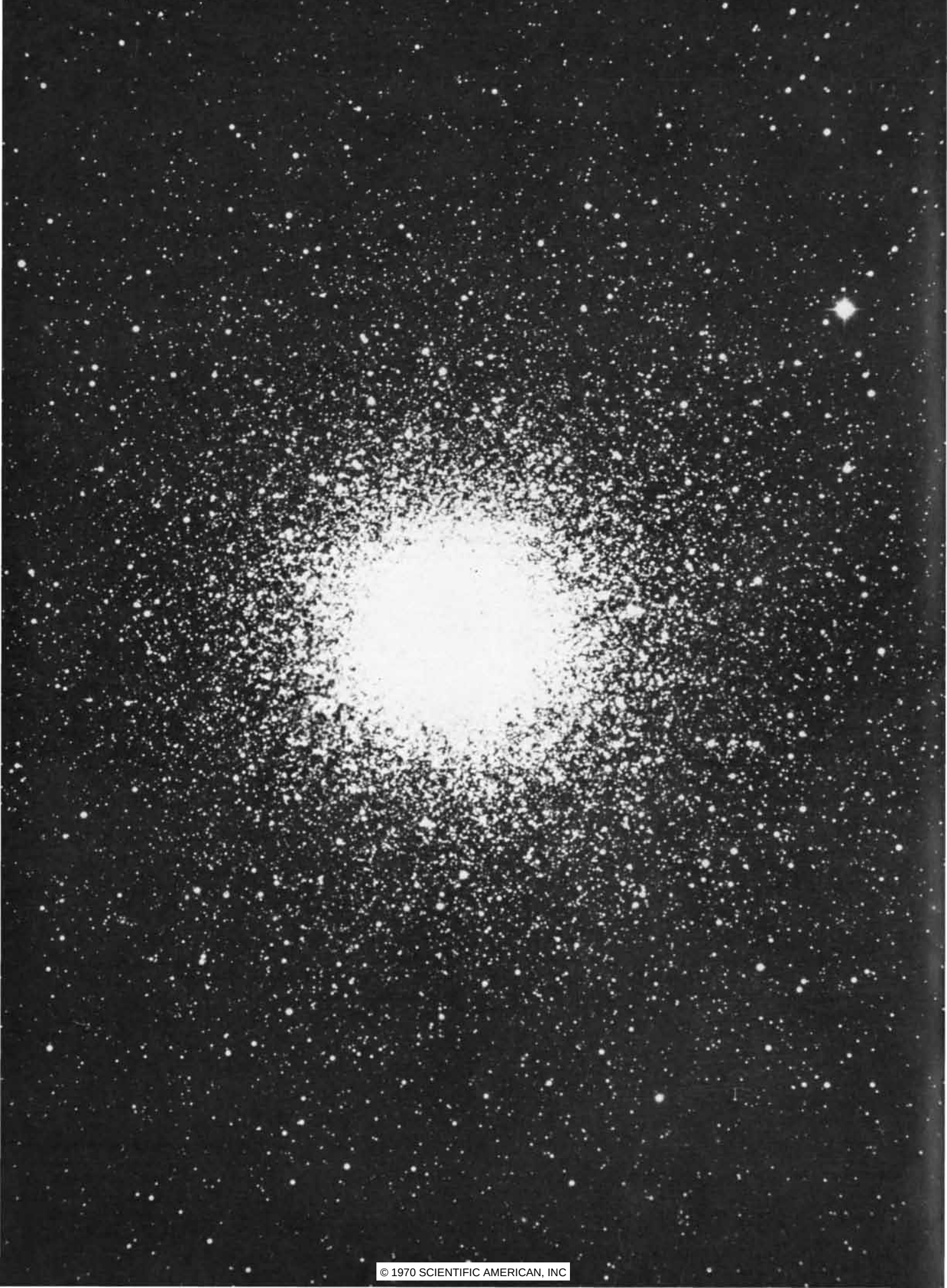
The bedding industry will be metric at the beginning of 1971. Footwear will be expressed in metric sizes in the fall of 1972; so will fibers and fabrics. By about that time the general public will become increasingly aware of the change to the metric system in the shops.

The Metrication Board is charged with seeing that all this activity proceeds as briskly and smoothly as possible. We have the function of a railway marshaling yard. The individual firms, industries or trade associations, under the guidance of the board's steering committees, send in the freight cars. The board does the switching, sets the signals and clears the track so that the train can make its connections with the other sectors. Our only sanction, which is a powerful one, is implicit in the warning that any company or industry that lags behind will be left behind.

Our role is in planning and persuasion. We can speed the laggards, including government departments. We can encourage the enterprising. We must create an understanding that will help the changing groups and reconcile the public to the changes, which, apart from their own merits, are historically, scientifically, technologically and economically inevitable.

PREFIX	SYMBOL	POWER		EXAMPLE
tera	T	10^{12} =	1,000,000,000,000	
giga	G	10^{9} =	1,000,000,000	gigahertz (GHz)
mega	M	10^{6} =	1,000,000	megawatt (MW)
kilo	k	10^{3} =	1,000	kilometer (km)
hecto	**h**	**10^{2} =**	**100**	
deca	**da**	**10^{1} =**	**10**	
deci	**d**	**10^{-1} =**	**.1**	
centi	**c**	**10^{-2} =**	**.01**	
milli	m	10^{-3} =	.001	milligram (mg)
micro	μ	10^{-6} =	.000001	microsecond (μs)
nano	n	10^{-9} =	.000000001	nanometer (nm)
pico	p	10^{-12} =	.000000000001	picofarad (pf)
femto	f	10^{-15} =	.000000000000001	
atto	a	10^{-18} =	.000000000000000001	

MULTIPLYING PREFIXES for use with the basic units of the metric system were approved by the General Conference on Weights and Measures, which promotes the use of the metric system and acts as a center for international cooperation. In nations using SI the four prefixes shown in boldface letters are supposed to be limited to uses where the other prefixes, all representing 10 raised to a power of three or a multiple of three, are inconvenient. Under SI it is also urged that only one prefix should be applied to any basic unit.



GLOBULAR-CLUSTER STARS

Within our galaxy are some 200 globular clusters, each consisting of 100,000 to a million stars. The study of models indicates they began life 10 to 15 billion years ago when the universe was young

by Icko Iben, Jr.

If we could view our galaxy from outside, we would see a giant pinwheel made up of billions of stars rotating slowly around a compact, brilliant nucleus. Looking more closely, we would detect that the spherical volume of space above and below the pinwheel was not empty but was filled with billions of much fainter stars, and scattered about in this "halo" we would also see some 200 fuzzy but brightly glowing globules made up of stars. Close inspection would reveal that each globule consists of from 100,000 to a million stars, most of which are of low luminosity. These are the globular clusters. If one could collapse some hundreds of millions of years into a few minutes, one would see that the 200 clusters are traveling in giant elliptical orbits around the nucleus of the galaxy, closely resembling the old-fashioned picture of electrons whirling around the nucleus of an atom [*see illustration on next page*].

The globular clusters are fascinating astronomical objects in their own right, but what has pushed them to the forefront of theoretical interest has been the recognition that the stars in globular clusters are exceedingly ancient and contain clues to the early history of the universe. Indeed, some of the first estimates of their age were so high (20 to 25 billion years) that they could not be reconciled with the apparently much younger age of the universe, as inferred from the recession velocity of distant galaxies. The velocity measurements imply that all the galaxies emerged from a primordial fireball no more than 12 or 13 billion years ago.

My own interest has been in making theoretical models of stars of various initial masses and compositions to see exactly how they would evolve and how rapidly they would evolve as their nuclear fuel was consumed. From such studies one can hope to refine estimates of the age of globular-cluster stars, to establish if the ages are really inconsistent with the estimated age of the universe and to draw some inferences about the cosmological conditions surrounding their origin. The conclusion, in brief, is that the stars in globular clusters are about as old as the universe is inferred to be, and that the initial abundance of helium in cluster stars is remarkably close to that formed in "big bang" models of the universe. Hence studies of cluster stars provide evidence that the universe has expanded from a tremendously hot primeval atom that was formed in one unique event 12 or 13 billion years ago.

Direct observation of individual cluster stars provides just enough information to make comparison with theoretical models worthwhile. Although one can determine none of the bulk properties (mass, radius or intrinsic brightness), one can estimate the surface temperature from the star's color and judge whether it is more luminous or less luminous than its neighbors. From the star's spectrum one can also estimate the abundance of elements heavier than helium. Unfortunately it is the helium content that most directly determines the course of a star's evolution. Like our own sun, cluster stars are so cool that helium atoms remain in their lowest states of excitation and thus yield little spectroscopic information. Elements heavier than helium, on the other hand, are readily excited, which makes their abundance easier to estimate. It turns out that cluster stars are strikingly deficient in heavy elements: they contain only from a tenth to a hundredth as many heavy atoms as stars of equal mass in the disk of the galaxy do. Moreover, the farther a cluster star (or other halo star) is from the center of the galaxy, the more deficient it seems to be in heavy elements.

Knowing the surface temperature of stars in a given cluster and their relative brightness, one can plot the distribution of cluster stars in a two-coordinate scheme known as a Hertzsprung-Russell, or H-R, diagram [*see illustration on page 29*]. The horizontal axis in an H-R diagram represents surface temperature (as inferred from its color); the vertical axis represents luminosity. After plotting a few thousand stars one sees that the great majority fall in a simple linear pattern (the "cluster locus") that rather resembles the head and beak of a bird. The head is thrown back so that the beak points to the upper right corner of the diagram. The tip of the beak is the tip of the "red giant" branch of the cluster locus: the branch containing the coolest (reddest) and brightest stars in the cluster. The shoulder and back of the bird consist of stars 100 to 10,000 times less luminous, whose surfaces range up to 2,000 degrees hotter on the Kelvin, or absolute, scale. Since the great majority of cluster stars (as well as galactic-disk stars) fall in this range, they are known as main-sequence stars. (The sun falls in the middle of the main sequence.) The breast of the bird, which forms an almost horizontal line in the diagram (the "horizontal branch"), is made up of stars approximately 50 times brighter than the sun, with temperatures ranging from the sun's temperature (6,200 degrees K.) to

GLOBULAR CLUSTER M 3 in the constellation Canes Venatici (*opposite page*) is one of the brighter clusters visible in the Northern Hemisphere. This cluster contains some 500,000 stars and is thus a typical member of its class. Photograph was taken with the 200-inch Hale telescope on Palomar Mountain.

more than 11,000 degrees K. In this branch stars with surface temperatures between 6,500 and 7,500 degrees fall within the "instability strip." Known as RR Lyrae stars, they dim and brighten rhythmically at precise intervals ranging from six hours to about a day.

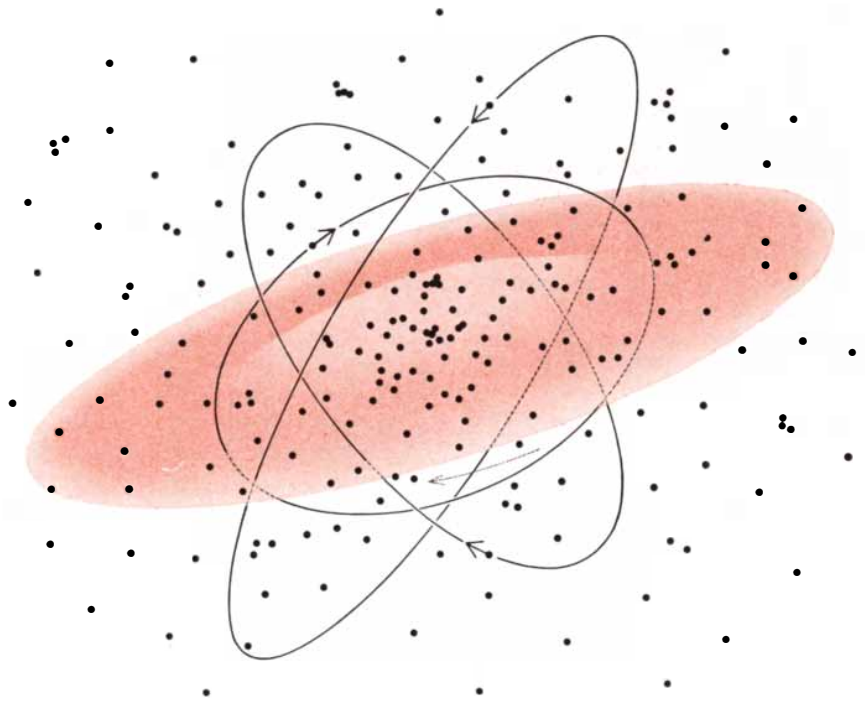
If one assumes, as seems reasonable, that all the stars in a cluster are nearly the same age, how can one account for their distribution along the cluster locus? How do they differ from one another? Since all seem to have about the same content of heavy elements, it is reasonable to assume they all started with the same relative abundance of hydrogen and helium. This leaves only mass as the probable variable.

This assumption, however, does not tell us whether mass simply correlates with a specific position on the locus, which would imply that a star is born to its "station" in life, or whether mass determines a rate of evolution through a succession of stations on the locus. Since stars must "burn" their substance in order to shine at all, the latter proposition seems more likely. We still do not know, however, in which direction a star may travel along the locus or whether stars at birth can appear at many different points or at only a few. To answer such questions one must construct model stars and see what kinds of answer emerge. Finally, model studies should be able to tell us if stars of the same age but of different initial mass would indeed distribute themselves as actual stars in an H-R diagram do.

In these idealized models we suppose that a star is a perfect sphere composed of a sequence of concentric shells. At any given instant, as a consequence of simple physical laws, the temperature, pressure and composition will vary from one shell to the next. As the star evolves and exhausts a given nuclear fuel, the shells change in physical structure and composition. During a star's lifetime the temperature in the core can vary from 10 million to a billion degrees K. Densities in the core can reach millions of grams per cubic centimeter. Matter inside the star is predominantly in the form of a plasma, consisting of free atomic nuclei and free electrons. At several stages in a star's life the electrons in the core become "degenerate," a quantum-mechanical state in which the frequency of electron collisions depends less on temperature than on electron density. Under these conditions the pressure in the core can be maintained not by an increase in temperature but solely by an increase in density. The significance of this will become clear as we proceed.

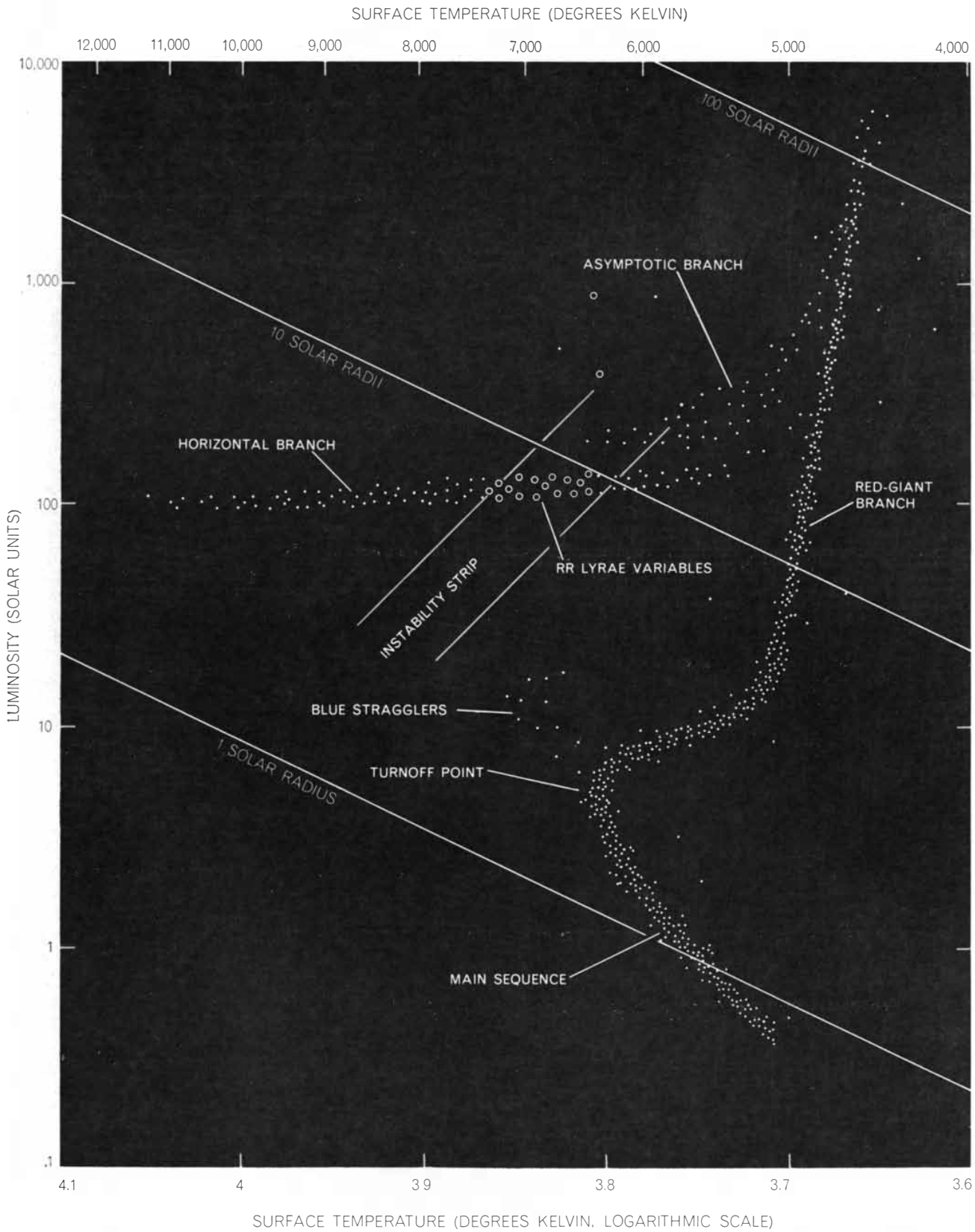
The design of a stellar model begins with a specification of composition and total mass. The symbols X, Y and Z are respectively used to indicate initial abundance of hydrogen, helium and heavier elements. For a typical globular-cluster star we can let $X = .699$, $Y = .3$ and $Z = .001$. (The choice of .3 for Y is actually the consequence of many such computations, as the reader will discover.) Let us also assume that the model cluster star has eight-tenths the mass of the sun. When it begins to burn nuclear fuel, it has .7 the sun's radius and luminosity [*see illustration on page 31*]. As in the sun, the principal nuclear reaction is the fusion of two protons (hydrogen nuclei) to form a deuteron, the nucleus of deuterium (heavy hydrogen). Subsequently a deuteron and a proton combine to create a light helium nucleus. Two light helium nuclei then combine to form a heavy helium nucleus, releasing two protons in the process. The net result is that for every heavy helium nucleus formed, four protons disappear.

In our model star significant amounts of energy are released throughout a fairly large central region. Near the center of this region and in much of the stellar envelope energy is transported by convection, by matter in turbulent motion. Throughout the main volume of the star, however, energy is carried outward by photons, which are scattered, absorbed and reemitted. This process is called radiative diffusion. As photons approach the surface they are scattered less and less until they finally emerge from the star in straight lines. Although we do not yet have a theory to predict how fast material particles boil off the surface, there is spectroscopic evidence that red giants may boil off as much as a tenth of a solar mass in 100 million years, or perhaps 10,000 times the rate at which particles leave the sun in the "solar wind."

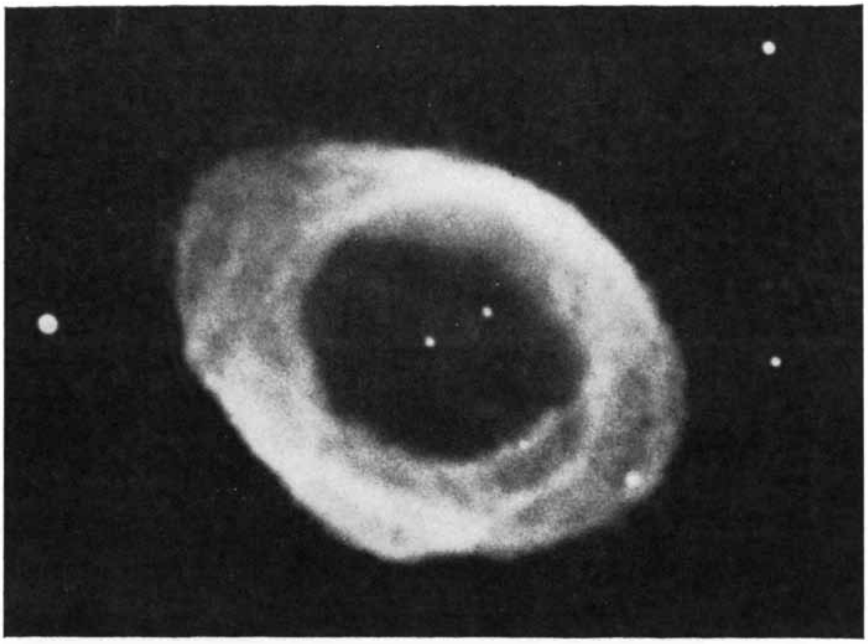


SCHEMATIC VIEW OF OUR GALAXY indicates hypothetical orbits for several of the globular clusters that evidently had enough mass to condense into gravitationally bound groups before the main body of the galactic disk was formed. The disk consists of about 100 billion stars, has a diameter of about 100,000 light-years and makes one revolution in about 200 million years. The globular clusters and several billion individual stars of low luminosity form a "halo" around the disk. Cluster stars and individual halo stars, unlike disk stars, are notably deficient in elements heavier than helium. This characteristic and others lead to the conclusion, based on theoretical models, that cluster stars (and presumably the individual halo stars as well) are older than the brightest stars that are found in the galactic disk.

What happens as time passes and the supply of hydrogen in the center of the star is depleted? The key to the answer rests on the simple fact that the fusion of four protons to create one helium nucleus reduces the number of free particles in the star's interior. At a given temperature the pressure in the interior is directly proportional to the number of colliding particles. With fewer particles in the center the star begins to contract until the compressive force of gravitation is exactly balanced by the resisting force,



TEMPERATURE-LUMINOSITY DISTRIBUTION of the stars in a typical globular cluster is plotted in a Hertzsprung-Russell (H-R) diagram. The surface temperature of the star is derived from a formula based on color: the redder the cooler, the bluer the hotter. Since the absolute luminosity of a star cannot be determined directly, the vertical scale simply indicates relative brightness. The mean curve defined by the distribution is the "cluster locus." A star begins to shine when the temperature in its core becomes high enough to fuse hydrogen nuclei into helium nuclei. At this stage it makes its appearance on the main sequence. As the star evolves it sometimes expands and sometimes contracts. On various grounds it is believed all the stars in a globular cluster are nearly the same age. They are distributed throughout the cluster locus, as we see them today, because some stars (the most massive ones) evolve much faster than others. The author's studies have helped to establish the rate of evolution in the globular-cluster stars.



RING NEBULA in the constellation Lyra is one of the most famous planetary nebulas. It consists of a small, hot star surrounded by a shell of glowing gas. The shell may represent material that was ejected during an unstable stage toward the end of the star's evolution.

which is a product of the particle density and temperature. As particles fall toward the center of the star they acquire kinetic energy from the gravitational field. This energy, in turn, is converted by collisions into the energy of thermal motion. Thus the temperature rises along with the particle density.

As a result of the increased temperature in the contracting core the fusion of the remaining hydrogen proceeds more rapidly, releasing more energy that must find its way to the surface. To accommodate the increased flow of energy the outer shell of the star is forced to expand. Hence the apparent paradox: as the interior of the star contracts, the surface volume increases.

This sequence of events, which can be traced in a stellar model, has its counterpart in the H-R diagram. While hydrogen still burns at the center the luminosity of the star increases more rapidly than its radius does. The surface temperature therefore rises, and the track of the star in the H-R diagram is upward and to the left [*see illustration on page 32*].

When all the hydrogen at the center is consumed, the outer envelope begins to expand more rapidly than the luminosity increases. When this happens, the star has arrived at the "turnoff point" where it leaves the main sequence and starts on its way toward the red-giant branch.

The rising pressure in the growing core, where hydrogen is exhausted, causes the electrons in the core to become increasingly degenerate. Two new phenomena become important. First, energy can flow efficiently through a "sea" of degenerate electrons by conduction, just as heat flows through a metal bar. Thus the temperature rises at about the same rate everywhere in the core. The second consequence is that within the core neutrinos and antineutrinos start being produced in pairs in huge numbers. Their source is a plasma process that cannot readily be duplicated in terrestrial laboratories. Because neutrinos can pass through ordinary matter almost as if it were not there, they escape unimpeded from the core of the red giant, carrying the energy that produced them. The curious net effect of the escape of neutrinos is that energy leaves the center so rapidly that temperatures do not rise as fast as they do somewhat away from the center, where the degeneracy is not so great.

If one makes a computer model of a typical red giant 70 percent as massive as the sun, one finds that the core temperature rises steadily from about 30 million degrees to 80 million degrees, a temperature at which the fusion of helium is kindled. Until that time comes hydrogen remains the nuclear fuel, but at the higher temperatures now prevailing the fusion process proceeds predominantly by the carbon-nitrogen cycle. In this cycle a carbon nucleus absorbs four protons in sequence; along the way two of the four protons are converted to neutrons by the emission of positrons. Finally a helium nucleus splits off, releasing a carbon nucleus to repeat the cycle. (The cycle owes half of its name to the transient appearance of an isotope of nitrogen.) Since there is no more hydrogen left in the core, the carbon-nitrogen cycle proceeds just outside the core in an exceedingly thin shell (about 2,000 miles thick); the radius of the shell is only about five times the radius of the earth [*see illustrations on pages 34 and 35*].

Because a young red giant consumes hydrogen at a high rate it evolves rapidly up the giant branch of the H-R diagram, growing steadily brighter. Throughout this period the size of the core remains constant while the envelope of the star expands by a factor of 10 or more as it grows slightly cooler. One can think of the hydrogen-burning shell in a red giant as a furnace fixed in space that draws in fuel from above and deposits the ashes (helium) in a bin (the core) that steadily grows denser and hotter.

We find in our studies that a star seven-tenths as massive as the sun requires about three billion years to move from the turnoff point on the H-R diagram to the tip of the red-giant branch [*see illustration on page 32*]. Thus we have evidence, based on physical principles, that the direction of stellar evolution indeed proceeds from the main-sequence portion of the H-R diagram to the red-giant portion. We also discover that if one considers a model only slightly more massive, say a star of .8 solar mass, the time needed to evolve from the turnoff point to the base of the red-giant branch is much shorter. A star of .7 solar mass traverses this portion of the H-R diagram in about 2.5 billion years; a star of .8 solar mass takes only half as long. Both stars, however, spend very nearly the same amount of time in the red-giant stage: about 300 million years.

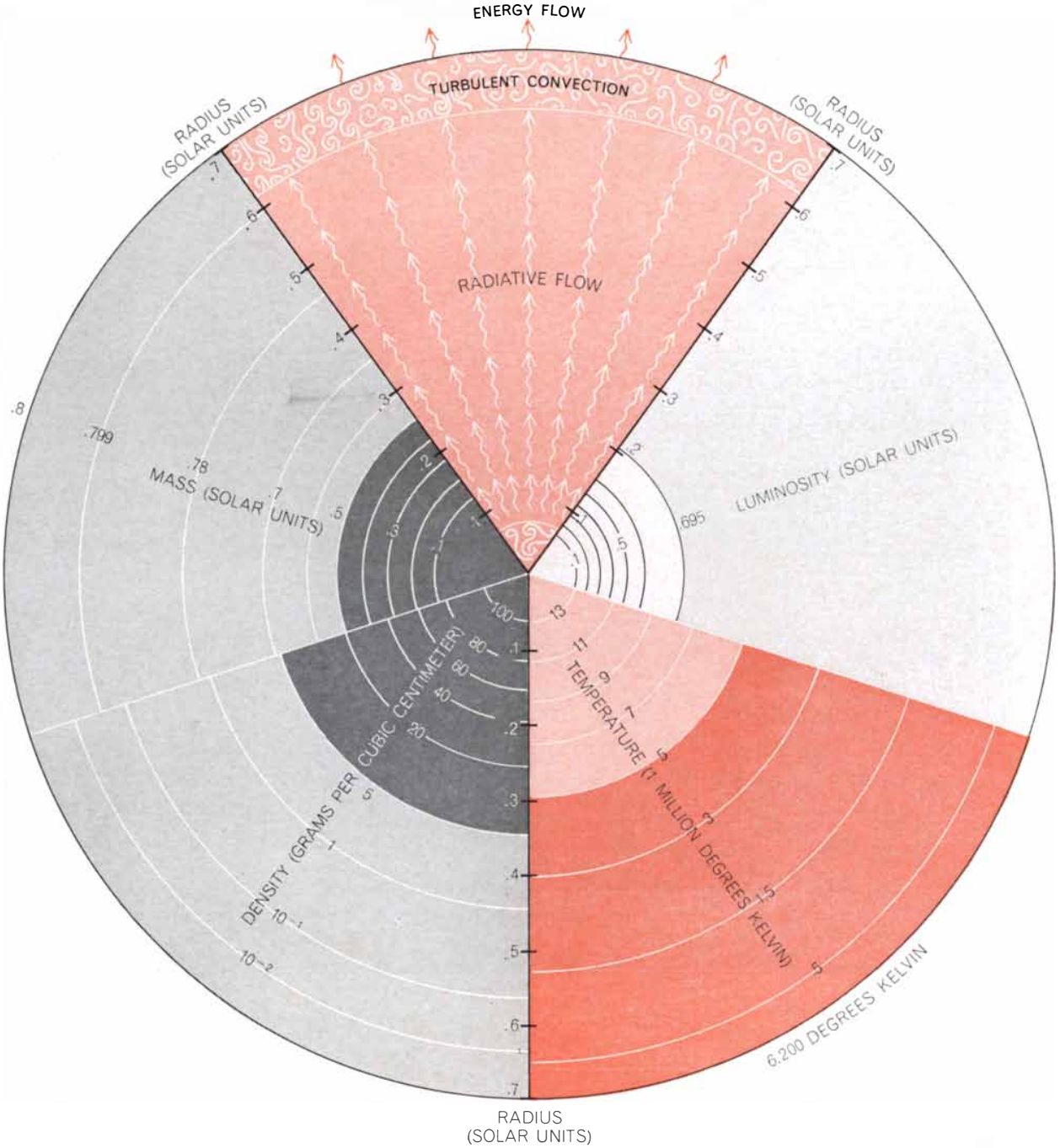
We have now reached the stage in a star's evolution where changes occur exceedingly rapidly. When this brief period ends, a star that has reached the tip of the red-giant branch reappears on the horizontal branch of the diagram, where, regardless of surface temperature, it is about 40 times brighter than the sun and about 25 times brighter than it was at the turnoff point. The pathway to this new location involves violent changes in the star's structure.

When the core of the red giant, which has been steadily growing in mass and getting hotter, approaches half the sun's mass, helium fusion is kindled. Pairs of helium nuclei fuse into carbon nuclei

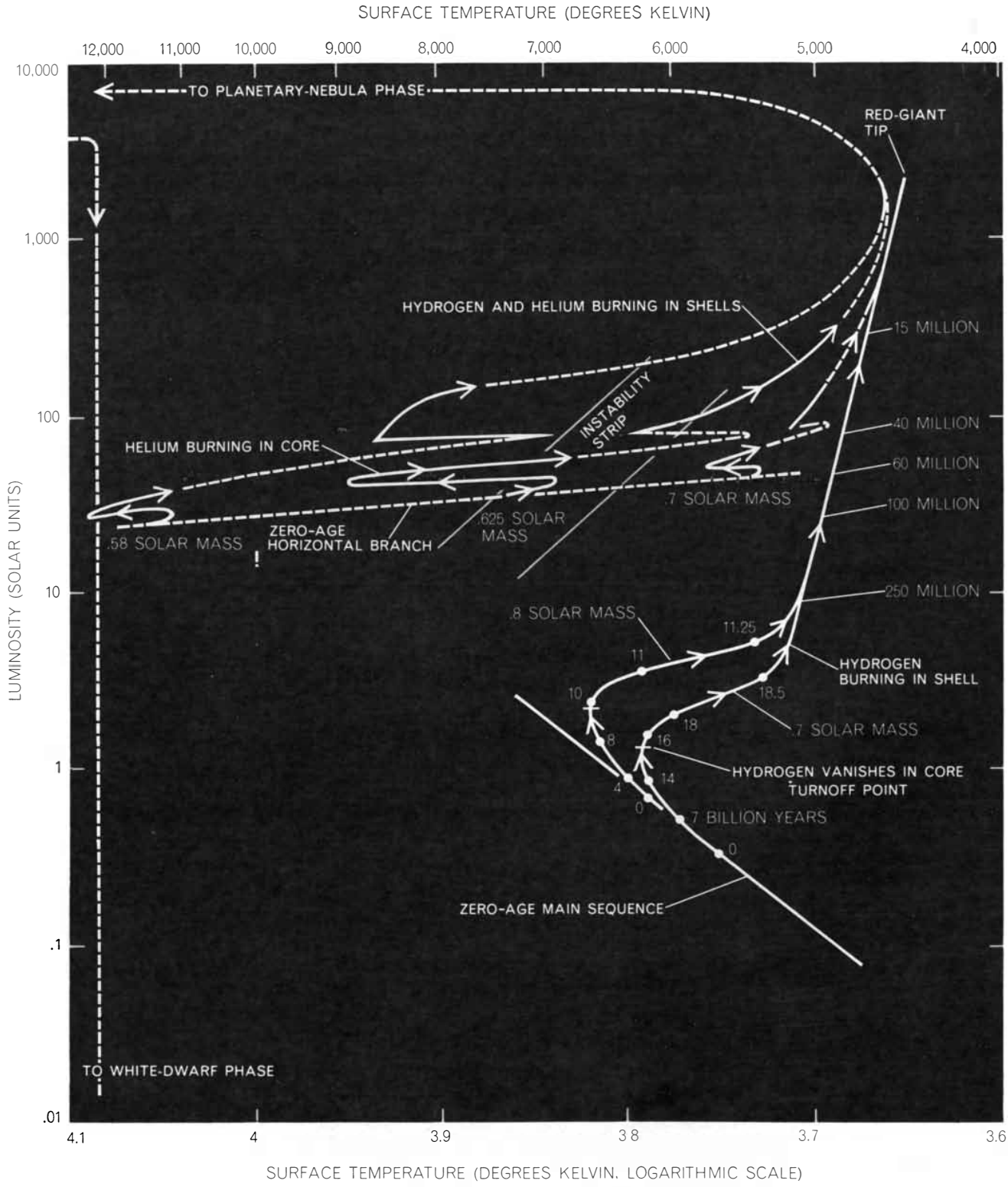
and some of the carbon nuclei take up another helium nucleus to form oxygen. With further slight increases in temperature and density more energy is produced than can be carried away even by degenerate electrons. The core becomes hotter and fusion proceeds still faster.

At first the pressure that supports the core is primarily due to degenerate electrons and is therefore nearly independent of temperature. Eventually, however, the portion of the electron pressure that increases with temperature begins to dominate the temperature-independent portion. At that point the core expands rapidly until the electrons are no longer degenerate.

The entire process is described as the "helium flash." At its peak the core temperature exceeds 300 million degrees; when the flash ends, the core temperature falls back to about 100 million degrees. The star, no longer a red giant, now embarks on a quiet period in which helium burns at the center of the core and hydrogen continues to burn at the edge of the core [*see illustrations on*



MAIN-SEQUENCE STAR is represented in a pie chart that shows how different properties vary with distance from the center. This model represents a typical globular-cluster star containing 69.9 percent hydrogen, 30 percent helium and .1 percent of elements heavier than helium. Its surface temperature is about 500 degrees Kelvin higher than the surface of the sun. Since it is only eight-tenths as massive as the sun it has only seven-tenths the sun's radius and luminosity. Luminosity is a direct function of energy production, 99 percent of which takes place in a burning hydrogen core that contains half of the star's total mass. In the very center of the core matter is mixed by turbulent convection. Throughout most of the star's volume energy is carried by radiative flow, a process involving the repeated absorption, reemission and scattering of photons. Turbulent convection reappears in the outer shell of the star.



TRACKS OF MODEL STARS in an H-R diagram closely follow the pattern of actual stars in a cluster locus (*see illustration on page 29*). "Zero-age main sequence" defines the location of all homogeneous models of a given composition that are burning hydrogen in their center. A model of .7 solar mass exhausts its central hydrogen and reaches the "turnoff point" (*horizontal bar*) in slightly less than 16 billion years. A model of .8 solar mass reaches the turnoff much more rapidly: in about 10 billion years. The two evolutionary tracks merge as both stars begin moving up the red-giant branch of the diagram. In this stage hydrogen burns in a thin shell surrounding a core of inert helium. The time required by either model to reach the tip of the red-giant branch from various lower points is shown in millions of years. At the tip of the red-giant branch the helium in the core of the model star becomes hot enough to burn, creating a violent "helium flash," after which the model turns up on the "zero-age horizontal branch," defined as the position of all stars of the same composition that have begun to burn helium quietly. The location of stars on this branch is very sensitive to mass. The models traverse the initial solid portion of the tracks in about 60 million years; they traverse the succeeding broken-line portion in only about five million years. At the end of that time the star's central helium is gone and energy is provided by helium and hydrogen burning in separate shells. The final broken portion of the track on the diagram ("To planetary-nebula phase") is speculative.

pages 36 and 37]. The hydrogen-burning shell is much thicker but also much weaker than it was during the final years of its red-giant period. The helium-burning core and the hydrogen-burning shell make roughly equal contributions to the star's luminosity. For a typical cluster star of .625 solar mass the phase in which helium burns in the core lasts about 60 million years.

At some time in this period a star of about .6 to .625 solar mass will pass through the instability strip, where, as an RR Lyrae star, it will dim and brighten periodically. The rhythmic variation can be traced to instability in a zone near the surface where helium and hydrogen are only partially ionized. The degree of ionization depends sensitively on pressure and temperature. Any instability that increases ionization subtracts energy from the flow of energy reaching the surface and the star dims. Ionization, however, increases the number of free particles in a given volume, so that the shell simultaneously starts to expand. As the shell expands it cools, and electrons recombine with nuclei. The recombination releases energy, and the star again brightens. As the star keeps "hunting" unsuccessfully for a stable balance its luminosity fluctuates.

A star of .625 solar mass will enter and leave the instability strip more than once as it moves back and forth across the horizontal branch, swinging first to the cooler red end, then back to the hotter blue end and finally back to the red. During this last swing the star develops a thick helium-burning shell in addition to its hydrogen-burning shell. Because of numerical complexities that arise, our understanding of what happens thereafter is not complete.

Evidently instabilities may become so severe that the outer hydrogen shell can collapse while the inner helium shell is expanding. A little later the two shells reverse their direction. One can see that if an oscillation develops and becomes strong enough, the material lying outside the helium shell may be thrown violently out into space. This is perhaps how a planetary nebula is formed [*see illustration on page 30*]. The remnant star may then survive as a brilliant blue star several thousand times brighter than the sun and with a surface temperature of more than 12,000 degrees. Thus it will lie outside the limits of the H-R diagram as depicted in the accompanying illustrations. We cannot predict whether or not all stars of low mass will eject an expanding nebular shell.

Even if no mass is ejected, a star as light as .625 solar mass cannot become hot enough to initiate the fusion of the carbon and oxygen nuclei that remain after all the hydrogen and helium are consumed. The star has no choice but to contract under the influence of gravitation until most of its interior electrons are degenerate. During contraction the star's luminosity is due almost entirely to the release of gravitational energy. The collapse is finally halted by the pressure of degenerate electrons. The only source of energy then remaining is the heat stored in carbon and oxygen nuclei. The star now dims rapidly along a track of constant radius. (The luminosity along the track is too low to be shown in the illustration on the opposite page.) The star has turned into a white dwarf with a radius scarcely larger than the radius of the earth. In this condition it emits so little radiation, even though its surface is hotter than the sun, that it can shine for billions of years.

From such model studies we can now say with some confidence that the way stars in a globular cluster distribute themselves in an H-R diagram is a consequence not of differences in age or of initial composition but solely of differences in initial mass. We have also followed the changes in surface temperature and luminosity of our models as they evolve, step by step, through the H-R diagram. We can now interpret the important features of the cluster locus and identify the evolutionary stage of a star at any particular point.

Let us imagine a cluster made up of model stars of many different initial masses but all the same age. For concreteness we shall assume that this common age is about 12.5 billion years [*see illustration on page 38*]. The main sequence is still populated by stars lighter than .7 solar mass. Stars that are only slightly heavier, about .75 solar mass, have reached the turnoff point, having exhausted the hydrogen in their core. Stars between .75 and .78 solar mass initially have rounded the turnoff and are moving rapidly up the red-giant branch. Stars of .8 solar mass initially have already been through a helium flash and are no longer red giants.

The horizontal branch is populated by stars that began life with at least eight-tenths the mass of the sun. Because they have lost mass along the way their present masses lie between about .6 and .8 solar mass. In general the red stars on the horizontal branch are more massive than the blue ones.

Finally we see that some of our stellar models are burning both helium and hydrogen in thick shells. In the H-R diagram these stars are spread in a broad band, the "asymptotic branch," that slants upward from the red end of the horizontal branch toward the red-giant branch. The area of this band and the distribution of stars within it depend on the amount of mass lost by these stars during their red-giant phase.

Hence we are able to account for the characteristics of most of the stars in a real cluster. Perhaps the most puzzling stars to explain are the "blue stragglers," a small number of stars that seemingly refuse to turn off the main sequence. Actually they may represent a brief transitional stage between the red-giant and horizontal-branch phases. Another suggestion is that they may be rare double stars so close together that their evolution is radically affected.

With this background we are now in a position to estimate how much helium was present in cluster stars at the time of their formation. Once the helium abundance is estimated we can use that value to make model stars coincide with the relative positions of actual stars in H-R diagrams of clusters and from this derive an estimate of the age of real clusters to see whether it conflicts with or supports estimates of the age of the universe derived from other evidence.

Of particular importance in making helium-abundance estimates is the upper section of the red-giant branch, beginning at a point directly opposite the horizontal branch. This point can serve as a luminosity reference point in comparing model stars with actual ones.

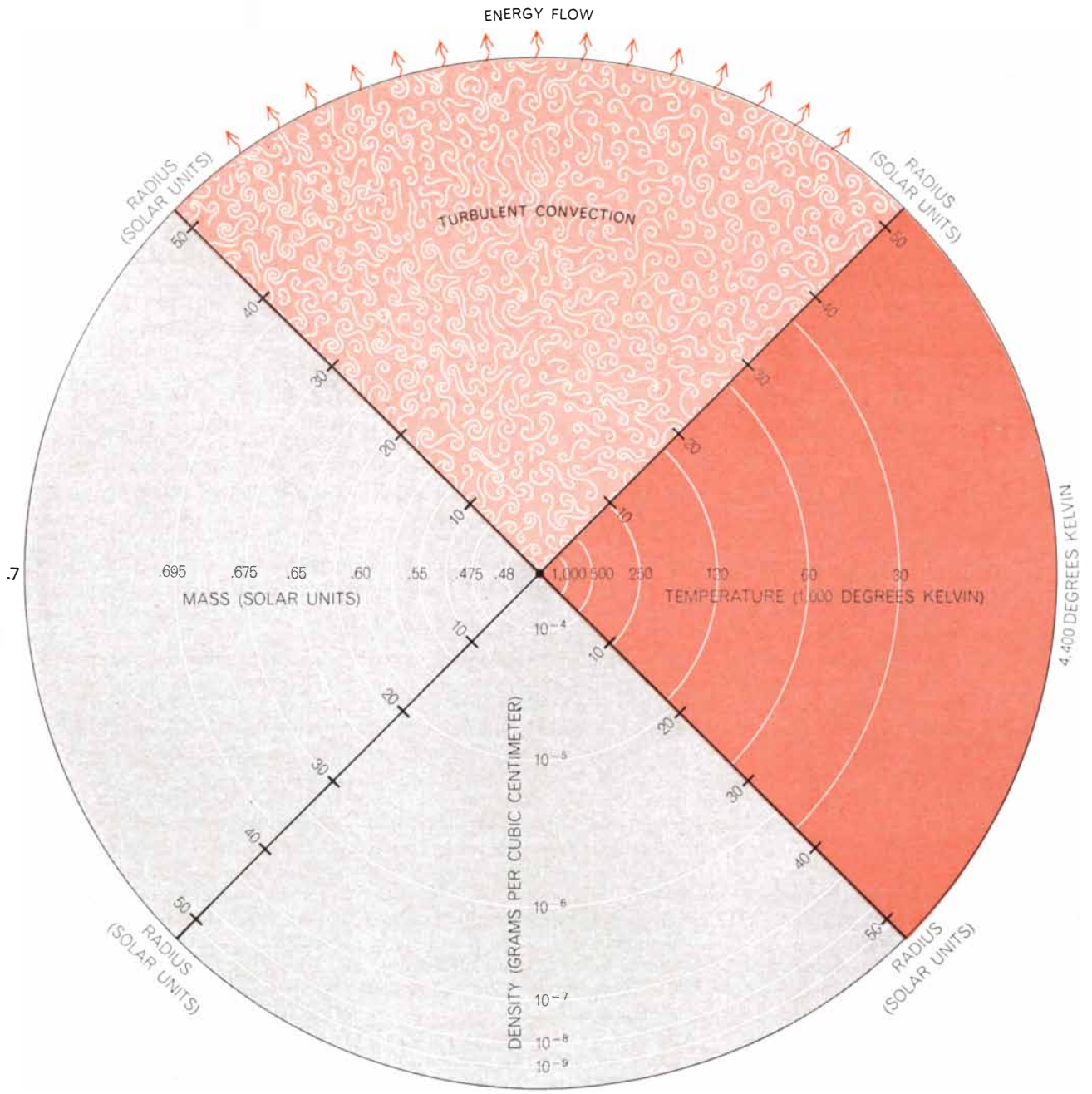
When astronomers sample the stars in a cluster to plot a cluster locus, they find that the great majority fall within the main sequence. Fewer than .1 percent fall on the segment of the red-giant branch that is of particular interest; a comparable number occupy the horizontal branch. The reason for the sparsity of stars in these regions is simply that as a star becomes older, the rate at which it moves along its evolutionary track accelerates.

It is a fair assumption that the number of stars in a given segment of the H-R diagram is proportional to the time a single star spends there. Moreover, the rate at which stars leave a given segment must be roughly equal to the rate at which stars arrive from a preceding segment. Direct observation shows that the ratio of the number of horizontal-branch stars to the number of red-giant-branch stars lies between .8 and 1. These, at least, are the ratios for the only two clusters that have been adequately stud-

ied. Partial data for 10 other clusters fall in the same range.

The trick now is to adjust the composition of theoretical models so that their rate of evolution provides a steady-state cluster in which the ratio of horizontal-branch stars to giant-branch stars matches the observed ratio: between .8 and 1. The important composition variables in these models are *Y*, the fractional content of helium, and *Z*, the fractional content of heavier elements. (Hydrogen, *X*, then constitutes the remainder.) Fortunately spectroscopic observations provide direct information about Z. The observed abundances of heavy elements lie between .001 and .0001 (between .1 percent and .01 percent). Using these values as limits in our models, we can obtain the desired ratios of horizontal-branch stars to red-giant-branch stars by adjusting the helium content, *Y*, between .26 and .32 (26 and 32 percent). The mean value, .29, represents the best guess for the initial helium abundance for stars in globular clusters [*see upper illustration on page 39*].

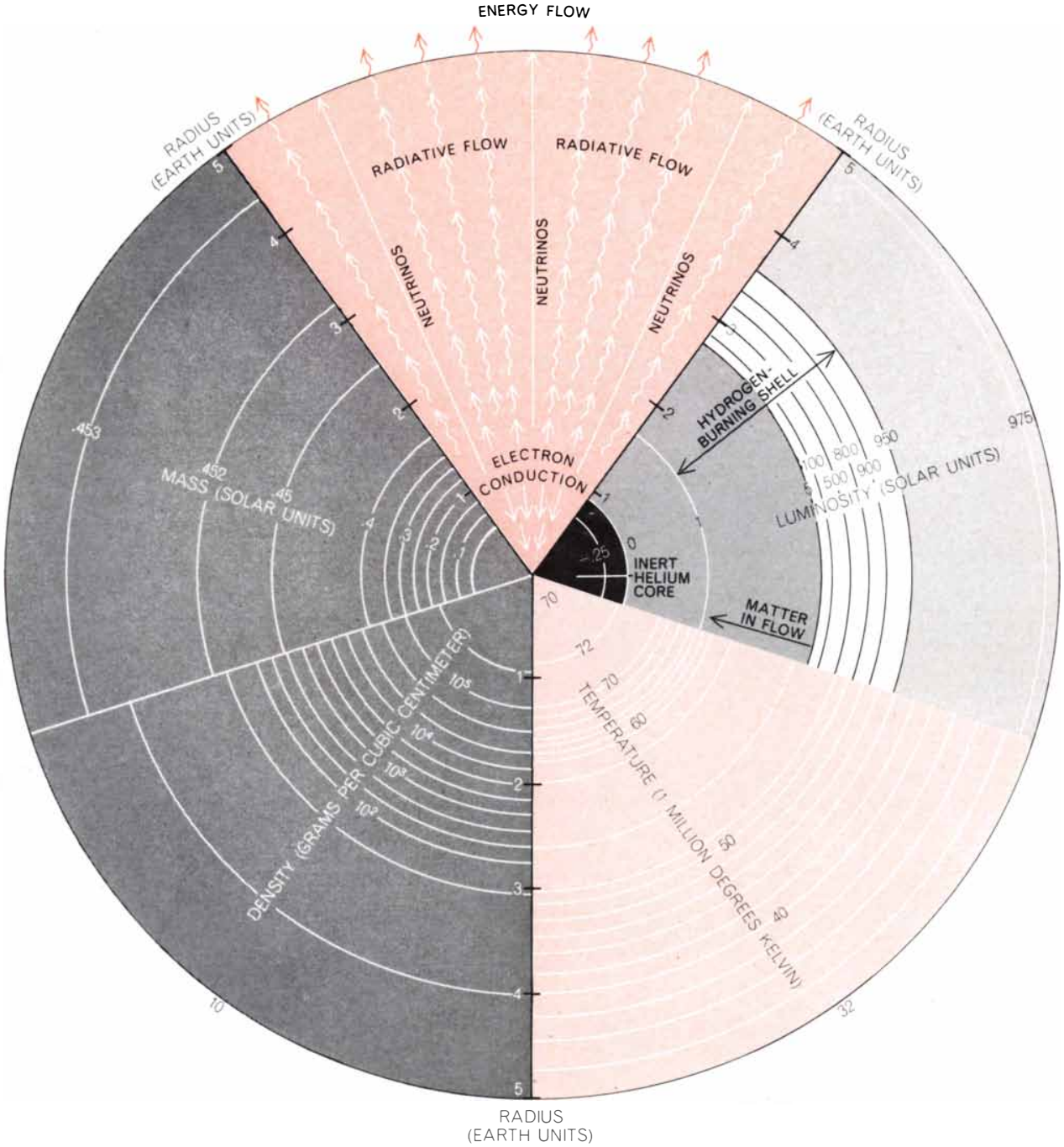
Now we are ready to estimate age. Here the critical ratio is between the luminosity of stars on the horizontal



RED-GIANT MODEL shows the dramatic changes that occur in the stellar envelope when a star of .7 solar mass exhausts the hydrogen in its center and turns off the main sequence. Now burning hydrogen in a thin shell, it enters a new stage of life on the red-giant branch of the H-R diagram. Whereas its radius had been smaller than the radius of the sun, it now expands rapidly to more than 50 solar radii and its luminosity jumps more than a thousandfold. The young giant is now 975 times brighter than the sun. The surface temperature, however, drops nearly 1,500 degrees, so that the color of the star is now red rather than yellow. Most of the star's mass and its entire energy source are concentrated within the small dot in the center of this illustration (*see illustration on opposite page*).

branch and the luminosity of more massive stars of the same age that have just exhausted the hydrogen at their centers and thus are just reaching the turnoff point. In actual clusters horizontal-branch stars are 25 times brighter than stars at the turnoff. In our model clusters this ratio of brightness varies with the age of the model: the older the cluster, the greater the ratio. When we choose values of Z, as before, between .001 and .0001, and introduce values of Y between .26 and .32, we obtain the desired luminosity ratio when the age of the model cluster lies between 10 and 15 billion years [*see lower illustration on page 39*]. The most probable value for the age of the cluster is 12.5 billion years with an uncertainty of ± 3 billion years.

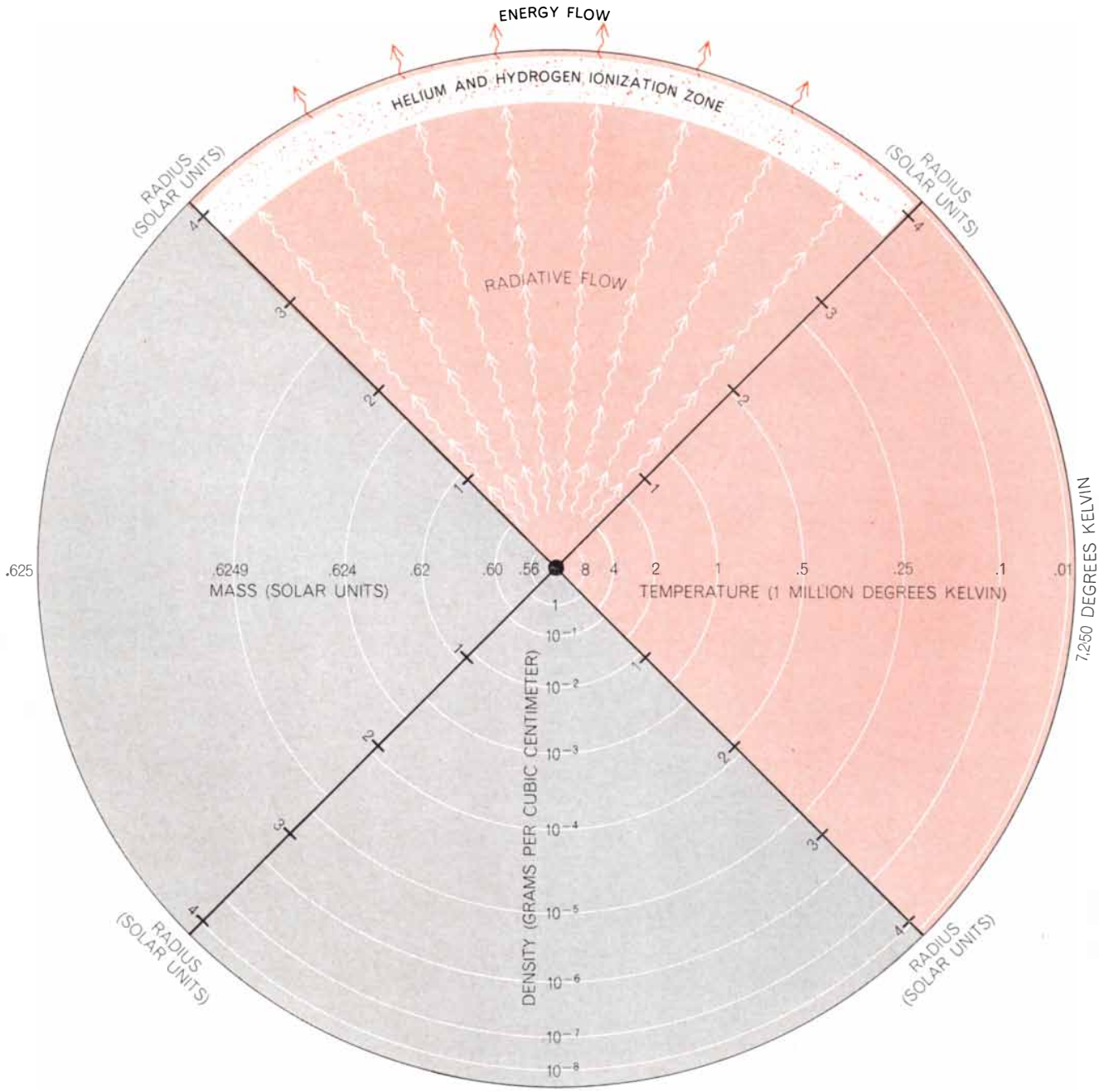
Even allowing for the uncertainty in these age estimates it seems clear that globular clusters, and probably other halo stars as well, are among the oldest objects in our galaxy. We know that metal-rich stars in the galactic disk are much younger; some of the brightest among our near neighbors are only a few hundred million years old. The sun itself was born no more than five billion years



CORE OF RED GIANT, whose envelope is diagrammed on the opposite page, is only five times the radius of the earth. All the star's energy is produced in a thin shell of hydrogen, barely 2,000 miles thick, whose mass is only about 300 times that of the earth. The maximum temperature in the inert core of helium is not reached at the center. This comes about because neutrinos, produced copiously by a plasma process near the center, escape through the giant's envelope almost as if it were transparent, removing energy at an enormous rate. The deficit is made good by energy conducted inward by "degenerate" electrons. The helium core finally becomes hot enough to ignite a thermonuclear reaction, producing the helium flash that ends the star's days as a red giant.

ago. It seems safe to say that in the galactic disk the overwhelming majority of stars are no more than about 10 billion years old and that they have been forming continuously throughout this period from clouds of gas comparatively rich in elements heavier than helium.

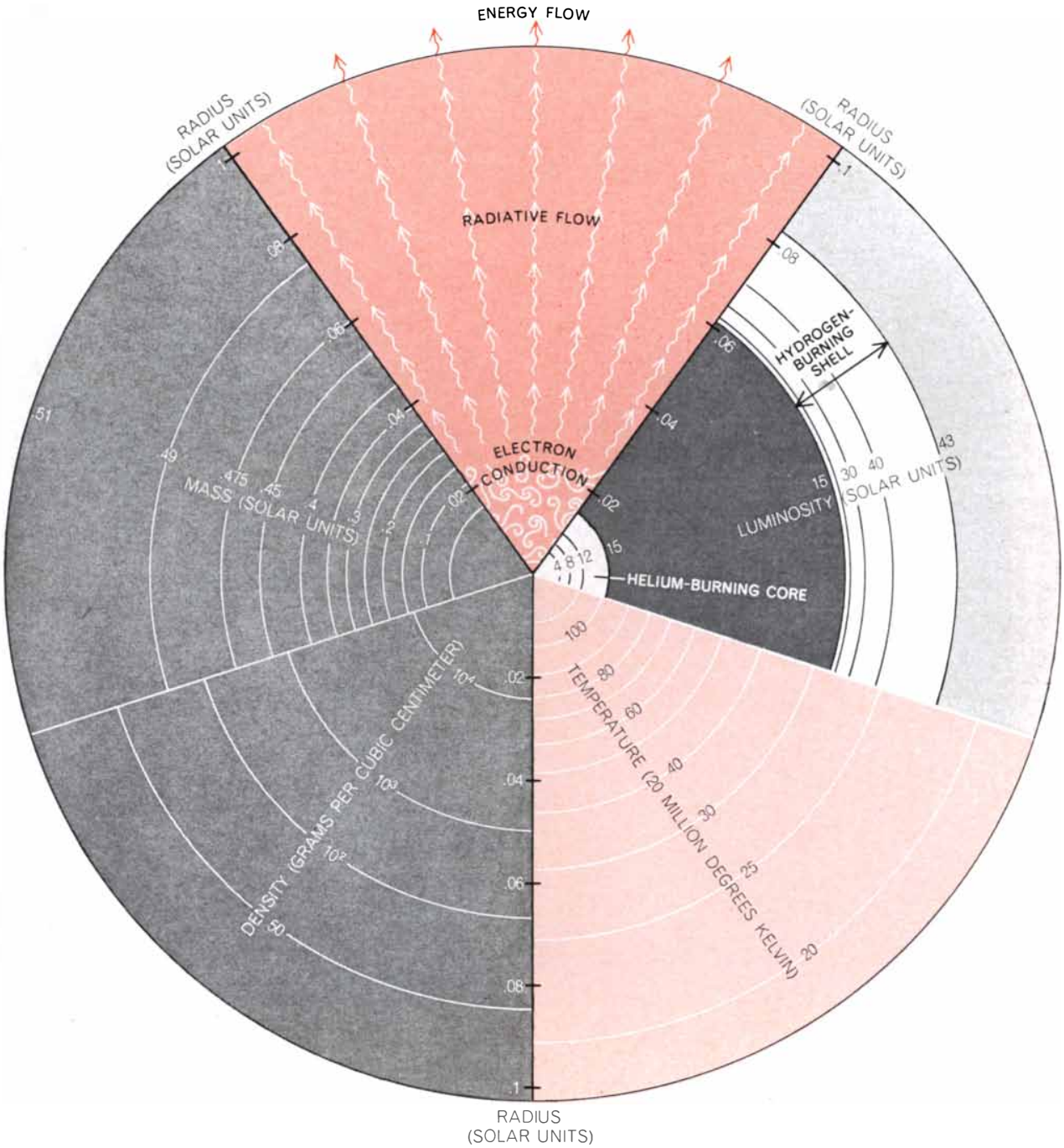
Hence there appears to be a sharp break in the history of our galaxy that coincides with copious production of heavy elements within a fairly short period of time. The concentration of metal-rich stars near the galactic center and in the galactic disk suggests that most of the production of heavy elements may have occurred in these regions *after* the gaseous spheres that became the globular clusters were already formed. An attractive hypothesis is that the galaxy collapsed from an original cloud of gas composed almost entirely of hydrogen and helium. The masses that became globular clusters separated out before the final stages of collapse, between 12 and 13 billion years ago. Soon thereafter densities near the center of the collapsing protogalaxy became favorable for star formation. Presumably some of the star masses that coalesced were so large they consumed fuel at an extremely high rate



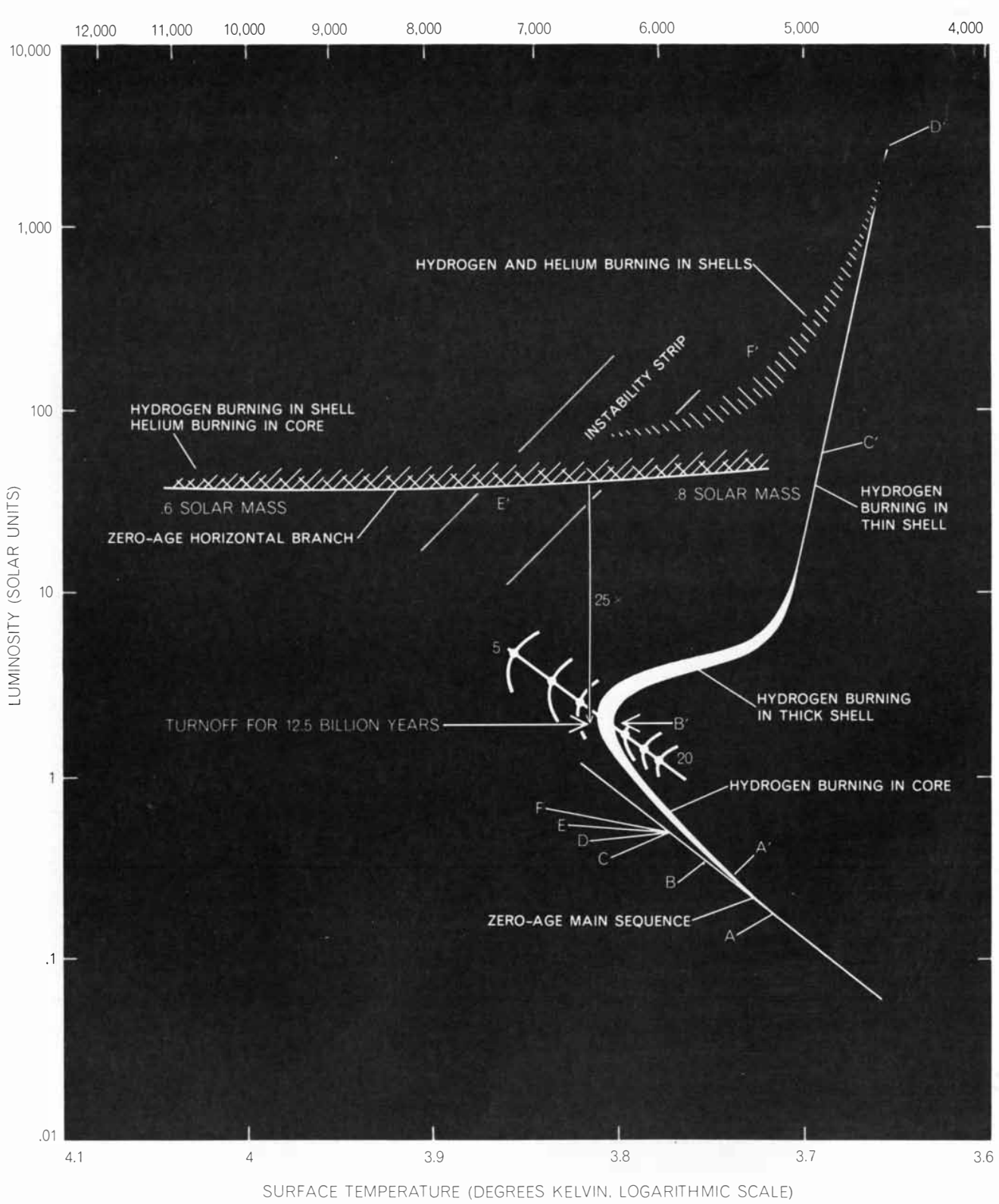
"HORIZONTAL BRANCH" STAR, once a red giant, has shrunk back to four solar radii, assuming a model whose mass is .625 times that of the sun. It is 43 times brighter than the sun and its surface is some 1,500 degrees hotter. The star's energy is produced in a volume represented by the central dot. This tiny volume encloses a central core of burning helium and a shell of burning hydrogen (*see illustration on opposite page*). **Helium fusion contributes about a third, and hydrogen fusion the balance, of the star's energy. In the outer mantle hydrogen and helium are only partially ionized. When this zone becomes deep enough, it induces the periodic fluctuations in brightness characteristic of stars called RR Lyrae variables, which populate "instability strip" in the H-R diagram.**

and thus burned themselves out in less than 100 million years. Such stars would have been hot enough to form elements heavier than carbon and oxygen. Some of them undoubtedly exploded, showering heavy elements out into space. The great majority of stars we see today in the disk of our galaxy were formed after these explosive events and so were able to incorporate sizable amounts (a few percent) of the elements heavier than helium.

According to this hypothesis, the subcondensations of gas remaining outside the galactic disk in the final stages of collapse remained isolated, each contracting under its own weight. At the center of some of these "protoglobular" clusters conditions may have become favorable for star formation even earlier than at the center of the galaxy. In others star formation may have begun later, in which case clusters might vary in age by as much as five billion years. We shall have to wait for more observations and further refinements in theory to decide whether or not this is so. For the present, however, it seems reasonable to assume that globular clusters are all nearly the same age and that this age—12.5 ± 3 bil-



CORE OF HORIZONTAL-BRANCH STAR is the core of the star of .625 solar mass depicted on the opposite page. More than 80 percent of the star's total mass is contained within a volume whose radius is only a tenth that of the sun. In this tiny core helium is converted to carbon by nuclear fusion, and some of the carbon nuclei capture helium nuclei to form oxygen nuclei. The hydrogen-burning shell is much thicker but less energetic than it was in the red-giant stage. Helium is exhausted about 100 times more rapidly than hydrogen was exhausted when the star was starting life on the main sequence. Thus a star spends only about a hundredth as much time (the length of time is typically 100 million years) on the horizontal branch of the H-R diagram as it does on the main sequence.



CLUSTER LOCUS OF MODEL STARS looks like this at the end of 12.5 billion years. The stars that are spread along the locus from *A′* to *D′* are all of the same age and initial composition but they differ in mass. Stars at *A′* began at point *A* 12.5 billion years earlier; stars at *B′* began at *B*, and so on. The turnoff points for theoretical clusters ranging in age from five billion to 20 billion years are also shown by dots placed at intervals of 2.5 billion years. The 12.5-billion-year age appears to be the most probable age for globular-cluster stars because model horizontal-branch stars are then brighter than model stars at the turnoff point by the factor actually observed in real clusters: a factor of roughly 25. The horizontal branch in the diagram is constructed from tracks for models with masses between .6 and .785 solar mass that are burning helium at the center. The masses chosen for these models are arbitrarily lower than the masses of red-giant models in order to compensate for the evidence that stars in the red-giant stage throw off significant amounts of mass. Thus stars lying between *D′* and *F′* originated as more massive stars between *D* and *F* on the main sequence.

lion years—marks the time when our galaxy collapsed to its present dimensions.

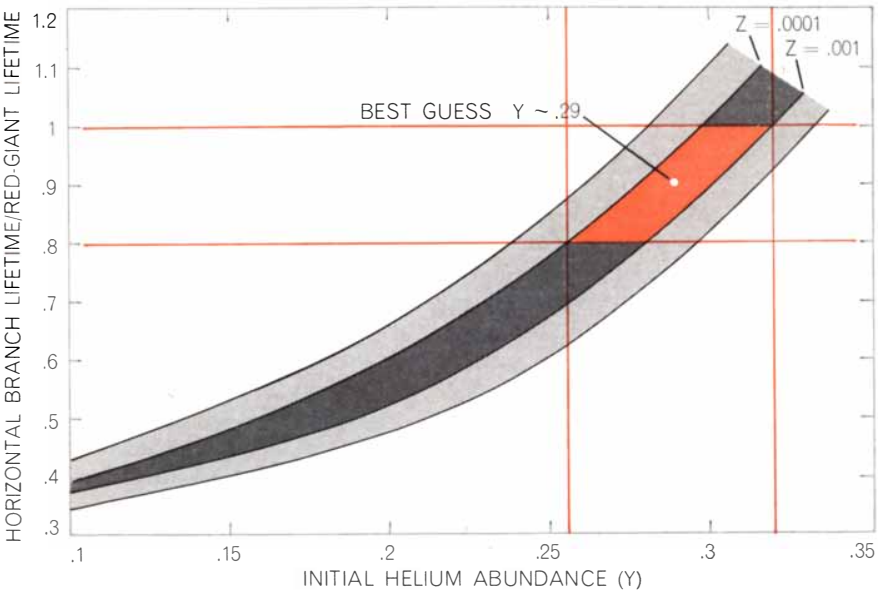
How does this interact with estimates of the age of the universe based on the observed rate of recession of galaxies? If we interpret the observed recession rate according to the simplest big-bang model, the universe began with the explosion of a primordial atom between seven and 13 billion years ago, the precise time depending on the average density one assigns to matter in the universe now. Current knowledge of the average density does not exclude either end of the age range. The lower the present average density, or the more "open" the universe, the closer the big-bang estimate approaches 13 billion years. In view of our estimate of cluster age, it seems that the universe may be very open indeed and that our galaxy and the cluster stars within it may have been formed very shortly after the "beginning."

The fact that galaxies are receding from one another does not by itself constitute proof that we are living in a big-bang universe; the fact of recession can also be accounted for by steady-state models, which assume that the average density of matter in the universe is maintained constant by the continuous creation of matter. An additional fact, however, would seem to make the big-bang model the more appropriate description. This fact is the abundance of helium in the primitive gas, when computed for models of the primordial fireball. The figure obtained agrees very closely with the values disclosed by our studies of the primitive gas from which the oldest stars were formed.

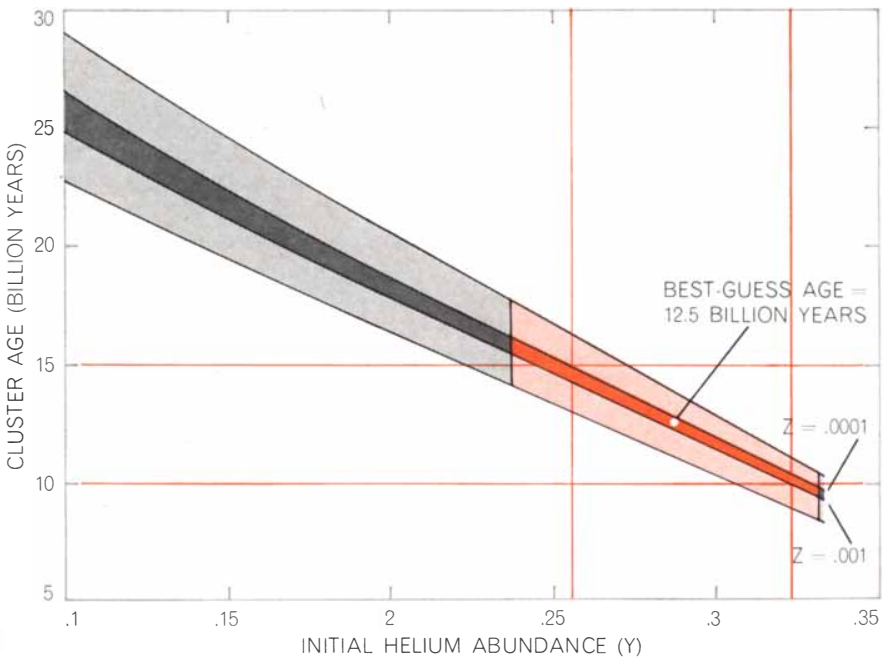
In big-bang models temperatures in the primordial fireball are at one stage so high (above 10 billion degrees) that matter is almost entirely dissociated into protons, neutrons and electrons. As the fireball rapidly expands and cools, protons and neutrons can begin to fuse into deuterium nuclei. Later deuterium nuclei can fuse into helium nuclei and nuclei of still heavier elements. For the simplest big-bang models, however, the matter emerging from the fireball contains between 23 and 28 percent helium; virtually all the rest is hydrogen. The fraction of elements heavier than helium is more than 100 times smaller than can be detected in the most metal-deficient globular-cluster stars.

The effective absence of heavy elements in the matter emerging from the fireball phase strengthens our inference from purely galactic evidence that most of the heavy elements were absent in the protogalaxy and were made in massive stars during a phase of galactic collapse.

The abundance of helium in matter emerging from the fireball phase is strikingly close to the initial helium abundance we have estimated for the oldest metal-poor stars. This agreement provides perhaps the strongest evidence yet uncovered for the hypothesis that the matter in the universe was once crammed together at extremely high densities and at temperatures in excess of 10 billion degrees. The fact that under such conditions most of the energy in the universe was in the form of electromagnetic radiation (photons) gives added meaning to the phrase "And God said, Let there be light."



HELIUM CONTENT OF CLUSTER STARS can be derived from model studies by specifying the abundance of elements heavier than helium (labeled Z) and by specifying the ratio between the time models spend on the horizontal branch and the time spent on the red-giant branch. In actual clusters this ratio seems to lie between .8 and 1. The most probable value of Z is between .001 and .0001 (.1 and .01 percent). The best-fitting models show an initial helium abundance (Y) between .26 and .32, with a mean of about .29.



AGE OF CLUSTER STARS can be computed once the initial abundance of helium and heavier elements (Z) is known. The ages of models are adjusted until stars in the horizontal branch of the cluster locus show a luminosity some 25 times higher than models that are just leaving the main sequence. The most probable age is 12.5 ± 3 billion years.

The Multiple Sclerosis Problem

The cause of this disease of the central nervous system is unknown. The variations in its incidence around the world, however, suggest that it results from infection by a virus with a long latent period

by Geoffrey Dean

The written record of multiple sclerosis begins in 1822, when a young English nobleman, Sir Augustus D'Este, noted in his diary unmistakable evidence that he was a victim of some mysterious disease. When he died at the age of 54, D'Este had spent at least 26 years of his life fighting this unseen enemy. In 1838, while D'Este was traveling all over Europe in a vain search for a cure, Sir Robert Carswell, a British medical illustrator, included in his *Pathological Anatomy* a watercolor of a strange-looking spinal cord he had seen during an autopsy. On the otherwise seemingly healthy cord were scattered spots of hardened and discolored tissue. Carswell described the condition as "a peculiar diseased state of the spinal cord accompanied by atrophy of the discoloured portions."

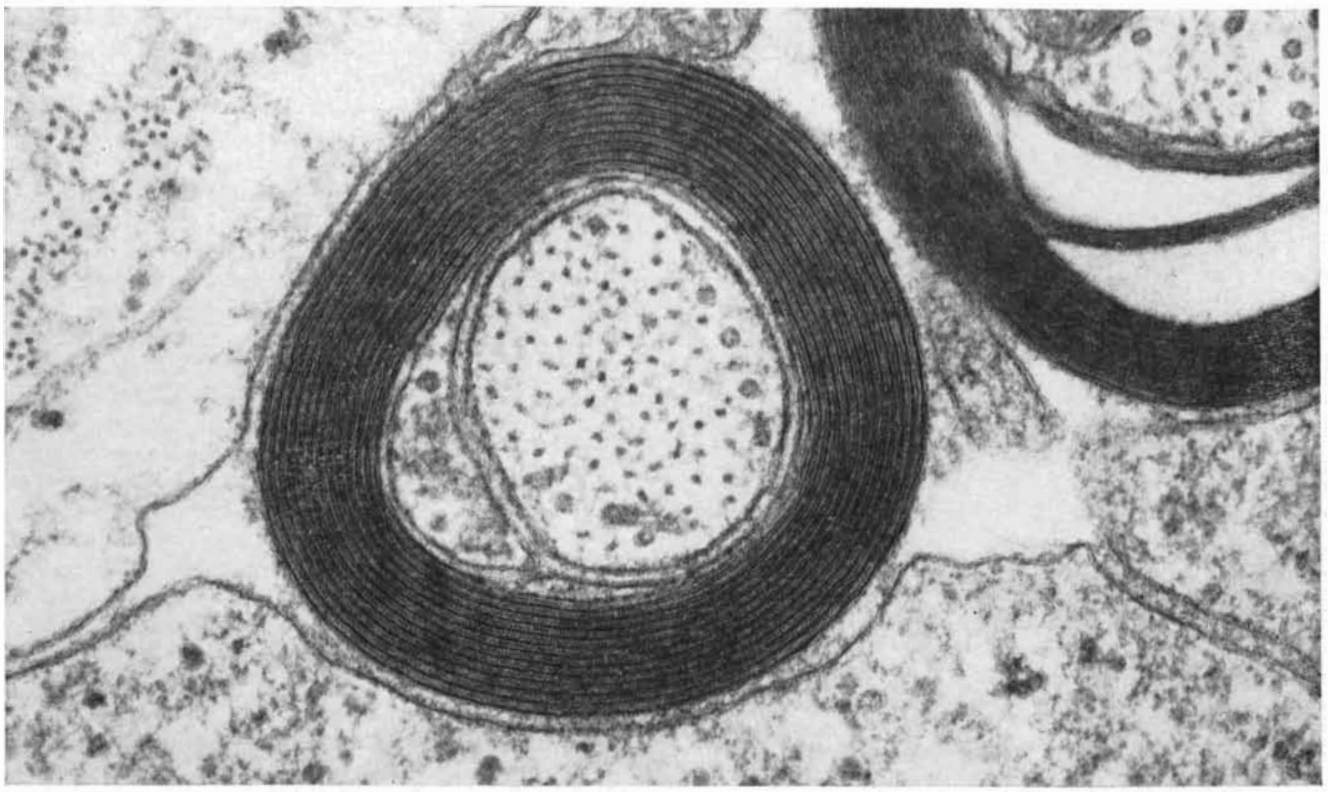
Almost at the same time a French physician, Jean Cruveilhier, observed similar spots on the spinal cord at the autopsy of a woman who had been hospitalized for many years. He called the spots "islands of sclerosis," from the Greek for "hardening," and speculated that they might have actually caused the disease. Thirty more years were to pass before multiple sclerosis was properly identified. It was Jean Martin Charcot, one of France's most famous medical investigators, who in 1868 gave the world a detailed description of the disease. Working at the Salpêtrière hospital in Paris, Charcot found that many of his patients were suffering from varying degrees of tremor and paralysis. Some of these patients suffered from the shaking palsy ("paralysis agitans"), which had been first described by James Parkinson in England in 1817. Charcot realized, however, that another distinct disease was present, evidenced by tremor and varying degrees of spastic paralysis, or stiff and jerky body and limb movements. At autopsy victims of this disease showed plaques, or hardened flat patches, scattered throughout the central nervous system. He called the disease *sclerose en plaques.*

Charcot found that patients frequently experienced alternating exacerbation and remission of symptoms, that the severity of the symptoms varied greatly from patient to patient and that often the illness appeared to remain stationary for many years. In its advanced stages he found that the disease was characterized by paralysis and by three additional symptoms now known as Charcot's triad. One of these symptoms is "intention tremor": the limbs, particularly the arm and hand, shake violently whenever the patient tries to control his movements. The second symptom is slow and "scanning" speech, characterized by a pause after each syllable. The third is ocular abnormalities, particularly nystagmus: an involuntary flicking of the eye back and forth, up and down or around and around. In England the disease is called disseminated sclerosis; in the U.S. it is called multiple sclerosis (sometimes abbreviated MS).

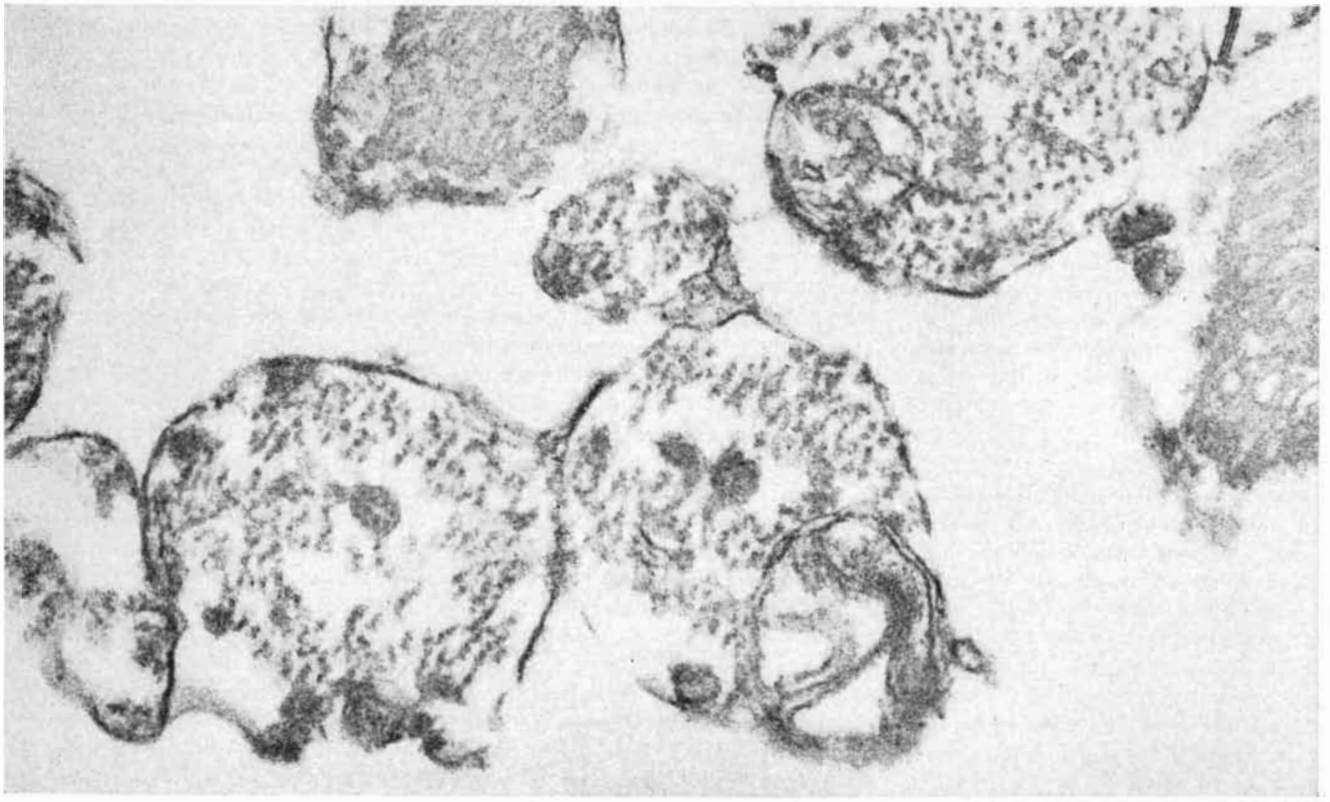
Today multiple sclerosis is the commonest disease of the nervous system affecting men and women in the prime of life in northern Europe and North America. We have learned that the neurological symptoms of the disease are apparently caused by the patchy destruction of the material known as myelin, a system of membranes wrapped in a spiral around the central component of the nerve fiber, the axon. This fatty substance can be compared to the insulation around a telephone wire. When the myelin sheath surrounding a nerve fiber breaks down, the conduction of impulses along the now exposed fiber is disrupted. Eventually scar tissue forms in the demyelinated areas, producing the hard plaques for which the disease is named.

Although its clinical history, its symptoms and signs and the characteristic scarring of the brain and spinal cord it produces have long been well known, the cause of multiple sclerosis has remained a mystery. I remember my old professor Henry Cohen (now Lord Cohen of Birkenhead) once asking a student what was the cause of disseminated sclerosis. The student scratched his head, looked up and down and replied: "Oh, sir, if you had only asked me ten minutes ago, I could have told you." Cohen said: "What a catastrophe. Ten minutes ago you knew and now no one knows."

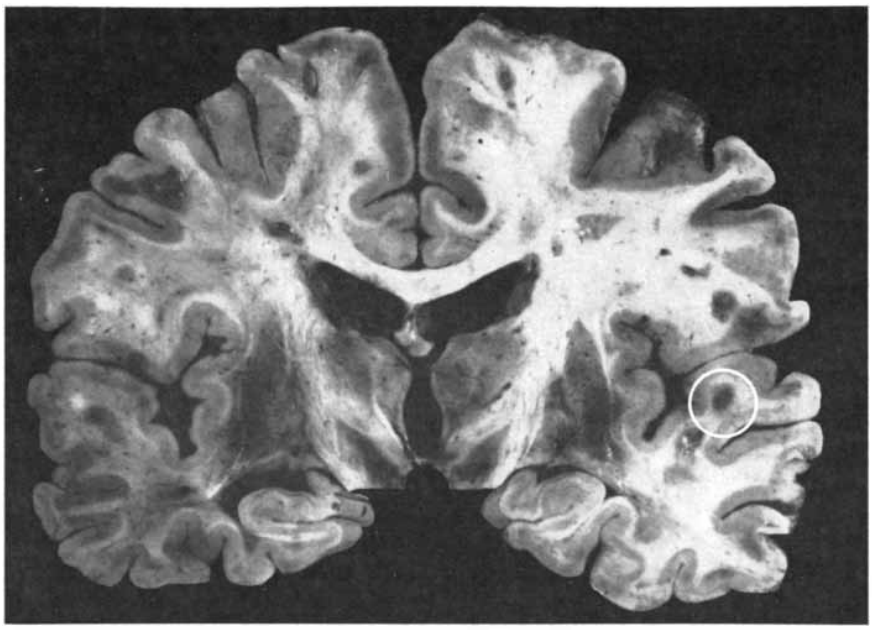
A social phenomenon that has developed in recent decades is the founding of societies to support intensive investigation into the causes of crippling diseases of humanity. Perhaps the best-known of these societies has been the National Foundation for Infantile Paralysis; it is because of the great research work organized and financed by this foundation that the problem of poliomyelitis has been solved. On May 1, 1945, a New York woman named Sylvia Lawry, whose brother had developed multiple sclerosis, inserted a small advertisement in *The New York Times* asking for information from anyone who had recovered from the disease. The response Miss Lawry received came instead from hundreds of multiple sclerosis patients and their families seeking the same information. From this very small beginning grew the National Multiple Sclerosis Society. Today the society has about 200 chapters throughout the U.S. and has helped to found the International Federation of Multiple Sclerosis Societies, which today has 18 member



SHEATHS OF MYELIN, single membranes that are wrapped in a tight protective spiral around the central axons of normal nerve fibers, are shown in cross section at a magnification of 105,000 diameters in this electron micrograph. Myelin, a fatty substance, is analogous in function to the insulation on wire; in its absence the conduction of nerve impulses along the axon is disrupted.



DENUDED NERVE FIBERS, lacking the protection of a myelin sheath, are shown in cross section in this electron micrograph at a magnification of 33,000 diameters. The fibers lie within one of the sclerotic, or hardened, patches whose abundance in the nervous tissue of sufferers from the disease is responsible for the term "multiple sclerosis." Like the micrograph at top of page, this one was made in the Department of Pathology at the Albert Einstein College of Medicine for a study of the neuropathology of the disease.



BRAIN CROSS SECTION shows the sclerotic patches, called "islands" and "plaques" by early observers of multiple sclerosis, that lie scattered throughout the subcortical white matter. For purposes of identification one of the plaques (*right*) has been circled in white.

societies in 17 nations around the world. Miss Lawry was also extremely active in convincing Congress to set up (in 1951) the National Institute of Neurological Diseases and Stroke as one of the National Institutes of Health. The society and the institute are today the chief supporters of multiple sclerosis research in the U.S. and elsewhere.

Among the great advances in medical science since World War II has been the growth of a new discipline: geographical medicine. This field of investigation involves intensive study of populations, particularly of populations that have migrated from one environment to another. In this way one may discover what environmental factors are responsible for various diseases. Much of the rather skimpy body of hard fact so far known about multiple sclerosis has been gathered in just this manner.

For example, in 1953 Leonard T. Kurland, who now works at the Mayo Clinic, compared cases of multiple sclerosis in Halifax, Nova Scotia, with cases in New Orleans and found that the disease was three times more prevalent in the northern city than in the southern one. Kurland also found that multiple sclerosis was more prevalent among the whites of each city than among the blacks (most of Nova Scotia's 14,000 blacks live in Halifax). A similar north-south gradient has been found in Europe, where multiple sclerosis appears to be three times commoner in the countries of northern Europe than it is in areas along the Mediterranean [*see illustration on pages 44 and 45*]. For a long time climate was thought to be the factor responsible for such differences. This hypothesis, however, seems to be contradicted by surveys of countries such as Japan; multiple sclerosis is equally uncommon in both the northernmost and the southernmost islands of that nation. Moreover, there are areas in northern Europe, such as the Faeroe Islands and the coast of Norway, where the disease is relatively uncommon.

In 1947 I emigrated from England to South Africa, where I was surprised to find a marked difference in the prevalence of certain diseases among white South Africans compared with my experience in England. Multiple sclerosis, for instance, was reported to be extremely uncommon; many neurologists who had been trained in Europe and knew multiple sclerosis well stated that they had not seen a single South-African-born white with the disease. Shortly after my arrival in the country I therefore undertook to study all the patients who had been diagnosed as having multiple sclerosis at the five main teaching hospitals in South Africa during the preceding 10 years. There were, I found, 27 probable cases of multiple sclerosis. Of these 14 were immigrants from the United Kingdom or northern Europe, although such immigrants comprised less than 10 percent of the total white population. Clearly multiple sclerosis was extremely uncommon among white persons born in South Africa.

Since 1956, with support from the U.S. National Multiple Sclerosis Society, I have undertaken an intensive study of the annual incidence, prevalence and mortality of multiple sclerosis in South Africa. This has involved the cooperation of many physicians, lay organizations, the press and the radio. The South African Multiple Sclerosis Society was formed in the course of this work and helped greatly with the research. I found that among immigrants from the United Kingdom and northern Europe multiple sclerosis is common, with an incidence of about 50 cases per 100,000 of population. This incidence is about the same order of magnitude recorded for northern Europe. Among immigrants from southern Europe only a third as many contract the disease. Among the English-speaking South-African-born whites only a fourth as many contract it, and among Afrikaans-speaking South-African-born whites only an eleventh as many do so. Multiple sclerosis does occur but is extremely uncommon among the Cape Colored and among the Asians of Natal (chiefly of Hindu origin). Not a single patient with multiple sclerosis has been found among South Africa's 12 million Bantu, although in the major cities of South Africa there are good Bantu hospitals with neurologists who are well able to diagnose the disease.

The elevenfold difference in the prevalence of multiple sclerosis between the immigrants from the United Kingdom and northern Europe and the Afrikaans-speaking South-African-born whites, who are also of northern European stock, shows that multiple sclerosis must be a disease of the environment. Its incidence within families is far too low for multiple sclerosis to be an inherited disease. Two members of the same family are occasionally affected; the number is a little more than would be expected by chance alone. The reason may be that the two have been subjected to the same home environment or perhaps that certain families are more predisposed to the disorder than others.

In 1962 Milton Alter of the University of Minnesota Medical School reported that Jews who immigrated to Israel from northern Europe also showed a high incidence of multiple sclerosis. Jews who were born in Israel, however, show only a sixth the incidence of the European immigrants, and Jews who came from the ghettos of North Africa run an even smaller risk of multiple sclerosis than the Jews born in Israel. The situation in

Australia, which is very similar to South Africa in climate, is quite different. The prevalence of multiple sclerosis is much the same among native-born white Australians as among immigrants to Australia from northern Europe; immigrants to Australia from southern Europe show a lower prevalence than either. In Britain and America those who are economically better off appear to run a slightly greater risk of developing multiple sclerosis than the poor. One more curious fact is that recent research by Alter and myself indicates that in both South Africa and Israel immigrants who come to those countries from northern Europe before the age of 15 run a much smaller risk of developing multiple sclerosis than those who arrive after the age of 15.

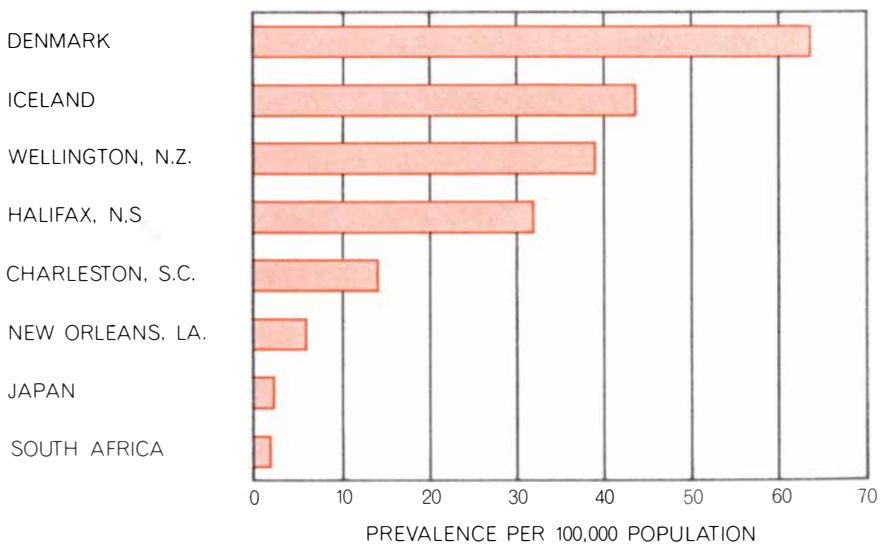
What hypothesis best fits these facts? The earliest one was that multiple sclerosis could be related to climate, particularly to the amount of sunshine. With growing knowledge of the geographical distribution of the disease this hypothesis becomes increasingly unlikely. John S. Barlow of the Massachusetts General Hospital has suggested that the geographical distribution of multiple sclerosis better fits geomagnetic latitude than geographic latitude; this too does not offer a really satisfactory answer to the problem. Some scholars have long considered the possible multiple sclerosis agent to be diet. R. L. Swank of the University of Oregon Medical School has suggested that the disease may be caused by a diet that is rich in animal fat; this is the kind of diet commonest in northern Europe and North America, both economically advanced regions. The hypothesis cannot be correct, however, because the South-African-born whites, particularly the Afrikaans-speaking ones, have a diet that is extremely high in animal fat. Although they have a high death rate from coronary thrombosis, these South Africans run a very low risk of developing multiple sclerosis.

There is a most interesting parallel between the epidemiology of multiple sclerosis and that of poliomyelitis. This was first pointed out by David C. Poskanzer of the Massachusetts General Hospital in 1963. Like multiple sclerosis, poliomyelitis in its paralytic form was a disease of the more advanced nations rather than of the less advanced ones, and of economically better-off people rather than of the poor. It occurred in northern Europe and North America much more frequently than in southern Europe or the countries of Africa, Asia or South America. Immigrants to South Africa from northern Europe ran twice the risk of contracting paralytic poliomyelitis that South-African-born whites ran, and the South-African-born whites ran a much greater risk than nonwhites. Among the Bantu of South Africa paralytic poliomyelitis was rarely an adult disease. During World War II in North Africa cases of paralytic poliomyelitis were commoner among officers in the British and American forces than among men in the other ranks. At the time various wild hypotheses for the difference were proposed; it was even suggested that it arose from the fact that the officers drank whisky whereas men in the other ranks drank beer!

We now understand very well the reason for the strange distribution of paralytic poliomyelitis. Until this century poliomyelitis was a universal infection of infancy and infants hardly ever suffered paralysis from it. The fact that they were occasionally so affected is what gave the disease the name "infantile paralysis." With the improvement of hygiene in the advanced countries of the world more and more people missed infection in early childhood and contracted the disease for the first time at a later age, when the risk that the infection will cause paralysis is much greater.

This explains why the first epidemics of poliomyelitis did not occur until this century and then only in the economically advanced countries. In South Africa examination of Bantu blood samples shows that the population is consistently infected with poliomyelitis virus in infancy. Most of the white people who grow up in South Africa, particularly the Afrikaans-speaking whites who as infants are usually cared for by Bantu or Cape Colored servants, are infected by the poliomyelitis virus early in life and seldom become paralyzed. Among the white immigrants, however, a larger number had missed this early infection and were infected for the first time in South Africa. Today we do what nature once did for all of us and deliberately infect our infants with a variety of living poliomyelitis viruses: the Sabin vaccine.

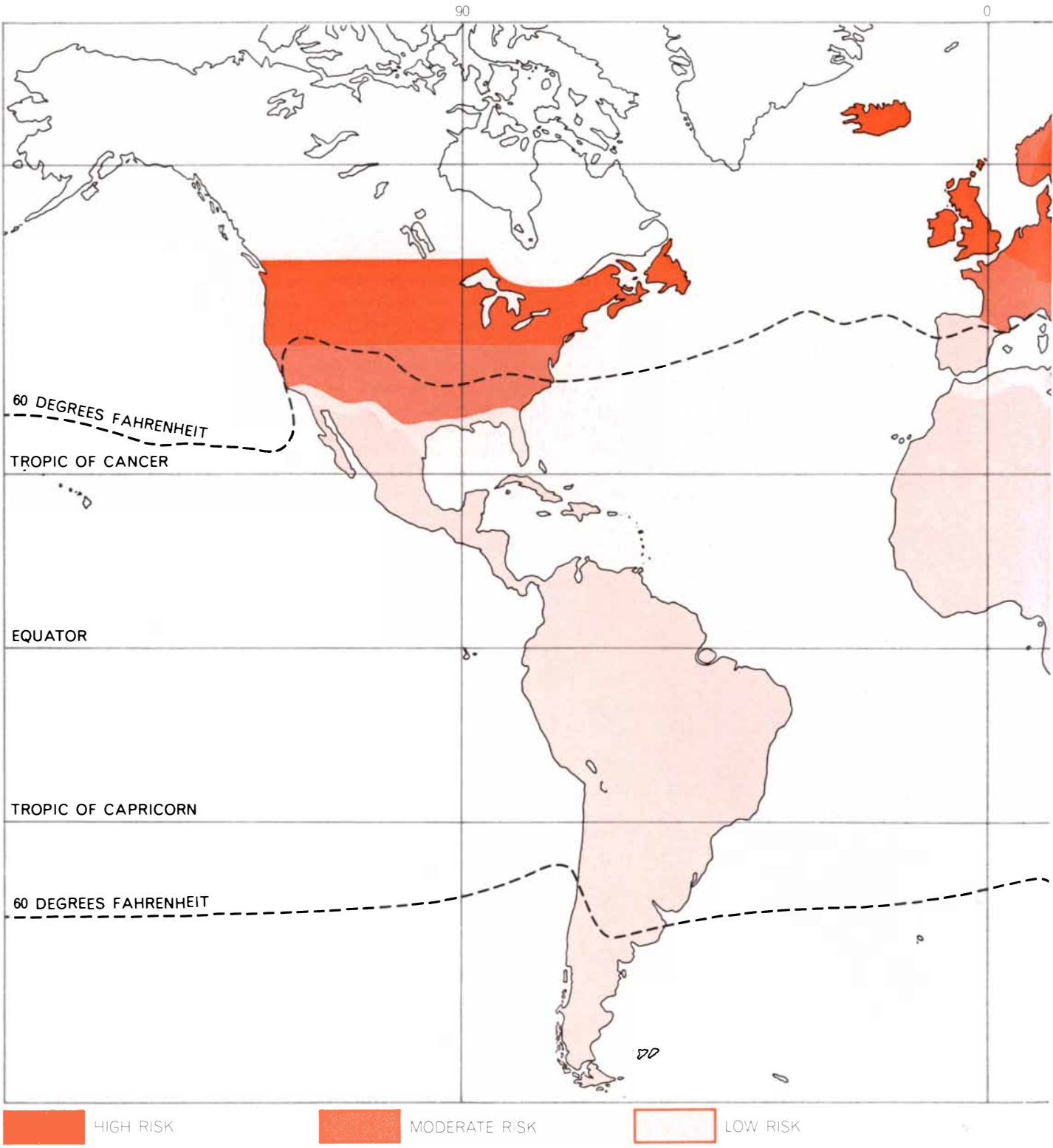
Let us compare prevaccine poliomyelitis with multiple sclerosis. Multiple sclerosis is a common disease of the nervous system in northern Europe and the northern part of the U.S. and is rather rare in the less economically advanced southern regions. Like paralytic poliomyelitis, multiple sclerosis is extremely uncommon in Japan, where until recently human excrement was used to fertilize crops. It is also uncommon in South America. In South Africa the native-born white children, particularly in Afrikaans-speaking families, are cared for by nonwhite servants whose home hygiene is primitive. The children are therefore likely to develop many infections early in childhood. White immigrants from Europe who come before the age of 15 would likewise be infected early, develop natural resistance and so be protected from a later infection. Conversely, in



PREVALENCE OF MULTIPLE SCLEROSIS is far from uniform. The graph records the number of victims of the disease per year per 100,000 population, but the data are not all for the same year. In the case of South Africa the figure applies only to native-born white residents and excludes white immigrants. An apparent connection between greater and lesser prevalence of the disease and the coolness or warmness of the climate has been disproved.

Australia the primitive nonwhite population has very limited contact with the white settlers, and the entire white population has a high level of domestic hygiene. In that country there is no great difference in the prevalence of multiple sclerosis between white immigrants from northern Europe and Australian-born whites. Within Australia, multiple sclerosis is commoner in the temperate state of Tasmania, far to the south, than in the more tropical northern states. It is even commoner in New Zealand.

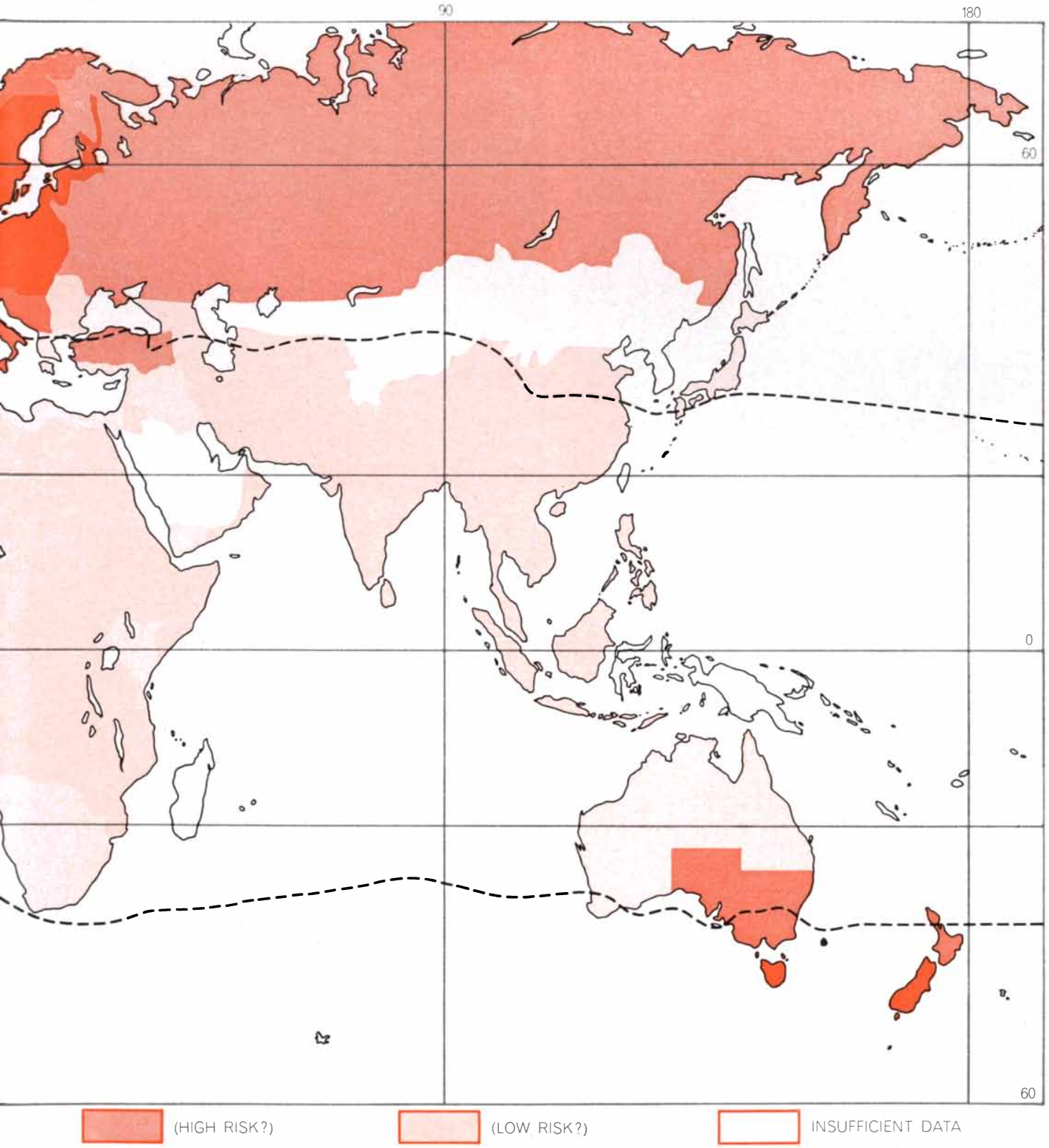
On the basis of the data from Israel and South Africa, John Kurtzke of the Georgetown University Medical School has suggested as an alternative hypothesis that multiple sclerosis is "caught" at about age 15 and that the risk of contracting the disease will depend on where one happens to be on attaining



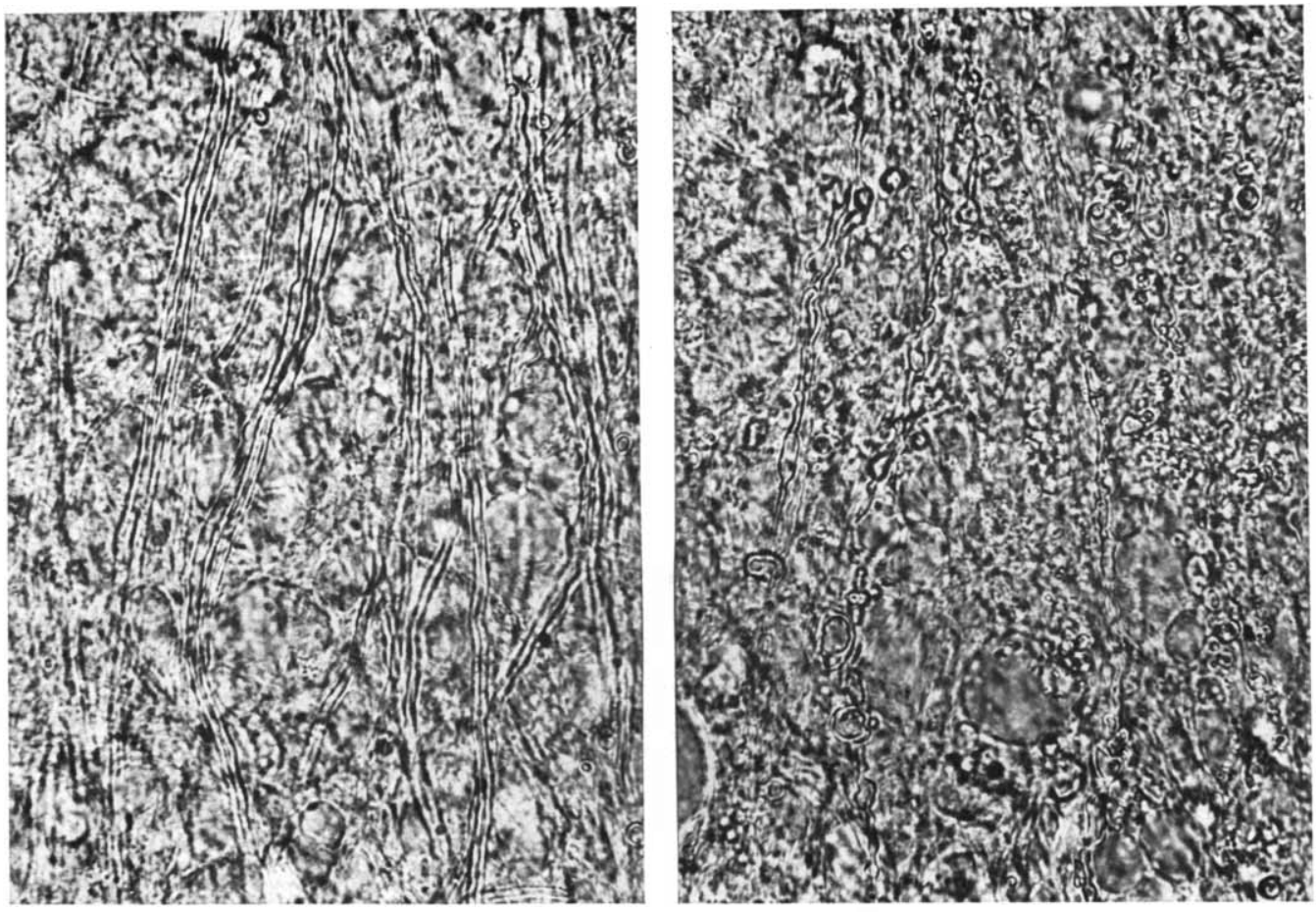
WORLDWIDE DISTRIBUTION of multiple sclerosis is plotted on this map by distinguishing between populations whose risk of contracting the disease is known to be high to moderate (*darker colors*) and those whose risk is known to be low (*lightest color*). For some regions, such as central Asia, Alaska and parts of Canada, the data are too few to allow this distinction. For others the as-

that age. If one is in a high-prevalence area at age 15, the risk is high; if one is in a low-prevalence area, the risk is low. Thus immigrants who come to South Africa from Europe after age 15 carry their high risk with them, and those who arrive before that age have the advantage of residence in a low-risk region. Whether my belief (that multiple sclerosis is normally an infection of early childhood) or Kurtzke's hypothesis (that it is caught wherever one happens to be at the age of 15) is correct, we both agree that multiple sclerosis is almost certainly a virus infection of childhood, and that the virus responsible is probably one of the little-understood "slow," or latent, viruses.

It is now known that certain unusual viruses can remain latent in the nervous system for many years without pro-



signment to either extreme is tentative. A comparison of the prevalence of the disease among the native-born and among immigrants to such nations as South Africa, Israel and Australia in the years following World War II suggests strong parallels between multiple sclerosis and infantile paralysis; both diseases are most commonly found among populations with high standards of domestic hygiene.



LABORATORY PROOF that a demyelinating factor is present in the blood of multiple sclerosis victims is seen in these two micrographs from an experiment by Murray B. Bornstein of the Albert Einstein College of Medicine. At left is a living culture of rat brain tissue; the numerous strands running generally from top to bottom of the micrograph are normal nerve axons, sheathed in myelin. The culture is seen at right in a micrograph made eight hours after serum from a patient undergoing an acute exacerbation of the disease was added to the nutrient medium. Most of the sheaths have been fragmented and the myelin has been reduced to droplets.

ducing symptoms. A latent virus is suspected as the causative agent of scrapie, a nervous disorder found in sheep. Another example is the agent responsible for "kuru," a disease that once brought death due to generalized paralysis to about 10 percent of the members of the Fore tribe in New Guinea. Originally thought to be the result of genetic factors, kuru has now been transmitted from human patients to chimpanzees and has subsequently been passed on from one chimpanzee to another. The incubation period is initially very long but shortens somewhat with repeated passage, a pattern that is typical of virus infections. It thus appears that kuru is caused by a virus that can remain latent in the nervous system for a long time. In New Guinea the disease was probably spread by the tribal ritual of "cerebrocannibalism," or the eating of human brains, in a situation where the brains were already infected with the kuru virus.

The virus responsible for kuru and other newly discovered latent viruses are strange forms of life indeed. They are difficult to study in our present state of knowledge; we do not yet know how to culture them nor have they been seen even with the electron microscope. They do not cause the biochemical reaction known as complement fixation, which normally assists in the detection of viruses, and what is even more remarkable, their infectivity is affected neither by boiling nor by lengthy immersion in Formalin.

Today, a century after Charcot recognized multiple sclerosis as a distinct medical entity, it remains a disease with an unknown cause, an unpredictable course, an undiscovered cure and not even a simple laboratory test to confirm its diagnosis. So far the best clue to the etiology of the disease has come from the study of multiple sclerosis among populations in different parts of the world. The hypothesis that best fits the few facts is that multiple sclerosis is normally a virus infection of childhood and that it has a long latent period. If the early infection is missed and occurs for the first time in early adult life, there may be a precarious balance between allergy and immunity to the virus that produces the alternating exacerbations and remissions that are characteristic of the disease. To investigate this hypothesis the National Multiple Sclerosis Society has channeled a significant proportion of its resources into improving virus technology and into the search for a multiple sclerosis virus.

The multiple sclerosis problem is going to be solved. All the evidence points to its being a disease of the environment. The most likely environmental factor is infection by one of the latent viruses, infective agents that can remain relatively quiescent in the central nervous system for many years. When the search for the virus eventually succeeds and leads to an effective treatment or preventive for this disease, the key to success will have been provided by geographical medicine.

Two small, captive chapters of the round earth society

(no caps, please)

Kodak

MACRO *Our company has become an important repository of skills in handling visual imagery on a massive scale. Within the extant structure of society, we seek ways to join these skills of ours most effectively with organized endeavors of others for rational assessment, sharing, utilization, and conservation of the earth's finite resources. Our point of initial contact with such organized efforts may be addressed as "Advanced Planning," Eastman Kodak Company, Rochester, N. Y. 14650. Contact is earnestly invited.*

Role seeking, role development, role guarding—all as basic as breathing, eating, sleeping. Do it without shame, but don't admit you're doing it. Roles should be honorable and important.

Nothing today is more honorable or important than coping with the new-found finiteness of global surface. Since Columbus, rumors have persisted that space to explore and exploit is theoretically less than infinite. Quite recently brave, camera-carrying tourists on their way to and from the moon proved the theory with color pictures of the whole earth.

Unbelievers grow scarce. A major stocktaking must be *organized*. Is there going to be enough fresh water? What temperature? How about the oil? Timber for housing? Ore? Must the mines come that close to the clam beds? Will there be enough *food?* What kind of food? What do the savants think—the geologists, the hydrologists, the agronomists, the climatologists, the forestry people, the fisheries people, the ecologists?

Introspection having gone out of style among savants, they ask for data, synoptic data—pictures. An observer on foot or in a surface vehicle gets only a traverse. That is hardly enough, nor is a single spectral band. Correlation of signals from several bands makes up the signature of a condition, sensed from a distance. Remote sensing from aircraft creates desire for the more synoptic view from satellites. Impressive as is a single frame of film bearing the entire earth, it permits relatively few judgements.

Thus roles shape up for the space technologists. Even if the earlier on-board sensors are all-electronic, down on the ground the message needs to be put into graphic—photographic—form for eye-to-brain judgement. If the resolution is fine enough to be useful, much photographic material will be required, and in colors for the additional channels of information to the brain that colors provide. There will have to be many copies for distribution to the various disciplines. Those of us who measure success in acreage of film made, exposed, and processed in machinery of our manufacture will experience the joy of role fulfillment.

Certain full-time innovators among us who made their mark with the Lunar Orbiter now seek role extension by advocating photography as the original sensor. They argue that this is the only proven way yet to get enough resolution for significant data from orbital altitudes. And with the awesome volume of data that would then need sifting, a role is in prospect for still others of us who dream of machines that will scrutinize film at great speed and with inhuman immunity to boredom.

MICRO *People of our Laboratory of Industrial Medicine. A minor recent accomplishment of theirs, quite incidental to their tasks, has been the discovery that Methyl 5,12-Diethylfluorindinium Methosulfate, when fed to housefly larvae, is retained as a blue abdominal stain into the adult stage. This may prove useful for studies of the dispersal patterns in agricultural pests. The dye has been made available through lab supply houses as* EASTMAN *Organic Chemical No. 11258. When the adult is crushed on filter paper, the stained viscera color the paper blue.*

The tasks are to probe into the biological consequences of what we do, make, or plan. The laboratory was set up for this purpose in 1936, a year when business in general was in poor condition to stand assault on its motivations. The chemist-physician who has run it since 1949 has built it to its present strength of 50.

His is not a highly competitive personality. He does not boast of more zeal than counterparts in certain other famous companies. They can speak for themselves. For all we know, other companies also illuminate with a xenon arc the bottles of effluvia in which laboratory fish live. Not just the discharge but the products of sunlight shining on it should be proved innocent.

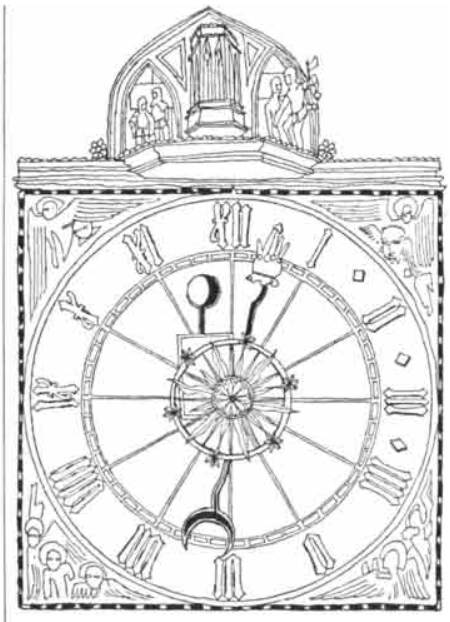
Certainly we are not the only chemical manufacturer who employs young ladies to count and measure bones in the skeletons of newborn rodents. Charge it to corporate curiosity about the fetus-deforming potential of a compound which might, for example, be proposed some day as a modifier of a plastic which could find its way into the construction of food-handling equipment.

Though this level of curiosity about the ultimate effects of micro-additions to the biosphere deserves all the attention it gets, it was to a deeper level that our Laboratory of Industrial Medicine further ventured some years ago. If that puts us into molecular biology, so be it. Could be helpful in the commercial interests of product development to learn what chemical doors to shun.

No point wasting development money in directions that are bad for the ecosystem and even worse for us ourselves who, unwarned and unguarded, might take into our personal bodies larger doses of intolerables for a mere paycheck.

Economically feasible way to publish in full color

Send us a set of slides (or other forms of color imagery), send us your text, and we can put it all on color microfiche in as many copies as needed. We have gone so far as to station all around the country persons well informed on the details of "micropublication," including viewing equipment. They can be reached through the Yellow Pages. Look under "Microfilming" for "Eastman Kodak Company, Business Systems."



The Ecology of War

War is always destructive of the environment, but in Vietnam war is being waged directly against it. The tactics of chemical defoliation and crop destruction are having a profound effect on the country's ecology, according to two biologists who visited Vietnam last year. In their report, published in *Science*, Gordon H. Orians of the University of Washington and E. W. Pfeiffer of the University of Montana urge the American Association for the Advancement of Science to help set up an international research program on the long-term effects of the military use of herbicides.

Orians and Pfeiffer interviewed military personnel, flew on spraying missions, surveyed defoliated areas from the air and by boat, talked with plantation owners, agricultural experts and scientists and studied records and photographs. The defoliation program, they write, was begun in 1962 and was stepped up sharply in 1966. Defoliation is now considered a potent weapon in guerrilla warfare, and "it is to be expected that in any future wars of this nature more extensive use will be made of it."

In forests, where most spraying operations are conducted in an effort to reduce concealment, a significant fraction of mature trees are killed by a single application; almost complete kills can be expected if spraying is repeated often. Orians and Pfeiffer estimate that from 20 to 25 percent of Vietnam's forests have been sprayed more than once. The mangrove forests characteristic of the river deltas southeast of Saigon are particularly susceptible to defoliation; one application kills most of the trees. Orians and Pfeiffer toured a large mangrove area and found it almost completely barren; the forest may never become completely reestablished, they report. As for upland forests, they cite earlier studies to the effect that two or three spray applications may kill about half of the commercially valuable timber. (The timber can be harvested, but then there is another complication: so much of it is studded with shrapnel that damage to saw blades has become a serious problem for lumber mills.)

Most of the spraying is directed against forest and brushland with two preparations called White and Orange, in which the active chemicals are 2,4-D and 2,4,5-T. Another agent, called Blue, in which the active chemical is cacodylic acid, an arsenic compound, is applied to cropland in mountainous parts of the country generally under the control of the National Liberation Front (Vietcong). U.S. officers consider this "resource denial" program successful because many captured soldiers from the sprayed areas are seriously undernourished. The authors remark that any such food shortage may well affect children, women and old people more than it does soldiers.

Military officials and some independent studies have suggested that herbicides do not often do damage beyond intended target areas. Orians and Pfeiffer disagree. In several instances they observed damage to fruit trees and other crops that could be traced to defoliation attacks or to the jettisoning of chemicals nearby. It was difficult to establish the extent of this damage, they observe, because claims for damages are discouraged by local officials. Orians and Pfeiffer did find that damage to rubber trees, one of South Vietnam's major resources, has been extensive. It has presumably all been accidental, caused by vaporized defoliant that is blown by the wind into plantations from nearby target areas, since permission to spray rubber plantations is said never to be granted. According to the Rubber Research Institute of Vietnam, repeated defoliation threatens the existence of rubber culture in the country.

As for the effect of herbicides on ani-

mal populations, the authors conclude that it may be primarily through the destruction of habitat, although there may also be direct toxic effects. In the defoliated mangrove areas, Orians and Pfeiffer found no local birds that live on insects or berries, and the number of fish-eating birds was smaller than expected. They cite a recent report on the production of birth defects by 2,4-D and 2,4,5-T, and they point out that U.S. manufacturers of these agents and the arsenic herbicide include warnings on their labels regarding possible toxic effects on humans and domestic animals.

Moving beyond defoliation, Orians and Pfeiffer mention other environmental effects: the pockmarking of Vietnam with bomb craters (2.6 million of them in 1968, they calculate), the upsetting of human ecology through destruction of villages and forced urbanization in Saigon, the sharp increase in air pollution and the promotion of forest fires. Not all animals have suffered from the war, however. Tigers seem to have benefited. "In the past 24 years they have learned to associate the sounds of gunfire with the presence of dead and wounded human beings in the vicinity. As a result, tigers rapidly move toward gunfire and apparently consume large numbers of battle casualties."

Making a Gene

The first complete synthesis of a structural gene, an achievement reported last month by H. Gobind Khorana and his associates at the University of Wisconsin, was a remarkable feat of chemical construction. DNA, the genetic material, had been synthesized in the sense of being replicated from natural DNA in the test tube by Arthur Kornberg of the Stanford University School of Medicine in 1968. Last year a complete gene had been isolated from a bacterium by Jonathan R. Beckwith and his group at the Harvard Medical School. The Khorana group at Wisconsin's Institute for Enzyme Research began with knowledge of the sequence of 77 nucleotides, or nucleic acid subunits, that specifies the synthesis in yeast of a certain molecule, alanine transfer RNA. Using off-the-shelf chemicals, they put those subunits together one by one. Their synthesis, in other words, was a process that the living

cell, which makes new genes by copying old ones, does not perform.

DNA is a double chain of nucleotides, each consisting of a backbone of sugar and phosphate and one of four organic bases: adenine, thymine, guanine and cytosine (*A, T, G* and *C*). RNA is made of similar nucleotides comprising *A, G, C* and uridine (*U*), which substitutes for thymine. RNA is synthesized on DNA as a template by base-pairing; each RNA nucleotide pairs with its complementary DNA nucleotide: *A* with *T, U* with *A*, and *G* and *C* with each other. If one knows the sequence of nucleotides in an RNA molecule, therefore, one can deduce the sequence of nucleotides in the gene that made it. In 1964 Robert W. Holley, then at Cornell University, determined the nucleotide sequence of alanine transfer RNA. Knowing that, Khorana knew the sequence of the nucleotides in the gene he sought to construct.

Putting them together was a complex task. The Wisconsin group set about joining nucleotides to one another to form short single strands of DNA. Then they joined other nucleotides to make overlapping segments of the complementary, or opposite, DNA strands. They allowed these complementary segments to bond, as DNA does naturally, by base-pairing. Finally they used the enzyme DNA ligase to tie the double strands together, forming the completed double chain. Throughout, the major problem was to coax the nucleotides to form the proper bonds and to protect against unwanted reactions.

Khorana confirmed the accuracy of his 77-nucleotide sequence by various tests, such as the "nearest neighbor" analysis developed by Kornberg. The next steps will be to get the artificial gene to replicate itself as natural DNA does and then to get it to act as a template for the synthesis of alanine transfer RNA. Meanwhile the group is proceeding with the synthesis of another gene, the gene for tyrosine-suppressor transfer RNA. This is a special form of transfer RNA that corrects for a certain mutation in the bacterium *Escherichia coli,* so that the biological activity of the artificial gene will be subject to a convincing test.

Addiction, Medicine and the Law

Drug addiction can be regarded as a medical problem requiring therapy prescribed by a physician according to his patient's needs or it can be seen as a law-enforcement problem best solved by strict regulation of the availability and use of the drug. These inherently contradictory attitudes have been brought into open conflict by publication of Federal guidelines for the use of the drug methadone in controlling addiction, and by a bill before Congress.

According to the guidelines, which have been drafted by the Bureau of Narcotics and Dangerous Drugs and the Food and Drug Administration, a physician may administer methadone only in a rehabilitation program, and he must keep detailed records of each case that can be inspected at any time. He may treat only those patients who have given "informed and voluntary consent," who are dependent on narcotics and who can prove that they have failed to withdraw from addiction at least once. Usually a patient must take the drug at a rehabilitation facility under medical supervision and give a urine specimen once a week to prove that he is not taking drugs from other sources. A primary purpose of such a program, according to the guidelines, should be to return the addict to a drug-free state at some unspecified future time. The Narcotics Bureau says that the guidelines are intended to prevent abuse of methadone and to ensure that therapy programs yield reliable data.

The bill (and two others like it) before Congress gives the Attorney General the power to make rules affecting the experimental and therapeutic use of a broad range of narcotics and other potentially dangerous drugs (which constitute about a third of all legal prescriptions). The bill covers sedatives and tranquilizers as well as stimulants and psychotomimetics. According to the language of the bill, none of its provisions are to be construed as interfering with the dispensation of these drugs in accordance with "legitimate medical practice." The bill implies, however, that the Attorney General can define such practice and decide whether the practitioner's actions conform to it. The penalties for violating the bill's provisions include fines and imprisonment.

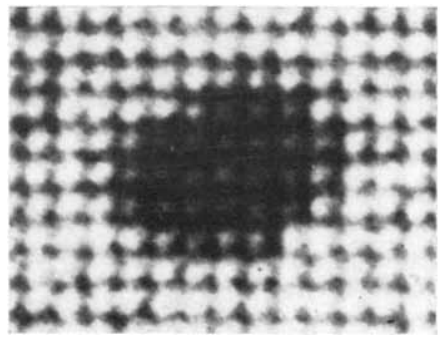
Both the guidelines (which are to be published in the *Federal Register*) and the bill have alarmed some physicians. These regulations, they fear, broaden police powers over narcotics research and therapy. In particular the physicians are concerned that Federal officials instead of clinicians will define therapeutic goals and procedures. There is no assurance, according to Vincent P. Dole of Rockefeller University, that the goals and methods described in the guidelines are appropriate to every patient. Moreover, the requirement that methadone cannot be administered for maintenance and may even be withdrawn if dependency does not lessen could threaten the original methadone-therapy program in New York. This program and similar efforts throughout the country maintain addicts on methadone, which satisfies physical craving but has no narcotic effect, so that they are better able to hold jobs and otherwise function normally.

Other physicians and investigators share Dole's concern that such regulations will discourage research and make effective treatment of addicts difficult. Physicians may also avoid addict patients if the procedures they regard as being necessary could lead to prosecution. Existing drug regulations and professional groups, the concerned physicians and investigators believe, can provide all the policing that is needed.

Oblong Moon

In three of the 126 full-frame photographs of Mars taken by *Mariner 7* last August a tiny black dot can be detected against the face of the planet. The dot is all but indistinguishable from 30 other dots, the tiny reseau spots placed symmetrically on the Vidicon faceplate for calibration purposes. By analyzing one of these specks in great detail, Bradford A. Smith of New Mexico State University has reached some striking conclusions about Phobos, the larger and brighter of Mars's two satellites. (The other is Deimos.) Both were thought to be less than 10 miles in diameter. The 1969 *Mariner* television experiment had been designed to catch a glimpse of at least one of them.

The image of Phobos studied most closely by Smith falls within a rectangular array of picture elements (pixels) measuring eight by 10 pixels (*see illustration below*). Each pixel, as recorded on the face of the Vidicon tube, is about 13 microns on a side. To clean up the image, Smith reports in *Science,* "first-order noise removal, geometric correction, and photometric decalibration [were] applied by hand, pixel by pixel." Next Smith regarded Phobos as a dim



The elongated image of Phobos

see what gypsum cement can do for you.

In this FREE illustrated and informative book you'll find some of the unusual applications of U.S.G. Gypsum Cement. Plus features of all 100+ known formulations...

• Odorless, benign, non-allergenic, non-toxic, non-corrosive, highly compatible, incombustible, economical. • Controllable strength, absorption, and specific gravity. • Outstanding dimensional stability. • Can be carved, cast, pressed or machined. • Adjustable setting times—3 minutes to 20 hours.

UNITED STATES GYPSUM

For the whole U.S.G. Gypsum Cement story, send this coupon in today.

UNITED STATES GYPSUM
Dept. SA-07
101 S. Wacker Drive, Chicago, Ill. 60606

Please send me your new brochure (IGL-155) on Gypsum Cements. I understand no salesman will call.

Name_____ Title_____

Firm_____

Address_____

City_____ State_____ Zip_____

Telephone_____

but not quite opaque object that came between the uniformly lighted surface of Mars and the *Mariner* camera, then about 137,000 kilometers from Mars. In this way he calculated the mean surface brightness of Phobos. Smith's surprising result is that the albedo, or reflectivity, of Phobos is lower than that of any other known body in the solar system. It reflects only 6.5 percent of the light striking it, compared with 11.5 percent for the moon and 15.4 percent for Mars itself. This suggests to Smith that Phobos has a dust-free surface "continuously sputtered clean by meteoritic impacts."

Simple inspection of the magnified image shows that Phobos is not spherical. After careful measurement Smith concludes that it measures 22 kilometers in the plane of its orbit and only 18 kilometers at right angles to the plane. Thus it is slightly larger than had been thought. Evidently, Smith concludes, "Phobos did not form by accretion as it orbited within the planetesimal cloud around primordial Mars, but may have been captured in its present form at some later time."

The Bristlecone Correction

A long-lived California tree apparently holds the answer to a puzzle that has troubled students of ancient history for some years. This is the discrepancy between dates derived from historical evidence and dates derived from carbon-14 measurements: the historical dates are considerably older. The best-known instance is found in Egypt. There the dynastic lists of kings make it possible to date the First Dynasty with confidence around 3000 B.C. Carbon-14 analysis of First Dynasty materials, however, suggests that they are 500 years younger.

The tree is the bristlecone pine, a slow-growing evergreen found on the barren slopes of the White Mountains in central California. A series of overlapping tree-ring counts from various bristlecone specimens extends back in time for more than 7,000 years. The wood in each ring thus provides a sample of organic material with a known date that can be subjected to carbon-14 analysis. This has been done by Hans E. Suess of the University of California at San Diego. When the two series of dates are compared, the carbon-14 underestimation of age increases with the increase in the true age of the specimen.

Reviewing other recent efforts to correlate historical and carbon-14 chronologies, the editors of *Current Archaeology* report that carbon-14 analysis of organic material from Egyptian tombs of various known dates shows discrepancies that agree generally with Suess's findings. They also note that students of medieval archaeology have found a carbon-14 "wobble" that affects dates in the 16th and 17th centuries: carbon-14 analysis of wood from tree rings that were formed in A.D. 1520 and 1640 both yield the same date: 1570. Suess's bristlecone studies show similar wobbles; there are several between 1800 and 2500 B.C. and earlier ones around 3000 and 3600 B.C. The causes of the underestimation and the wobbles are not yet understood.

Rocket Laser

A laser of extremely high power, capable in principle of generating thousands of kilowatts of infrared radiation in a continuous beam, has been developed by workers at the Avco Corporation. The new device, which resembles a rocket motor in operation, is called a gas-dynamic laser. It was invented by Arthur Kantrowitz, Edward T. Gerry, Donald A. Leonard and Jack Wilson of the Avco Everett Research Laboratory. The first working model of a high-power gas-dynamic laser was described recently by Gerry at the annual meeting of the American Physical Society in Washington. This model can produce 30 kilowatts of power in a very narrow beam and 60 kilowatts in a more divergent beam; both values are about triple those of the most powerful continuous-beam lasers in operation.

The gas-dynamic laser operates by burning fuel (a gaseous mixture of carbon dioxide and nitrogen) in one or more combustion chambers. The hot gas is then rapidly cooled by expansion as it passes at supersonic speed through an array of small nozzles. In the process the thermally excited gas molecules undergo a temporary "population inversion," which is the essential prerequisite for laser action. A spontaneously emitted photon of light can then strike a molecule in the overpopulated upper energy state, stimulating the emission of an additional photon, which in turn stimulates the emission of more photons from other excited molecules in the upper state. Mirrors facing inward from opposite ends of the supersonic channel collect and direct this cascade of stimulated coherent emission back and forth across the channel until the full available power is extracted from the gas medium by focusing the beam out through a hole in one of the mirror arrays.

An important feature of the gas-dynamic laser is that waste energy in the form of heat is removed from the device by the high-speed gas flow instead of by diffusion to the outer walls. The power output of most conventional lasers is limited by the amount of waste heat that can be removed by wall cooling.

The projected applications of high-power gas-dynamic lasers include cutting or trimming various materials during a continuous on-line production process; cutting, welding or brazing large metal or ceramic structures at short or long ranges; weakening rock in preparation for mining or tunneling operations; transmitting power to otherwise inaccessible locations (for example, to charge the batteries of an artificial earth satellite), and serving as the medium for a high-bit-rate, long-range optical communications system.

Marathon Mounds

The plain of Marathon, where a Greek force attacked and routed a Persian expedition bent on the destruction of Athens in 490 B.C., has been found to contain concrete evidence of the celebrated battle: a ceremonial burial of some of the fallen Greeks. An earth mound, 40 feet high, was raised on the battlefield soon after the victory. When prehistorians dug into the mound late in the 19th century, however, they found only a layer of ashes and some pottery. This spring archaeologists of the Greek Department of Antiquities, who have been surveying the numerous Neolithic and Bronze Age sites in the area, decided to dig a test trench into a lesser mound, some 10 feet high, that lies a mile west of the better-known battle monument. Within a few yards they uncovered an orderly burial composed of young adult males. Beside each skeleton was a pottery dish and a small phial for oil. One skull showed the mark of a sword or spear thrust.

After the remains of six individuals had been found but before the excavation was completed, the Inspector General of Greek Antiquities, Spyridon Marinatos, announced the discovery and suggested that the small mound covered the dead of the Athenians' allies at Marathon, the Plataeans, and the larger one the Athenian dead. The total of Plataean dead, if proportional to the Athenian losses, would have been roughly 20 men. Marinatos proposes reinvestigation of the larger tumulus to determine whether the 200-odd Athenian casualties lie undiscovered within it or whether, as the ashes unearthed nearly a century ago suggest, those warriors were cremated.

Able-bodied fun machine.

The Marlex® Marketing Team helped create it.

It's a car. It's a boat. It's a super-buggy.

All-terrain vehicles must suffer indignities in forests, swamps, sand dunes, streams and snow banks. Giving them super-tough bodies that wouldn't crack, rust, chip or peel—wherever they go—was the problem.

A Marlex Sales Engineer helped solve it.

And this is how he did it. First, he recommended the right Marlex resin . . . one with superior environmental stress cracking resistance, impact strength and weatherability.

Then he called in other Marlex Marketing Team experts in extruding and thermoforming high density polyethylene. The result: a durable plastic body with performance to spare.

That's the way our Marlex Sales Engineers work. And that's why, if you have any reason to work with plastics, you have every reason to work with Phillips—inventors of the leading process for producing high density polyethylene. The only producers of genuine Marlex, the resin that's worth more.

Get help from the Marlex Marketing Team. Call:

Robert F. Lynch, Fairfield, N.J. 201 227-6600.
Gerald Tucker, Southfield, Mich. 313 353-6666.
Ralph Porter, Villa Park, Ill. 312 834-6600.
Harry House, Memphis, Tenn. 901 682-8317.
David Day, Los Angeles, Calif. 213 685-6380.

PHILLIPS 66

At Phillips 66 it's performance that counts

What is U.S.Steel doing about

"Chicago Picasso," Chicago Civic Center
(© 1967 Public Building Commission of Chicago All Rights Reserved)

art and architecture?

USS COR-TEN Steel is one of the things we're doing.

Add the imagination of architects and this steel becomes a handsome, enduring building. The Chicago Civic Center. Or the Ford Foundation Building in New York. Or any of the thousands of other COR-TEN buildings, bridges and structures all across the country.

Add the genius of Picasso and this handsome material becomes a sculpture. It broods majestically over the city of Chicago like a modern Sphinx.

"Morningside Heights" by William Tarr.

Battelle Memorial Institute, Seattle, Wash.
Architect: Naramore, Bain, Brady & Johanson, Seattle, Wash.

USS COR-TEN is a high strength steel developed and pioneered by United States Steel. It's as natural-looking as wood and as permanent as stone. Artists use it. Architects use it. The secret of its beauty, as well as its durability, is the rich, russet brown coating it forms to protect itself from corrosion. It doesn't need painting. And with age, it only grows more handsome.

Because there's a need for the pleasing as well as the practical, we produce steel that is both.

C. F. Murphy Associates was the architect for the Chicago Civic Center.

USS and COR-TEN are registered trademarks of United States Steel.

Ford Foundation Building, New York City
Architect: Kevin Roche & John Dinkeloo & Associates, Hamden, Conn.

We're involved.

USS

CalComp presents a brilliant new film producer.

Up until now, you could buy your computer either a brilliant and versatile microfilm plotter.

Or a brilliant and high speed microfilm printer.

But you couldn't buy, for any reasonable amount of money, one film output system that was both.

Now you can.

CalComp presents the 1670 COM system. It's a plotter.

CalComp's new offline 1670 COM system is the finest Computer Output Microfilm plotter you can buy.

It draws at a rate of 500,000 increments a second.

Features a programmable raster of 16,384 by 16,384 positions. The finest resolution available.

And cuts tape write time to a tenth or better over previous microfilm systems.

With twenty usable line widths, the 1670 COM system is ready to draw practically anything.

And with the support of CalComp's extensive library of basic, functional and application software, it practically can.

CalComp presents the 1670 COM system. It's a printer.

If your primary need is for a high speed printer, the CalComp 1670 COM system is that, too.

System throughput, using hardware character generation, is 7,000 to 12,000 lines per minute.

And with CalComp software you can put it in any type font, size, or language you like.

With an optional forms projector, you have 16 program selectable forms.

And take your choice of cameras. 16 or 35mm. Sprocketed or unsprocketed. Even a 105mm microfiche option.

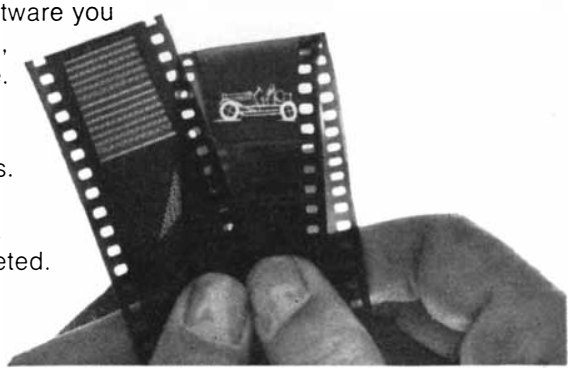
So if you haven't been able to decide between a microfilm system that *plots* or one that *prints,* call your nearest CalComp man today. (CalComp has offices in 34 cities.)

And tell him you want to discuss the brilliant new producer that produces both.

CALCOMP

TEACH YOUR COMPUTER TO DRAW.

California Computer Products, Inc., Dept. N-7
2411 W. La Palma Ave., Anaheim, California 92801



NERVE CELLS AND BEHAVIOR

Can memory and learning be studied at the level of nerve cells and their interconnections? A beginning is made by relating changes in the activity of nerve cells to the modification of simple behaviors

by Eric R. Kandel

The application of biological techniques to behavior has its roots in the writings of Charles Darwin, who argued that since man had evolved from lower animals, human behaviors must have parallels in the behaviors of lower forms. Darwin's radical insights stimulated studies of animal behavior, opening the way to experimentation that was not feasible in man. The studies dealt both with the comparative aspects of behavior in different species and with the analysis of brain mechanisms, two lines of investigation that for a while developed independently.

The comparative studies of behavior proved almost immediately rewarding. It soon became evident that in spite of great variations in the behavioral capabilities of different animals, certain basic response patterns essential for survival—such as feeding, escape and sexual behaviors—were almost universal. Surprising parallels in the details of these behaviors were found in higher and lower animals. Even simple forms of learning seemed to be governed by principles applicable to widely different species.

The analysis of brain mechanisms progressed more slowly at first, but after World War II the development of techniques for studying individual nerve cells revolutionized the neural sciences and made it possible to analyze progressively more complex neuronal processes. As a result a very good understanding has been achieved of the biophysical functioning of the nerve cell and of synaptic transmission, the mechanism whereby one nerve cell communicates with another [see "The Synapse," by Sir John Eccles; SCIENTIFIC AMERICAN, January, 1965]. In addition insights have been gained into the organization of interconnected groups of neurons from cellular studies of sensory and motor systems [see "The Visual Cortex of the Brain," by David H. Hubel; SCIENTIFIC AMERICAN, November, 1963].

In recent years the investigative tradition of comparative behavior and the tradition of brain mechanisms have begun to converge. Advances in the concepts and techniques for studying individual nerve cells and interconnected groups of cells have encouraged neural scientists to apply these methods to studying complete behavioral acts and modifications of behaviors produced by learning. Guided by the lessons of the comparative behaviorists, investigators have sought to study very general behaviors, characteristic of those found in most animals, in species whose nervous systems are amenable to detailed cellular analyses. This led to an interest in certain invertebrates, such as crayfish, leeches, various insects and snails, that have the great advantage that their nervous system is made up of relatively few nerve cells (perhaps 10,000 or 100,000 compared with the trillion or so in higher animals). In these animals one can begin to trace, at the level of individual cells, not only the sensory information coming into the nervous system and the motor actions coming out of it but also the total sequence of events that underlies a behavioral response. By combining psychological techniques for demonstrating the behavioral capabilities of simple animals with cellular techniques for analyzing neural mechanisms, it is now becoming possible to clarify some relations between neural mechanisms and learning. Although cellular concepts and techniques are still far from explaining behavior and learning in higher animals, they are beginning to be useful in understanding elementary forms of behavioral modification in simple animals.

In this article I shall first outline some of the theoretical issues that have influenced modern studies of the mechanism of learning and describe how cellular neurophysiological techniques helped to clarify these issues by revealing that the synapses of nerve cells are functionally modifiable. I shall then try to illustrate how the combined use of cellular and behavioral techniques makes it possible to relate synaptic modifications in certain nerve cells to short-term modification of behavior in an invertebrate.

Cellular-Connection Approach

When investigators first began applying biological techniques to the study of the neural mechanisms of learning, several quite different strategies evolved, but only one of them—the cellular-connection approach—has proved consistently useful. The cellular-connection approach assumes that both the transformation of neural information and its storage as memory involve only nerve cells and their interconnections. This approach derives from morphological studies of the nervous system by the Spanish anatomist Santiago Ramón y Cajal, who held that the nervous system was constructed from discrete cellular units, the neurons, and that the way to understand the brain was to analyze its functional architecture—its wiring diagram. Proponents of this view have therefore focused on the properties of individual neurons, paying particular attention to the synapse, the connection between the nerve cells.

The importance of specific neuronal interconnections in behavior was first demonstrated impressively by R. W. Sperry at the University of Chicago in the 1940's. In a series of studies on the regeneration of neural connections in lower vertebrates, Sperry showed that visual perception and motor coordination could best be explained in terms of highly specific cellular interconnections.

Moreover, these connections seemed invariant and appeared not to be affected by experience [see "The Growth of Nerve Circuits," by R. W. Sperry; SCIENTIFIC AMERICAN, November, 1959].

These studies presented an interesting paradox. If the development of connections between most neurons in the nervous system is rigidly determined, how then is behavior modified? How does one reconcile the known malleability of behavior with a preprogrammed and rigidly "wired" nervous system? One of the characteristic features of learning and other behavioral modifications is their long time course; even a simple behavioral modification lasts for several minutes and certain types of learning may endure for many years. How is the modified neural activity sustained in a set of prewired connections? Do memory and learning require some further additions to the wiring diagram?

Plastic and Dynamic Change

A number of solutions for this dilemma have been proposed. The two that have proved most interesting experimentally are based on notions of the plastic and the dynamic capabilities of neurons. The plasticity hypothesis was first put forward by Cajal and several other neuroanatomists and then in more modern form by two psychologists, Jerzy Kornorski in Poland and Donald O. Hebb in Canada. A current version of this hypothesis states that even though the anatomical connections between neurons may develop according to a rigid plan, the strength or effectiveness of the connections is not entirely predetermined and the effectiveness of synapses and other properties of neurons can be altered by experience. This hypothesis predicted that neurons, and in particular their synapses, should be able to change their functional properties as a result of altered activity.

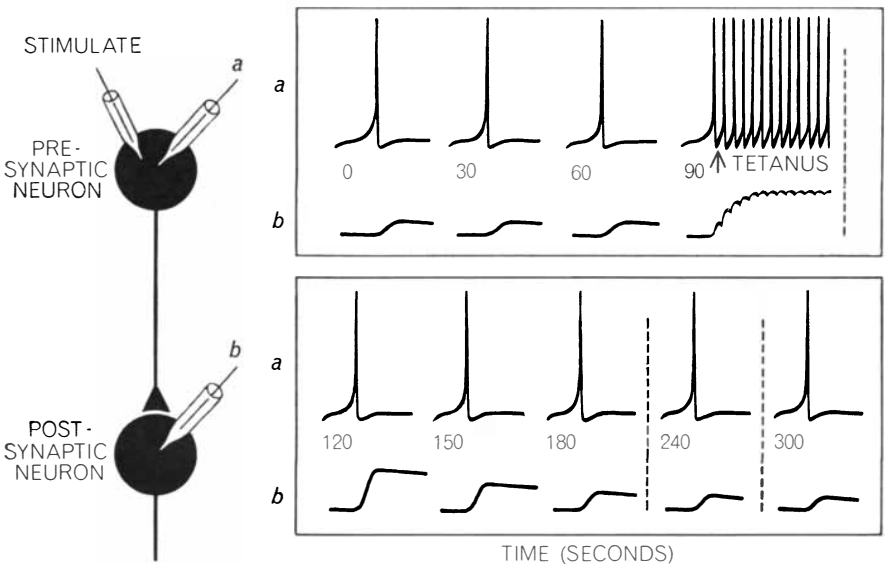
The dynamic hypothesis derived from the anatomical studies of Rafael Lorente de Nó of Rockefeller University, who showed that neurons are often interconnected in the form of closed chains. Neural activity could therefore be sustained by the circulation or reverberation of impulses within a closed chain of interconnected, self-reexciting neurons. This again would not require anatomical change; in fact, the hypothesis does not even require a functional change in the properties of neurons.

The possibility that dynamic changes could account for persistent neural activity was initially very attractive to neurophysiologists because there were many examples in the nervous system of neurons connected to one another in circular paths. As studies of the physiological functions of neural networks advanced, however, it became clear that neurons can mediate inhibition as well as excitation—can tend to quench as well as to fire a nerve impulse. What appeared anatomically to be a self-reexciting loop might therefore contain one or more inhibitory connections that could prevent reexcitation. In addition, memory was shown to survive a number of drastic experimental manipulations, such as cooling of the brain and epileptic convulsions, that would be likely to interrupt the circulation of impulses in closed chains of neurons. As a result of these findings the possibility that dynamic activity provides the neural basis of even short-term memory now seems less likely, although it has not been excluded. Studies of the plastic capabilities of neurons, on the other hand, have turned out to be surprisingly rewarding, since experiments have shown that a remarkable capacity for short-term functional modifications is built into the structure of many synapses.

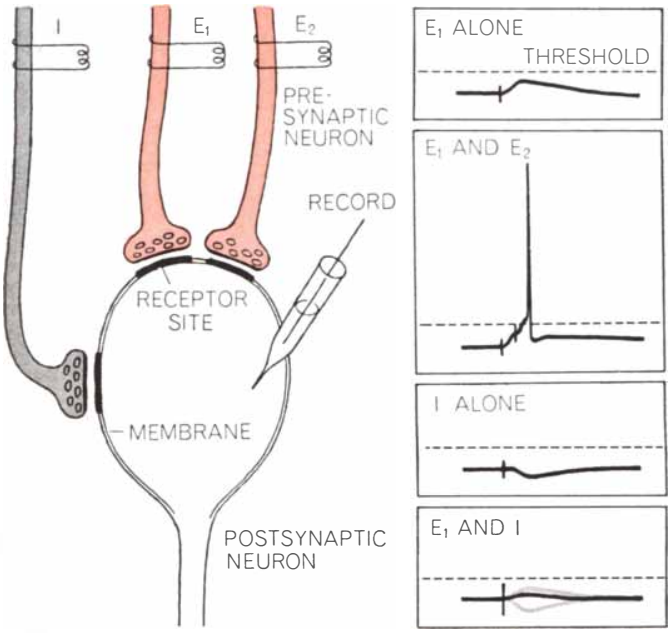
The first demonstration of the plastic capability of a synapse was provided in 1947 by Martin G. Larrabee and Detlev W. Bronk, then at Johns Hopkins University. They studied a simple "monosynaptic" pathway, that is, one consisting of a single class of neurons directly connected to another class of neurons through a single set of synapses. They tetanized (repetitively stimulated at a high frequency) certain fibers leading to the stellate ganglion in the autonomic nervous system of the cat and found that the responsiveness of the stimulated monosynaptic pathway was greatly facilitated, or enhanced, for a few minutes whereas neighboring, unstimulated pathways were unaffected [*see illustration on this page*]. They called this phenomenon posttetanic potentiation.

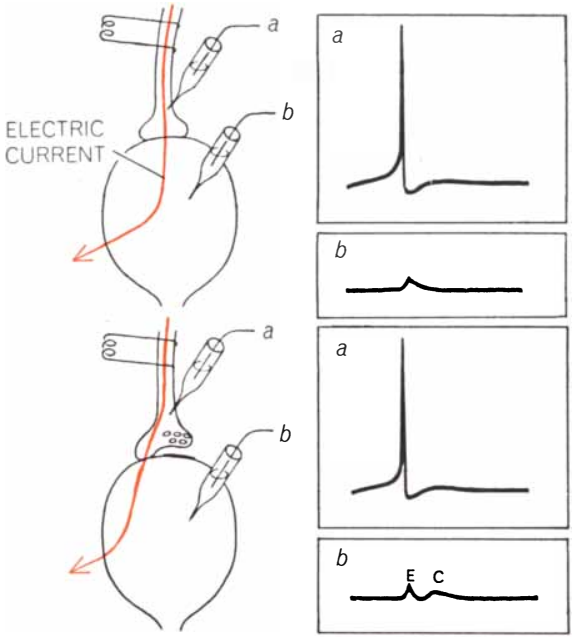
A few minutes is clearly not very long compared with the duration of most learning processes, but it is long compared with the millisecond events that had characterized nerve actions known up to that time. Some behavioral modifications are indeed relatively short-lived, and a plastic mechanism such as posttetanic potentiation might underlie them. Moreover, Larrabee and Bronk showed that the duration of posttetanic potentiation could be extended by longer periods of stimulation. Recently W. Alden Spencer and Reuben Wigdor of the New York University School of Medicine and F. B. Beswick and R. T. W. L. Conroy in Britain have found that posttetanic potentiation can last for as long as two hours after a period of tetanization lasting for from 15 to 30 minutes.

After the discovery of posttetanic potentiation, a number of investigators examined other monosynaptic pathways and encountered an opposite phenomenon, a posttetanic depression, whereby



POSTTETANIC POTENTIATION is a form of plastic change in neural activity. It can be studied in some monosynaptic pathways in which both presynaptic and postsynaptic cells can be impaled with microelectrodes. Stimulation of a presynaptic neuron produces an action potential in it (*a*) that propagates to the nerve terminal and leads to generation of a synaptic potential in the postsynaptic cell (*b*). Repetitive high-frequency stimulation greatly increases the effectiveness of the stimulated pathway, as indicated by the increase in the amplitude of the postsynaptic potential. The increase persists several minutes after tetanus.



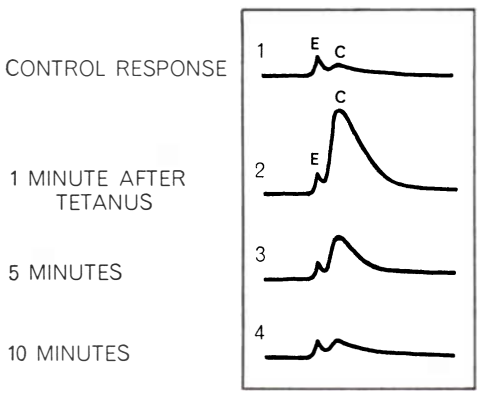


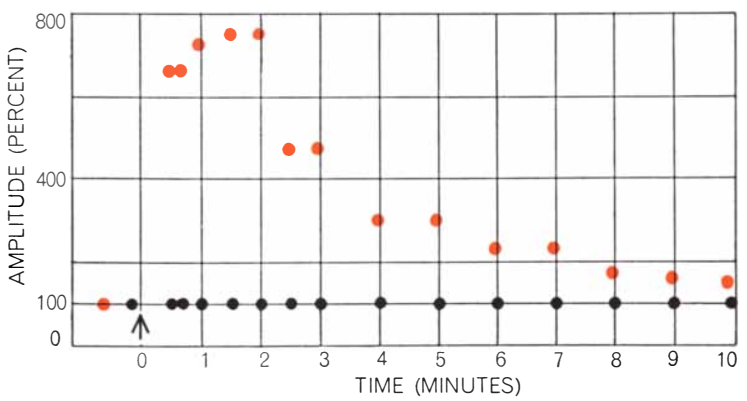
SYNAPSE is the junction between two nerve cells. Most synapses are chemical (*left*), but some are electrical (*top right*) or conjoint (*lower right*). A chemical synapse may be excitatory (*light color*) or inhibitory (*gray*). After a characteristic delay an arriving nerve impulse releases transmitter molecules, thought to be stored in vesicles in the synaptic knob of the presynaptic fiber, that diffuse across the synaptic gap to receptor sites on the postsynaptic cell membrane, increasing its permeability to certain ions and thus changing the electrical potentials of the postsynaptic cell. An impulse in one excitatory fiber (E_1) may produce only a small excitatory postsynaptic potential (EPSP) in the postsynaptic cell. Sequential impulses from two excitatory endings (E_1 *and* E_2) may combine to depolarize the membrane enough to reach threshold, firing the cell. An inhibitory fiber (*I*) changes the membrane potential in the opposite direction, counteracting any excitatory action arriving at the same time (E_1 *and I*). At an electrical synapse the gap between the presynaptic and postsynaptic neuron is reduced and the current produced by the action potential in the presynaptic neuron flows directly into the postsynaptic cell and out across its membrane, depolarizing it. A presynaptic impulse (*a*) can therefore produce an immediate electrical postsynaptic potential (*b*). In a conjoint synapse, both chemical and electrical, the presynaptic impulse (*a*) produces (*b*) first an electrical synaptic potential (*E*) and then, after a delay, a chemical synaptic potential (*C*).

repetitive stimulation leads to a decrease in synaptic effectiveness that sometimes lasts for an hour or more. Other synaptic pathways were found that do not require high-frequency stimulation to undergo plastic changes; they undergo a low-frequency depression when they are stimulated at low rates. For example, Jan Bruner and Ladislav Tauc, at the Marey Institute in Paris, recently described a profound low-frequency depression in a monosynaptic system in the marine invertebrate *Aplysia*. In some synapses one stimulus frequency produces depression and another frequency leads to facilitation. Moreover, plastic changes are not limited to excitatory synapses; they can also occur at inhibitory synapses.

The Synapse

The great advantage of monosynaptic systems is that their anatomical simplicity allows them to be examined directly. It was in monosynaptic systems that Sir Bernhard Katz and his colleagues at University College London and Sir John Eccles and his collaborators at the Australian National University first worked out the general principles that underlie chemical synaptic transmission. They showed that at the synapse the action potential, or nerve impulse, in the termi-





CONJOINT SYNAPSE experiment by Robert Martin and Guillermo R. J. Pilar demonstrated that the chemical synapse is subject to posttetanic potentiation but the electrical synapse is not. The electrical and chemical potential are shown in their usual relation (*1 at left*). After a brief tetanus the chemical synaptic potential is increased but not the electrical (*2*); the chemical synaptic potential remains elevated for more than 10 minutes (*3 and 4*). The chart (*right*) shows the quantitative results of the complete experiment. The arrow indicates the tetanus. The electrical EPSP (*black*) is unaffected; the chemical EPSP (*color*) rises and declines.

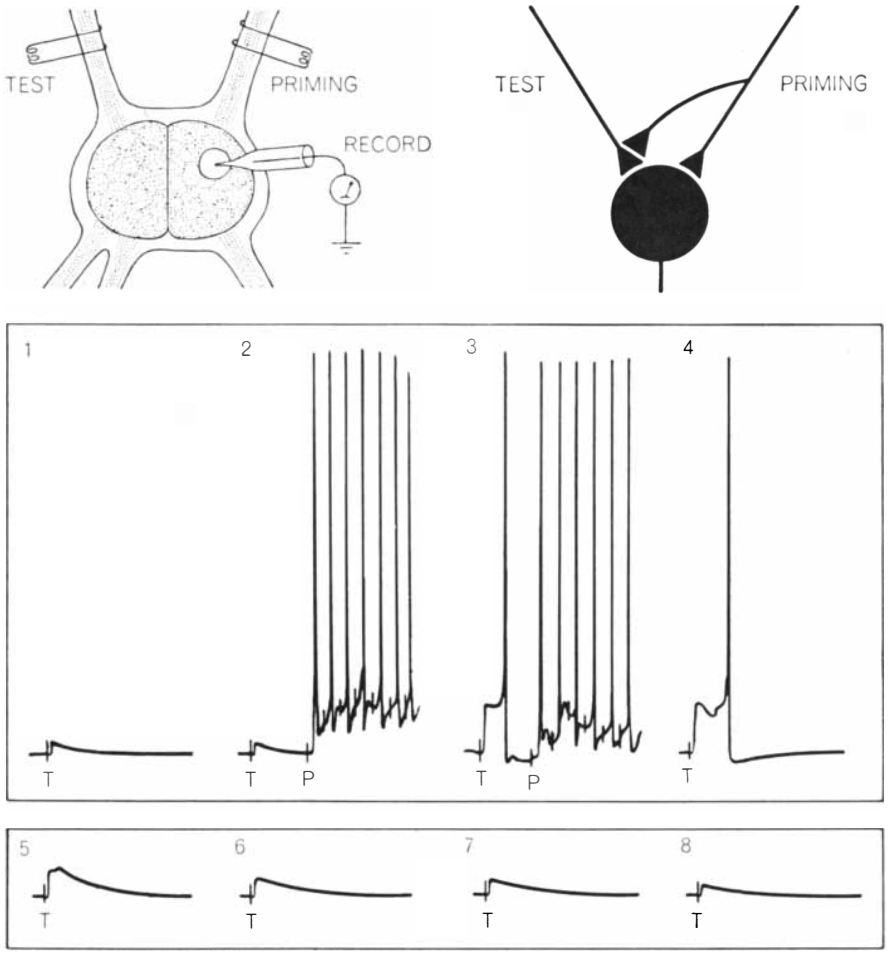
nal part of the neuron leading to the synapse (the presynaptic neuron) triggers the release of a "chemical messenger," or transmitter substance, such as acetylcholine [*see top illustration on preceding page*]. The chemical transmitter diffuses across the gap separating the two neurons at the synapse and interacts with a receptor site on the outer surface of the membrane of the postsynaptic cell. This interaction leads to a change in the permeability of the membrane to certain ions.

At an excitatory synapse the transmitter-receptor interaction produces an increase in the permeability of the membrane to sodium and potassium ions, resulting in a depolarizing potential change: the excitatory postsynaptic potential (EPSP). This makes the membrane potential less negative, moving it toward the threshold at which a new action potential will be discharged in the postsynaptic cell. If the EPSP is large enough, the threshold is exceeded and an all-or-none action potential is triggered. At an inhibitory synapse the transmitter-receptor interaction increases the permeability to potassium or chloride ions, resulting in a hyperpolarizing potential change: the inhibitory postsynaptic potential (IPSP). This moves the membrane potential away from the critical threshold potential, thereby decreasing the possibility of the postsynaptic cell's discharging an action potential.

Studies of monosynaptic systems provided a basis for analyzing the cellular mechanisms of plastic change. In all instances examined plasticity of synaptic pathways has been shown to involve a change in the amplitude of the postsynaptic potential. Detailed study of mechanisms has been possible in the case of posttetanic potentiation. O. F. Hutter of University College London found that the characteristic increase in the EPSP results from an increased amount of transmitter substance released by the presynaptic terminals with each impulse; the sensitivity of the receptor to the transmitter was unaltered. Similarly, posttetanic depression and low-frequency depression seem to be due to a decrease in the amount of transmitter released and not to a change in receptor sensitivity, but the evidence here is less complete.

In 1957 Edwin J. Furshpan and David D. Potter, then at University College London, found that the central nervous system contained not only chemical but also electrical synapses. The two classes of synapses have a number of properties in common, but the electrical synapses (which are less numerous) do not utilize a chemical transmitter; there is a direct flow of current from the presynaptic to the postsynaptic cell [*see top illustration on preceding page*]. The finding of a second class of synapses made possible further exploration of the mechanisms of plastic change. If synaptic plasticity resulted from changes in the action of transmitter substances, then electrical synapses, lacking these transmitters, should have restricted capabilities for plastic change. This hypothesis was soon tested by Robert Martin and Guillermo R. J. Pilar of the University of Utah, who compared electrical and chemical transmission in a "conjoint" synapse of the ciliary ganglion of the chick, where electrical and chemical transmission occur together. Martin and Pilar found that following repetitive stimulation only the chemically mediated EPSP changed; the electrically mediated synaptic potential was not affected. Michael V. L. Bennett of the Albert Einstein College of Medicine reached the same conclusion independently after detailed studies of more than a dozen electrical synapses.

These and other experiments suggest that the predominance of chemical over electrical transmission in the central nervous system may in part be related to the ability of many (perhaps all) chemical synapses to undergo prolonged alterations in efficacy as a result of earlier activity, and thereby to serve as elementary sites for information storage. The striking distinction between the plastic capabilities of electrical and of chemical synapses also strengthens the impression, which emerges from studies of different chemical synapses, that different synaptic pathways can vary greatly in the type and amount of the plastic change they are capable of.



PROLONGED HETEROSYNAPTIC facilitation is studied in the abdominal ganglion of the snail *Aplysia* (*top left*). A test pathway is weakly stimulated and produces a small EPSP (*1*). When a different, "priming" pathway is strongly and repetitively stimulated, it produces a larger synaptic potential, firing the cell (*2*). Priming facilitates the test pathway; it too produces a large EPSP that fires the cell (*3*). Ten seconds after priming, the test pathway alone fires the cell (*4*). It continues to produce a larger potential than control after 3.5, 10 and 20 minutes (*5, 6, 7*), reverting to control size after 30 minutes (*8*). A model postulated for this facilitation has a terminal of the priming pathway ending on the presynaptic terminal of the test pathway, controlling its transmitter release (*top right*).

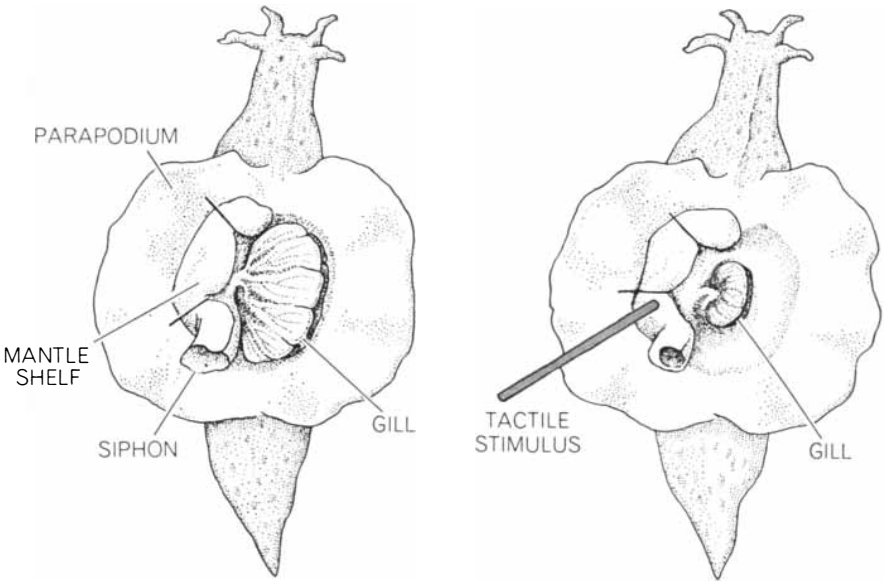
The finding that the pattern of activity along a given pathway can lead to changes in synaptic efficacy at chemical synapses in that pathway was an important step toward understanding how behavior might be modified. With posttetanic potentiation or low-frequency depression, however, the changes in effectiveness are restricted to the repetitively stimulated synaptic pathway. Even simple behavioral modifications characteristically involve activity in two different pathways. Might activity in one pathway produce plastic changes in the synaptic activity of another pathway?

Heterosynaptic Facilitation

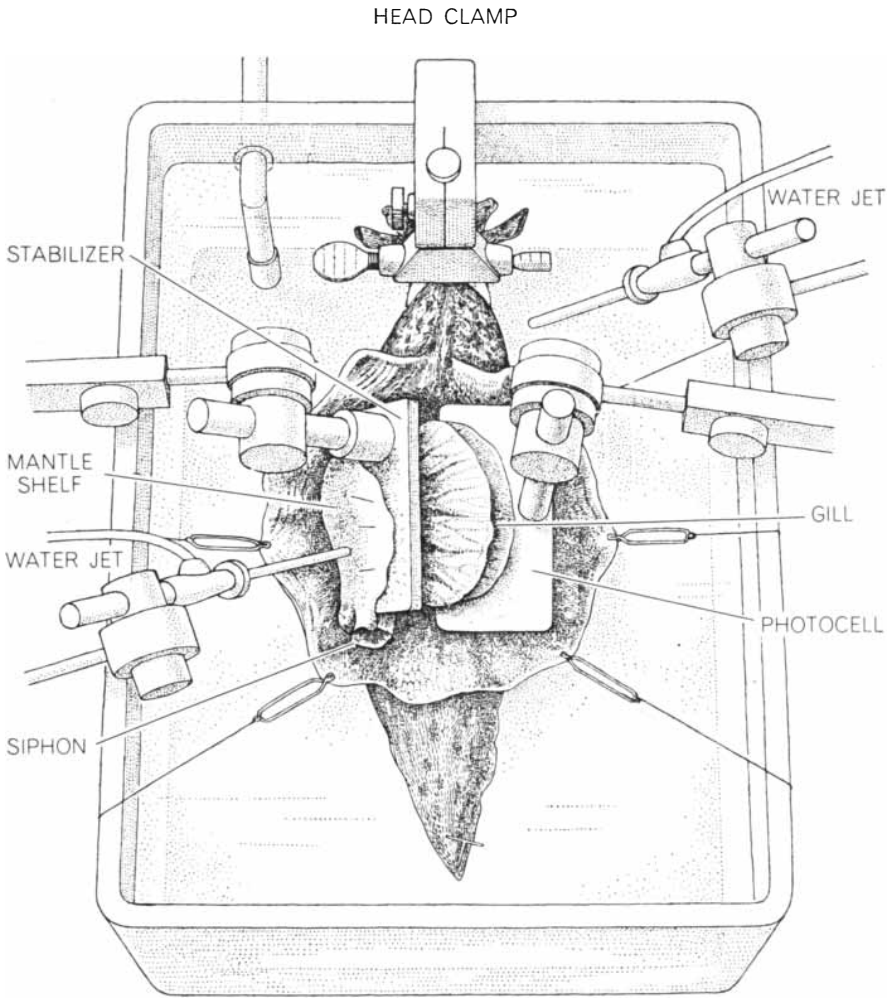
To examine this question, Tauc and I, working at the Marey Institute, investigated the effects of a more complex stimulus sequence on a number of different cells in the abdominal ganglion of *Aplysia*. In a particular cell we found that an EPSP produced by stimulation of one pathway could be greatly facilitated by activity in another pathway. We called this process heterosynaptic facilitation. Although not specific to or dependent on a precise pairing of the stimuli, the stimulus sequence used in these experiments begins to resemble the sequences used in learning experiments. The results provide a neural analogue of quasi-conditioning, or sensitization, the process whereby a strong stimulus enhances other responses.

As was the case for posttetanic potentiation, heterosynaptic facilitation did not result from a change in the threshold or in the biophysical characteristics of the postsynaptic cell membrane during facilitation. The facilitation could be produced in a pathway consisting of only a single fiber (as shown by the fact that it produced an elementary, or unitary, EPSP in the postsynaptic cell), indicating that the synaptic efficacy of the pathway was being directly controlled by the activity of another pathway. Tauc and I therefore proposed that fibers of the facilitating pathway synapse on the presynaptic terminals of the facilitated pathway, and that this presynaptic synapse acts as a governor regulating the long-term release of the chemical transmitter substance in the facilitated pathway [*see illustration on opposite page*].

The idea that activity in one pathway could control activity in another one was not new. A few years earlier Joseph Dudel and Stephen W. Kuffler of the Harvard Medical School and others had reported that some synapses undergo presynaptic inhibition, a mechanism whereby activity in one pathway de-



***APLYSIA*, a marine snail, has a vestigial shell in the mantle shelf. At rest, the gill is partially covered by the mantle shelf (*left*). When the siphon or shelf is touched, the siphon contracts and the gill retracts underneath the shelf in a typical defensive reflex (*right*).**



GILL-WITHDRAWAL REFLEX was studied by clamping the animal in a small aquarium and recording the extent of gill contraction with a photocell. When the siphon or shelf was stimulated by a jet of water, the gill contracted, exposing the photocell to light from above.

presses synaptic transmission in another [see "Inhibition in the Central Nervous System," by Victor J. Wilson; SCIENTIFIC AMERICAN, May, 1966]. That process generally lasts for only a few hundred milliseconds. The studies of *Aplysia* showed that certain pathways could exert a prolonged influence on the activity of others, and that this influence could be facilitatory as well as inhibitory.

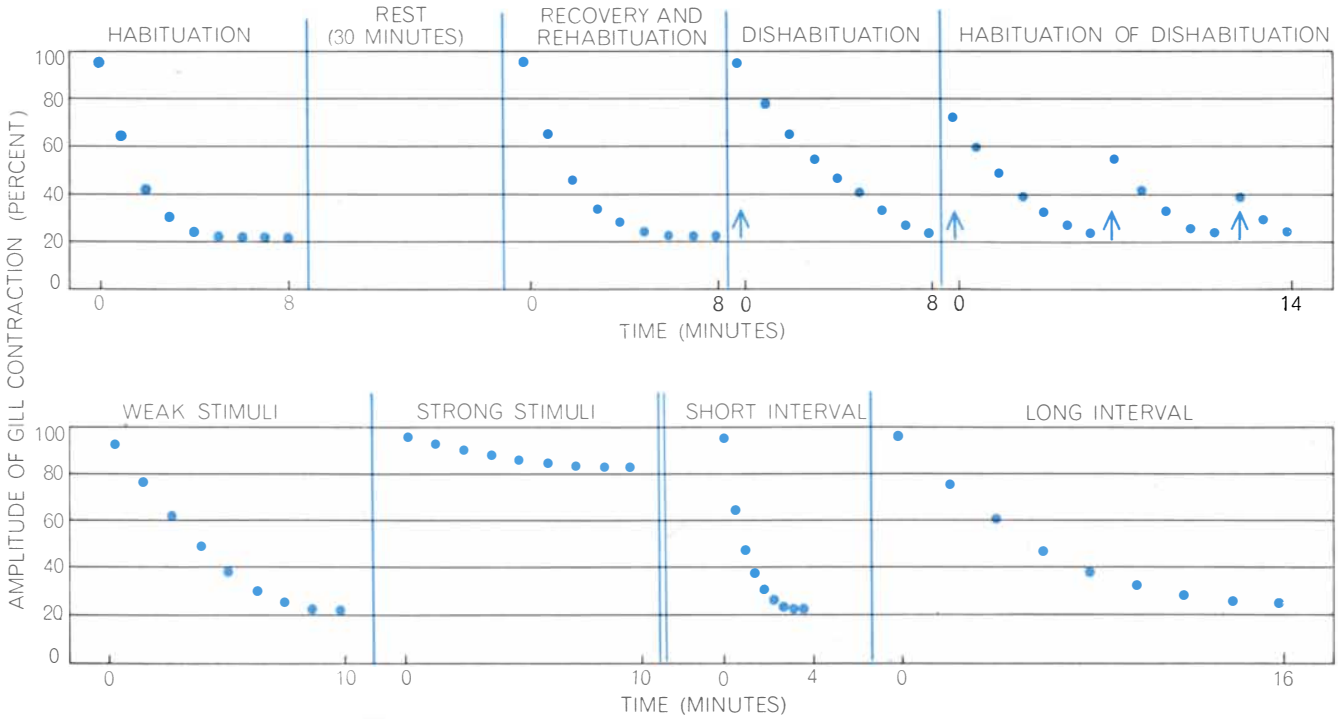
To summarize, different types of plastic change lasting for many minutes and even hours occur at chemical synapses, and there is some evidence that the changes involve two types of regulatory mechanism within the presynaptic terminals that control transmitter release. One class of regulatory mechanisms underlies low-frequency depression, posttetanic depression and posttetanic facilitation and is responsive to activity limited to the stimulated (homosynaptic) pathway. The second class underlies heterosynaptic facilitation and heterosynaptic depression, and involves alterations in one synaptic pathway following activity in an adjacent pathway. The evidence for presynaptic mechanisms is incomplete, however; in several instances there remains an alternate possibility that the change in synaptic effectiveness may be due to a change in the receptor-site sensitivity of the postsynaptic cell.

The detailed study of the plastic capabilities of neurons is only beginning. It is likely that other aspects of neuronal function—such as a cell's threshold or spontaneous firing pattern—may also prove capable of plastic change. It is clear, however, that an important prediction of the cellular-connection hypothesis has been supported: neurons and their synapses have at least some capabilities for plastic change.

Plasticity and Behavior

Given that some synapses can show plastic capabilities, how do these changes in a single nerve cell relate to behavior and its modification? In order to bridge the gap between cellular plasticity and behavior, a detailed knowledge of the anatomy and physiology of the specific neural circuit that mediates the behavior is required. This requirement is very difficult to meet in the intact brain of higher animals because these brains contain an enormous number of cells and an even larger number of interconnections. Moreover, the vertebrate brain mediates highly complex behaviors. One way around these problems is to study simple behaviors under the control of numerically reduced neural populations—either isolated portions of the vertebrate nervous system or simple invertebrate ganglia (discrete collections of cells). The most consistent progress has come from studies of habituation and dishabituation in the spinal cord of the cat and the abdominal ganglion of *Aplysia*.

Habituation, sometimes considered the most elementary form of learning, is a decrease in a behavioral response that occurs when an initially novel stimulus is presented repeatedly. A common example is the habituation of an "orienting response" to a new stimulus. When a stimulus such as a sudden noise is presented for the first time, one's attention is immediately drawn to it, and one's heart rate and respiratory rate increase. As the same noise is repeated, one's attention and bodily responses gradually diminish (which is why one can become accustomed to working in a noisy office). In this sense habituation is learning to accommodate to stimuli that have lost novelty or meaning. Besides being important in its own right, habituation is frequently involved in more complex learning, which consists not only in acquiring new responses but also in eliminating incorrect responses. Once a response is habituated two processes can lead to its restoration. One is spontaneous recovery, which occurs as a result of withholding the stimulus to which the animal has habituated. The other is dishabituation, which occurs as a result of changing the stimulus pattern, for example by presenting a stronger stimulus to another pathway. Similar types of response decrement and restoration, last-

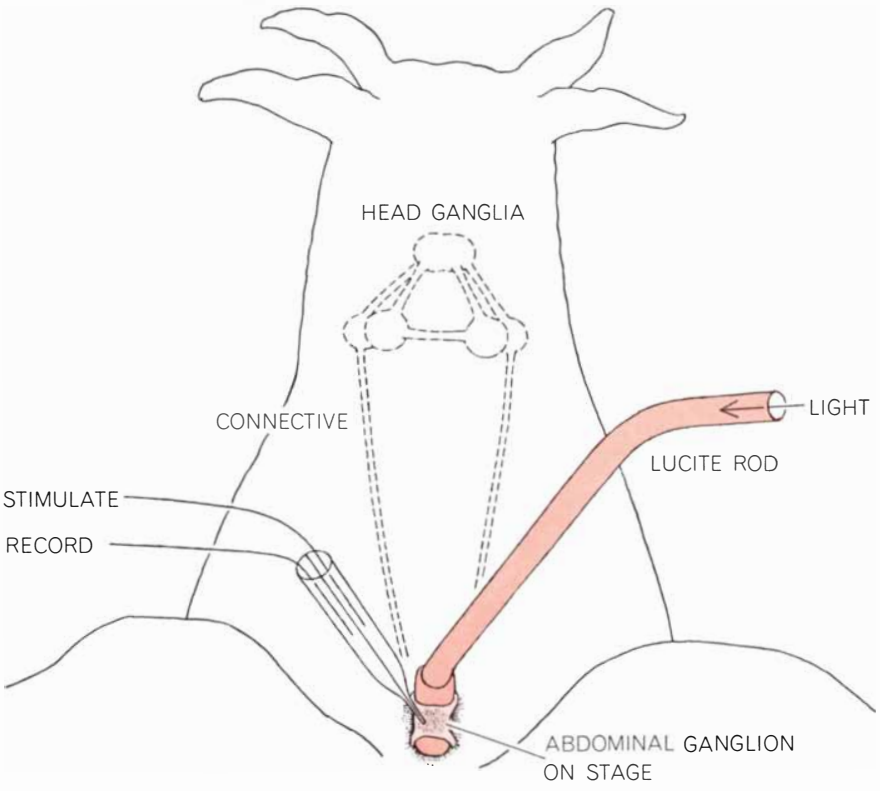


HABITUATION AND DISHABITUATION in the gill-withdrawal reflex have properties characteristic of these behavioral modifications in higher animals. The response decreased (habituation) with successive stimulations, recovered slowly after a rest period and then rehabituated when repeated stimuli were reintroduced. It could also be restored rapidly to control value (dishabituation) by the presentation of a strong new stimulus (*arrow*). The dishabituatory stimulus habituated in turn after repetition. The reflex also habituated more rapidly with weak rather than strong stimuli and with short rather than long intervals between stimuli.

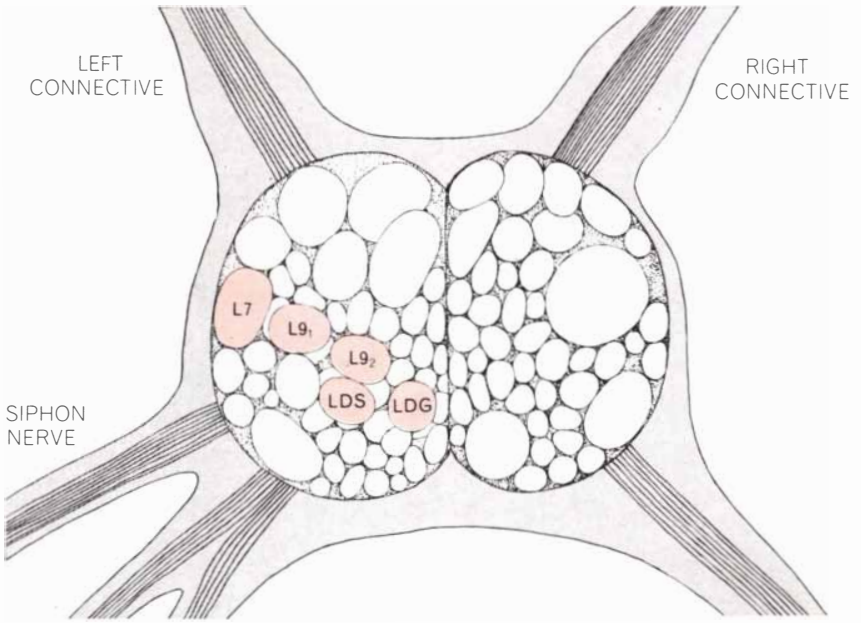
ing for from several minutes to several hours, have been demonstrated for a wide variety of behaviors in all animals that have been examined, including man. The existence of such general ways of modifying behavior suggests that their neuronal mechanisms may also prove to be quite general.

The first neural analysis of habituation was undertaken in the spinal cord of the cat. The spinal cord mediates the reflex responses underlying posture and locomotion in vertebrates. In the course of analyzing the neural mechanism of spinal reflexes, Sir Charles Sherrington, the great British physiologist, found that certain reflex responses, such as the flexion withdrawal of a limb in response to stimulation of the skin, decreased with repeated stimulation and recovered only after many seconds of rest. Sherrington was greatly influenced by Cajal's work, and he tried to develop his notions of reflex action in relation to the anatomical diagrams Cajal had worked out. (In fact, it was Sherrington who coined the term "synapse" for the zone of apposition between two neurons, which Cajal had described.) Sherrington had the great insight to appreciate that neural reactions within the spinal cord differed from those in the peripheral nerves because of the numerous synaptic connections within the cord. It was therefore perhaps only natural for him to attribute the decrease in the responsiveness of the withdrawal reflex to a decrease in function at the synapses through which the motor neuron responsible for flexion was repeatedly activated. In other words, as early as 1908 Sherrington had suggested that a plastic change at central synapses could underlie response decrement. He was able to show that the decrement was not due to fatigue of either the muscles or the sensory receptors but, with the neurophysiological techniques then available, he could not fully test his intriguing synaptic hypothesis.

The problem was reinvestigated by C. Ladd Prosser and Walter Hunter of Brown University, who found that the habituated flexion withdrawal can be restored to full size (dishabituated) by the application of a strong new stimulus to another part of the skin. More recently Alden Spencer, Richard F. Thompson and Duncan R. Neilson, Jr., then at the University of Oregon, found that dishabituation is not simply a transient wiping out of habituation but an independent facilitatory process superimposed on habituation, and they established that spinal-reflex habituation resembled the habituation of more complex behavioral responses. Spencer and his colleagues also began the cellular analysis of habituation. By recording intracellularly from a motor neuron they showed that response decrement involved a change not in the properties of the motor neurons but in the synaptic impingement on them. They could not localize the critical changes precisely, however, because the central synaptic pathways of the flexion withdrawal reflex in the cat are complex, involving many connections between sensory and motor cells through interneurons that have not yet been worked out. Barbara Wikelgren of the Massa-



ABDOMINAL GANGLION of *Aplysia* is a group of some 1,800 cells at the forward end of the animal's abdomen, near the head. It is connected to the head ganglia by connective nerves. To study the individual cells of the abdominal ganglion, the author and his colleagues cut a slit in the skin and lift ganglion out on a stage lighted by a curved Lucite rod.



CELLS OF GANGLION have been classified and a number have been identified as recognizable individuals. By firing cells one at a time with a microelectrode, five cells (*color*) in one part of the ganglion were identified that produced movements of the withdrawal reflex.

chusetts Institute of Technology has gone on to examine several populations of interneurons and has shown that decremental changes occur only in certain classes of them, but the complexity of the pathways has still prevented her from distinguishing between plastic and dynamic changes in neuronal activity or specifying whether inhibitory or excitatory synaptic transmission is altered.

Gill Withdrawal in *Aplysia*

The further analysis of habituation required a still simpler system, one where the components of the behavioral response could be reduced to one or more monosynaptic systems. In search of such systems, a number of researchers have been attracted to the invertebrate nervous system because these systems contain relatively few cells, thereby reducing the task of a behavioral analysis [see "Small Systems of Nerve Cells," by Donald Kennedy; SCIENTIFIC AMERICAN, May, 1967]. The nervous system of marine gastropod mollusks is particularly advantageous because it contains cells that are unusually large (some reach almost one millimeter in diameter) and therefore are suitable for study with intracellular microelectrodes. For this reason my colleagues and I at the New York University School of Medicine have worked on the abdominal ganglion of *Aplysia*, a giant marine snail that grows to about a foot in size. An opisthobranch, *Aplysia* differs from the better-known pulmonate snails in having only a very small residual shell. Like some other marine mollusks, such as the octopus and the squid, *Aplysia* gives off a brilliant purple ink when it is overly perturbed.

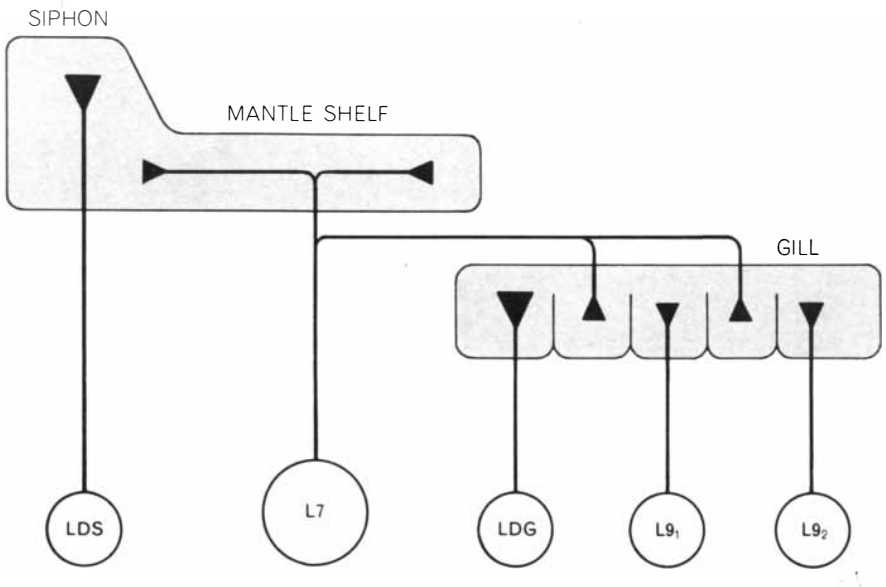
The abdominal ganglion of *Aplysia* has only about 1,800 cells and yet is capable of generating a number of interesting behaviors. Wesley T. Frazier, Irving Kupfermann, Rafiq Waziri and I collaborated with Richard E. Coggeshall, an anatomist at the Harvard Medical School, to classify the cells in this ganglion into different functional clusters. We also identified 30 cells as unique individuals and mapped a number of their central connections. This provided a good starting point for an attempt to examine the cellular changes in habituation and dishabituation. Our plan was to delineate a behavior that is controlled by cells in the abdominal ganglion and that undergoes habituation, to analyze the neural circuit of this behavior and then to try to specify the locus and the nature of the functional change in the neural circuit that underlies the modification of behavior. Kupfermann and I initiated this work and we were soon joined by Harold Pinsker and Vincent Castellucci.

An *Aplysia* shows a defensive withdrawal response that is in some ways analogous to the flexion withdrawal response in the cat. The snail's gill, an external respiratory organ, is partially covered by the mantle shelf, which contains the thin residual shell. When either the mantle shelf or the anal siphon, a fleshy continuation of the mantle shelf, is gently touched, the siphon contracts and the gill withdraws into the cavity underneath the mantle shelf. If the stimulus is very strong, it brings in other behavioral components, such as ink production. The defensive purpose of this reflex is clear: it protects the gill, a vital and delicate organ, from possible damage. Gill withdrawal is therefore analogous to the withdrawal of a man's hand from a potentially damaging stimulus. As in the case of these other defensive responses, the gill-withdrawal reflex habituates when it is repeatedly elicited by a weak or harmless stimulus.

When Thompson and Spencer reviewed the literature on habituation in vertebrates, they described nine features that characterize this simple behavioral modification. We found six of these characteristic features in the gill-withdrawal reflex in *Aplysia:* (1) response decrement, typically a negative exponential function of the number of stimulus presentations; (2) recovery with rest; (3) dishabituation; (4) habituation of the dishabituatory stimulus as it is repeated; (5) greater habituation with weak rather than strong stimuli, and (6) greater habituation with short rather than long stimulus intervals.

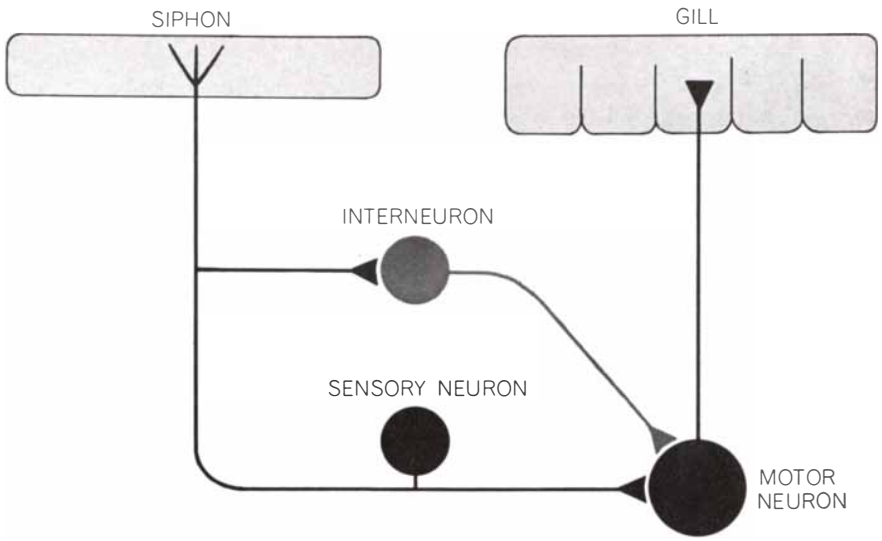
The fit between short-term habituation in *Aplysia* and in mammals was encouraging, and we went on to analyze the neural circuit of this behavior. To this end Kupfermann and I made a small slit just forward of the snail's mantle region through which the abdominal ganglion, with its connectives and peripheral nerves attached, was lifted out on an illuminated stage [*see top illustration on preceding page*] so that various neurons could be impaled with double-barreled microelectrodes for recording and direct stimulation.

First we searched for the motor neurons of the reflex—the component leading outward from the central nervous system. We did this by firing different cells one at a time with the microelectrode and observing whether or not they produced movements of the external organs of the mantle cavity. By this means we identified five motor cells, clustered together in one part of the ganglion, that produced contractions of the gill, the siphon or the mantle shelf. Three of the five cells produced movements limited to the gill, one cell produced movement of the siphon only and one cell produced movement of the gill, the siphon and the mantle shelf [*see illustration at left*]. The motor component of this reflex, then, consists of individual elements with both restricted and overlapping distributions; it is redundant, as are other motor systems that have been described in invertebrates. Turning to the afferent component of the reflex, which leads toward the central nervous system, we next mapped the sensory receptive field of the motor cells by stimulating the surface of the animal's body with light brush-



INDIVIDUAL ACTIONS of the five cells are shown. Three produce movement of the gill alone; one, movement of the siphon alone, and one, movement of gill, siphon and shelf.

strokes. We found that all five motor cells received large EPSP's, producing a brisk repetitive spike discharge, when the siphon or the mantle shelf was stimulated. The tactile receptive field of these five motor cells was identical with that of the defensive withdrawal reflex. Our analysis suggested a simple pathway for the excitatory input, in part monosynaptic and in part mediated by interneurons [*see top illustration on this page*]. At this early stage we could not exclude the possibility of an underlying inhibitory component.



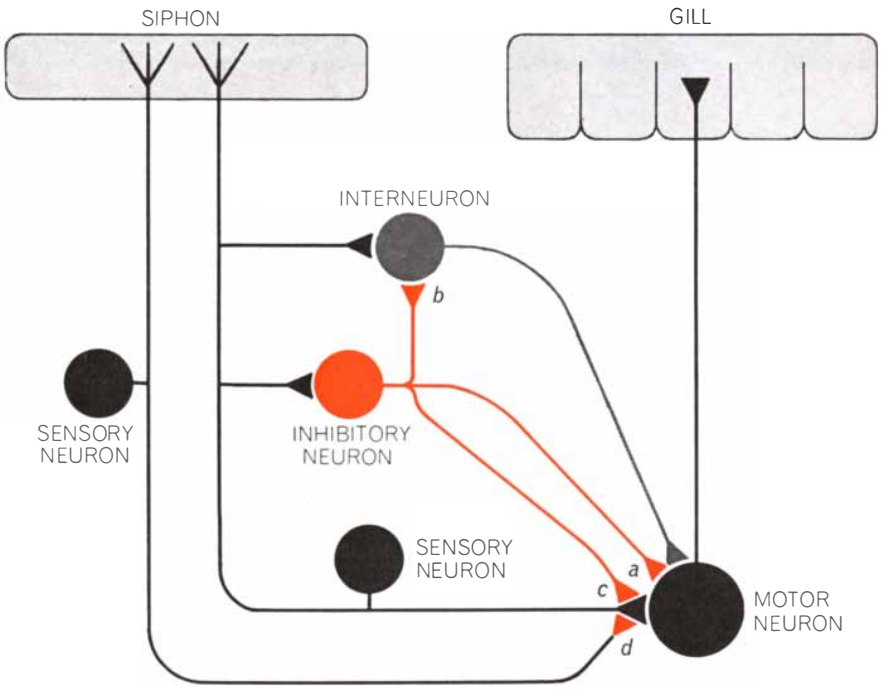
SIMPLE WIRING DIAGRAM for the gill component of the withdrawal reflex shows how the excitatory input from the sensory cells could fire the motor cells both directly and via intervening excitatory interneurons. (In *Aplysia* and other invertebrates the presynaptic neuron normally ends on the beginning of the fiber of the postsynaptic cell, as in the upper illustration on the next page. Here the synapse is indicated schematically, on the cell body.)

The Nature of the Changes

Having analyzed at least part of the wiring diagram of the total withdrawal reflex, we next examined the changes that accompany habituation and dishabituation of the gill component of the total reflex. These changes could occur at a number of points in the circuit. The first possibility, that habituation was due to changes outside the central nervous system either in the muscle or in the sensory receptors in the skin, was quite readily excluded. Stimulation of the individual motor neurons or of the peripheral nerves innervating the gill at rates that are effective in producing habituation did not give rise to a significant decrease in gill contraction. Moreover, the size of this directly evoked gill response was the same before, during and after habituation or dishabituation of the reflex response. Finally, single sensory axons recorded in the peripheral nerves did not change their rate of firing when the sensory stimulus was repeated at rates that produced reflex habituation.

Since neither muscle fatigue nor sensory accommodation could account for the habituation of the gill reflex, habituation must result from some functional change within the central nervous system. We therefore examined the EPSP produced in the major motor neurons of the gill by tactile stimulation of the siphon or mantle shelf and found that it underwent characteristic changes that were causally related to habituation, to recovery and to dishabituation. When a tactile stimulus was repeated to produce habituation at the behavioral level, the resulting EPSP in the gill motor neuron gradually decreased in size, and the amount and frequency of the spike activity it evoked decreased correspondingly. Recovery of reflex responsiveness, produced either by rest or by a dishabituatory stimulus, was associated with an increase of the EPSP and a corresponding increase in spike activity.

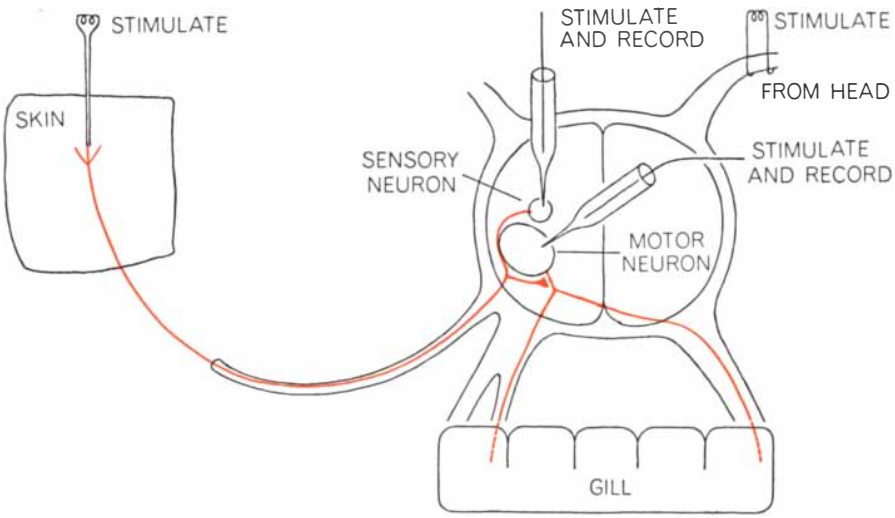
We next began a more detailed examination of the mechanism underlying these changes in the synaptic potential. First we examined whether the synaptic changes resulted from alteration of the synaptic impingement on the motor neuron or from a change in the biophysical properties of the motor neuron itself. The amplitude of any synaptic potential is determined by the amount of synaptic current that flows across the resistance provided by the cell membrane outside the synaptic area [*see top illustration on page 59*]. A synaptic potential can therefore be altered in two ways: (1) by a direct change in synaptic current, produced either by a change in the amount of transmitter released by presynaptic neurons or by a change in the sensitivity



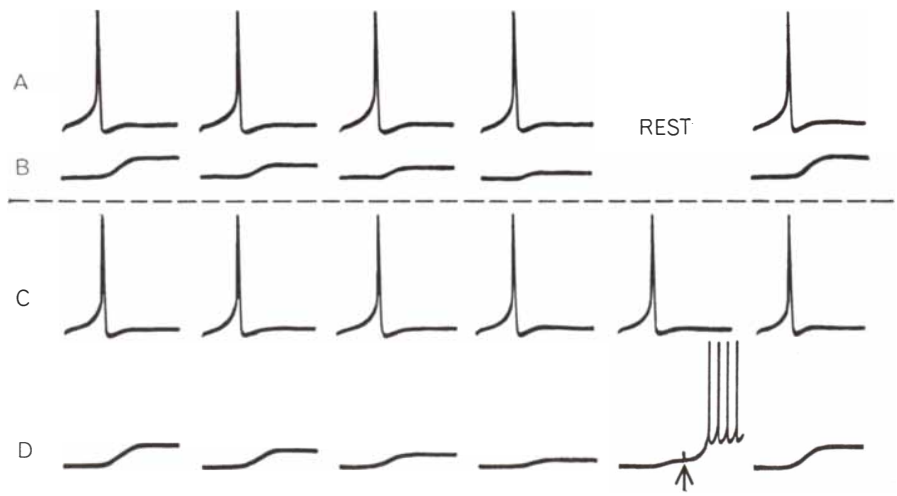
INHIBITION might reduce the excitation of motor neurons and thus cause the decreased responses of habituation. If an inhibitory interneuron were involved in the pathway, it might operate on the motor neuron itself (*a*), on an excitatory interneuron (*b*) or on the terminals of the afferent sensory fiber (*c*). Alternatively, parallel sensory fibers in a different pathway, activated by the same stimulus, might effect presynaptic inhibition (*d*).

of the postsynaptic neuron's receptor sites; (2) by a change in the resistance of the membrane of the postsynaptic cell. To distinguish between these two possibilities we measured the resistance of the postsynaptic cell and found that it was unaltered during habituation and dishabituation. The changes in the synaptic potential during habituation and dishabituation must therefore reflect changes in the synaptic current. This could have been caused by a change in receptor sensitivity or, what is more likely, by a change in the amount of transmitter that is released by the total synaptic impingement on the motor neuron.

Many elements impinge on a motor neuron and a change in synaptic impingement could therefore be caused by one of a number of factors. If for the moment we consider only habituation, then the decrease in the amplitude of the EPSP might be due to a decrease, caused by inhibition, in the number of active excitatory elements impinging on the motor neuron and thus contributing to the total synaptic potential. If the afferent fibers that excite the motor neuron also activated a side chain consisting of one or more inhibitory neurons, an inhibitory neuron could reduce the excitatory activity by acting on the motor neuron directly, on an excitatory interneuron or on the excitatory terminals of the sensory fiber [*see bottom illustration on preceding page*]. If repeated stimulation then produced a sustained increase in inhibitory action (by either dynamic or plastic means), it would result in progressive suppression of either the motor neuron or the excitatory interneuron that helps to discharge the motor neuron. Alternatively, habituation could occur without any increase of active inhibition; there might simply be a progressive decrease in the effectiveness of the individual excitatory synaptic actions.



INDIVIDUAL EXCITATORY ELEMENT was isolated to eliminate inhibitory neurons in the pathway as sources of habituation. A piece of skin was probed to produce in the motor neuron an "elementary" EPSP, one clearly evoked by a single cell. This EPSP showed decrement. In another experiment the cell body of an individual sensory neuron was identified and stimulated directly, and its decrement was observed. Dishabituation was also studied in this pathway, with electrical stimulation to the connective as the dishabituatory stimulus.



NEURAL DETERMINANTS of habituation and dishabituation are seen in these traces from an individual sensory neuron and motor neuron. Stimulation produces an action potential in the sensory neuron (*A*) that induces an elementary EPSP in the motor neuron (*B*); repeated stimulation produces a decrement of the EPSP; rest brings about recovery. Dishabituation was demonstrated by firing the sensory nerve (*C*) to produce habituation (*D*), then delivering a stronger stimulus (*arrow*) to a head connective, increasing the EPSP.

Eliminating Inhibition

In order to distinguish between these two possibilities it was necessary to examine an individual excitatory element in isolation to see whether, in the complete absence of inhibition, the EPSP it produced would show the characteristic decrement, paralleling behavioral habituation, in the course of repeated stimulation. To this end we radically simplified the afferent component of the reflex pathway by separating a small piece of siphon skin that is part of the receptive field of the reflex together with its afferent nerve to the ganglion, and isolating the ganglion from the rest of the nervous system by removing it from the animal to an experimental chamber [*see upper illustration at left*]. In this simplified preparation we searched the skin until we found a responsive region that when stimulated produced an "elementary" EPSP in the motor neuron, that is, a unitary signal reflecting the synaptic response produced by a single sensory fiber. When this elementary EPSP was repeatedly evoked, it paralleled the gill response, decreased with repeated stimulation and recovered following rest.

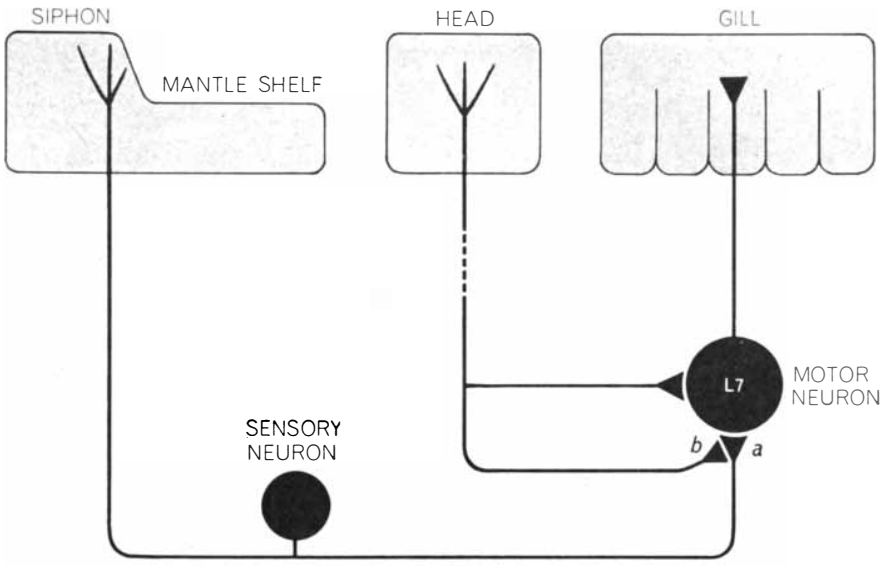
This established that the change in potential was not the result of inhibition in the direct pathway from sensory to motor nerve. Inhibition was not yet ruled out, however. Since the stimulus that produced the elementary EPSP might also be activating other sensory fibers, the decrement in the synaptic potential might be due to presynaptic inhibition from parallel afferent fibers [*see bottom illustration on preceding page*]. To eliminate this last possibility we had to stimulate a single sensory neuron in isolation, thereby ensuring that it was acting alone.

This appeared at first to be a difficult experimental objective. It could best be

achieved by stimulating the cell body of a sensory neuron with an intracellular electrode, and the cell bodies of mechanoreceptor neurons were believed to be located in the skin, where they are difficult to isolate for recording and stimulation. Soon after we started work on this problem, however, John G. Nicholls and Denis A. Baylor of the Yale University School of Medicine reported that the cell bodies of mechanoreceptor neurons in the leech were located not in the skin but in the central ganglia; fortunately we found a similar situation in the abdominal ganglion of *Aplysia*. We identified a group of sensory cells that were excited by mechanical stimulation of small regions of the skin of the siphon and that had properties consistent with their being primary sensory neurons of the gill-withdrawal reflex. By stimulating and recording from one of these mechanosensory neurons and simultaneously recording from a gill motor neuron, we finally were able to reduce the gill reflex to its most elementary monosynaptic components and to examine the two cells in turn as well as their interaction [*see upper illustration on opposite page*].

Stimulation of one of the mechanoreceptor neurons produced a fairly large elementary EPSP in the motor neuron. Repeated stimulation produced a dramatic decrease in the amplitude of the potential; as in the case of stimulation of the skin, the potential produced by direct stimulation of the sensory neuron sometimes diminished so markedly that after a few stimuli it was barely visible. Rest led to recovery [*see lower illustration on opposite page*]. These data make it very unlikely that inhibition is involved and suggest that decrement of the EPSP is due to a plastic change in excitatory synaptic efficacy. This could result from either a decrease in the amount of transmitter released with each impulse or a decrease in the sensitivity of the postsynaptic receptor sites.

Dishabituation, the facilitation of a decreased response that occurs when a strong new stimulus is applied to another part of the receptive field, has also been studied in the isolated ganglion. We found that an EPSP that had decreased through habituation could be restored by a strong stimulus to the right or left connective nerve, which carry fibers from the head ganglia; at times the facilitated synaptic potential was even larger than it had been before habituation. In studies of the individual mechanoreceptor neurons we found that the dishabituation took place without a change in the frequency of firing of the sensory neuron [*see lower illustration on opposite page*]. This excludes posttetanic potentiation as a mechanism for the facilitation and suggests that dishabituation involves a heterosynaptic (presynaptic) facilitation.



PROPOSED WIRING DIAGRAM for the elementary withdrawal reflex does not involve inhibition. The plastic change of habituation occurs at the excitatory synapse between sensory and motor neuron (*a*). Stimulation of head causes dishabituation, apparently by acting presynaptically at the synapse (*b*), increasing transmitter release from sensory terminals.

As in the case of the gill-withdrawal reflex in the intact animal, the elementary monosynaptic EPSP's produced by individual sensory neurons had many of the accepted characteristics of habituation and dishabituation: (1) decrement with repeated stimulation; (2) recovery with rest; (3) facilitation following a strong stimulus to another pathway; (4) decrement of the facilitatory effect with repetition; (5) greater decrement with short rather than with long interstimulus intervals. (We have not examined the effects of the sixth variable, stimulus intensity, on the elementary EPSP.)

The Wiring Diagram

We can now propose a simplified circuit diagram to illustrate the locus and mechanism of the various plastic changes that accompany habituation and dishabituation of the gill-withdrawal reflex [*see illustration above*]. The motor neuron is *L7*, on which the work with the elementary EPSP was done. Repetitive stimulation of sensory receptors leads to habituation by producing a plastic change at the synapse between the sensory neuron and the motor neuron. The exact mechanism of the synaptic change is uncertain. We still cannot exclude a change in receptor sensitivity but, by analogy with the brief low-frequency synaptic depression analyzed in detail in vertebrate and crayfish neuromuscular junctions, we tend to think the decrement of the elementary EPSP represents a decrease in the release of excitatory transmitter from the presynaptic terminal. Unlike the brief low-frequency depression, however, the decrement in the habituation pathway is remarkably large and prolonged. These features suggest that perhaps additional elements operate in synapses of the habituatory pathway. Stimulation of the head leads to dishabituation by producing heterosynaptic facilitation at the same synapse. This facilitation appears to operate presynaptically, perhaps by enhancing the release of transmitter substance.

We have used the monosynaptic pathway between the mechanoreceptor neurons and one of the motor neurons (*L7*) as a model for studying the total reflex. Comparable experiments with monosynaptic inputs still need to be done on other gill motor neurons and on the polysynaptic pathway in order to account quantitatively for the complete behavioral modifications. It is clear already that the unhabituated elementary EPSP's produced by a single spike in individual sensory neurons are relatively large and often trigger an action potential in the motor neuron. In addition there are at least 10 such sensory neurons that synapse on the motor neuron. It therefore seems likely that a substantial portion of the total EPSP in the gill motor neuron *L7* is due to the monosynaptic EPSP's

Silk on Silk on Polaroid Land Film

George Silk, the world's best-known outdoor and sports photographer, has been on Life magazine's staff for many years.

He has been acclaimed Photographer of the Year four times, an unmatched record.

"I was trying to take shots of a young girl Olympic diver doing a perfect dive... making a perfect entry.

To get what I wanted, I decided to stay in a tiny room at the side of the pool, and shoot through the window the coach uses to view dives from under water.

The pool level had to be lowered until the water was halfway up the window, a very slow process.

But the shooting itself couldn't take too long. If the girl made too many dives, and got tired, they'd no longer be perfect.

My timing had to be precisely coordinated with hers. And the worst problem was that I'd never be able to see if I got it just right until too late.

It was a real dilemma.

But then I remembered Polaroid Land film. Previously, I'd only used it experimentally. Just for fun. But it seemed to be the answer I was looking for.

And as you see, it worked beautifully. I tried several shots. Each time I had the results in my hands in a few moments, and could adjust my timing where it was off. After only a few tries, I had what I wanted.

I was literally able to lay out the picture story as I shot it. With the exact, final pictures.

Best of all, Polaroid Type 55 P/N film gave me the negatives I needed that would allow beautiful blow-ups of the shots.

I guess the kind of resolution I got is self-evident. This picture was hung in the Baltimore Museum of Art, in their "Man in Sports" show.

I'll say this for Polaroid Land film. When you're stuck with certain problems, there just isn't anything else that can solve them for you."

Photograph by permission of LIFE Magazine ©Time Inc.

from mechanoreceptor neurons. Furthermore, the spike activity of *L7* contributes a substantial part of the total gill contraction. Since changes in the spike activity of the motor neuron *L7* are directly produced by changes in EPSP amplitude, a substantial part of the habituation and dishabituation of the early component of the withdrawal reflex can be explained by alterations in the efficacy of excitatory synapses between the sensory and the motor neurons. Indeed, if similar processes occur on the other motor neurons and on the interneurons, these mechanisms could explain all the habituation and dishabituation.

Our data lead to some more general conclusions. First, habituation and dishabituation appear to involve a change in the functional effectiveness of previously existing excitatory connections. In these simple cases, at least, it therefore seems unnecessary to explain the behavioral modifications by invoking the growth of new connections or dynamic activity in a group of cells. The capability for behavioral modification seems to be built directly into the neural architecture of the behavioral reflex. Since the architecture is redundant and distributed, the capability for modification is redundant and distributed too.

Second, although a number of investigators have postulated on the basis of indirect evidence that habituation involves active inhibition, in *Aplysia,* where a major component of the synaptic mechanism of habituation can be studied directly, neither presynaptic nor postsynaptic inhibition appears to be critically involved.

Third, these experiments indicate that habituation and dishabituation are separate processes and that dishabituation is not merely an interruption of the decrease associated with habituation but an independent facilitatory process superimposed on the habituation. (Seen in this perspective, dishabituation is essentially a special case of quasi-conditioning.) Although habituation and dishabituation are independent, they do not involve different neurons with overlapping fields of action. Rather, the two processes appear to represent two independent regulatory mechanisms acting at the same synapse.

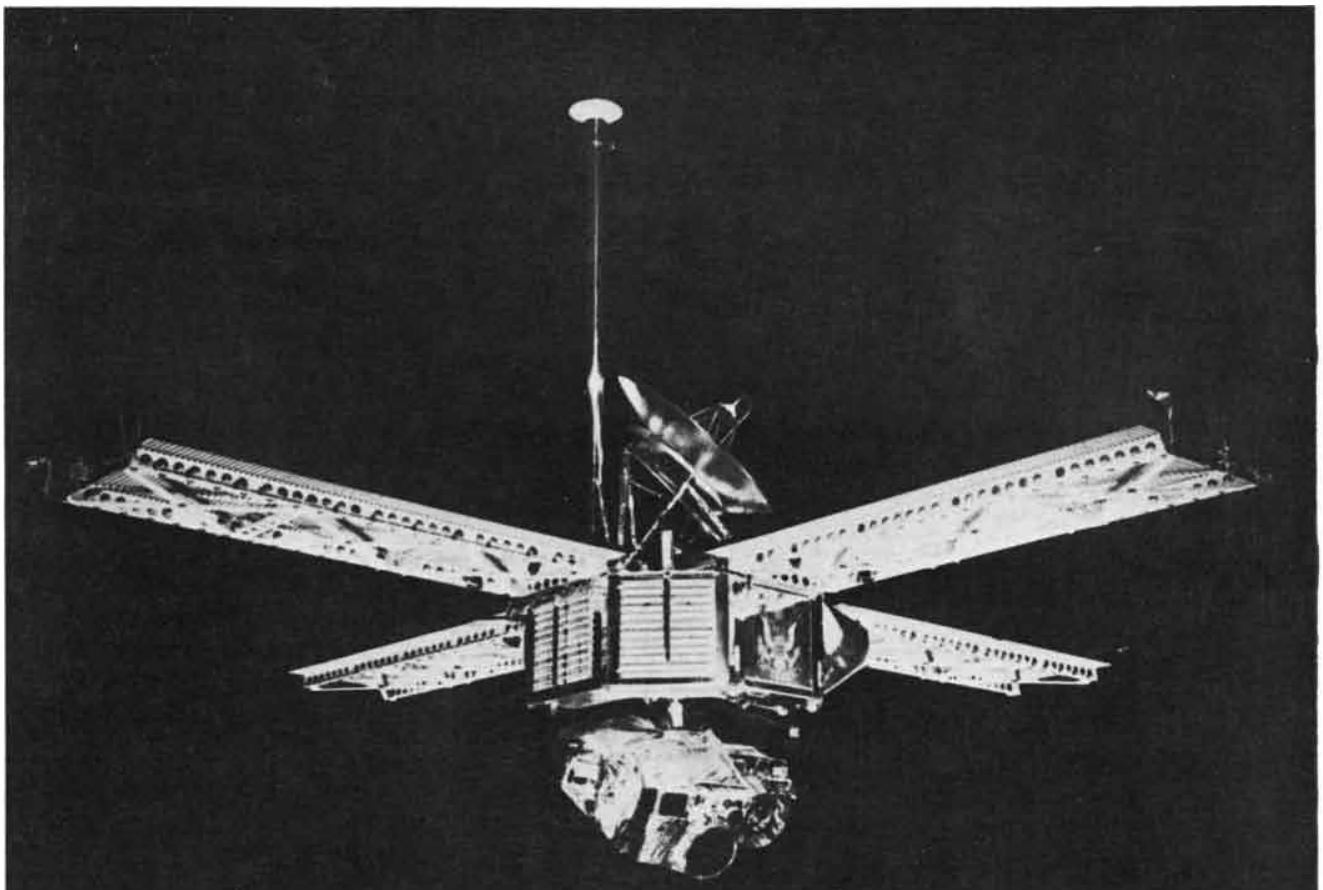
Finally, these studies strengthen the assumption that a prerequisite for studying behavioral modification is understanding the neural circuit underlying the behavior. Once the wiring diagram of the behavior is known, analysis of its modifications is simplified. The same strategy may be applicable to biochemical studies of learning that will eventually describe its mechanisms at a molecular level.

More Complex Behaviors

Studies of habituation of defensive reflexes in the cat and in *Aplysia* span six decades of research, but they represent only a small beginning; we are still far from understanding the neuronal mechanism of long-term memory and of higher learning. It would seem nevertheless that cellular approaches directed toward working out the wiring diagram of behavioral responses can now be applied to more complex learning processes. I have here considered only reflexive behaviors. Animals also display a wide variety of instinctive behaviors, or "fixed-action patterns," that are sometimes triggered by sensory stimuli but that do not require sensory information for the patterning of their behavioral sequence.

For example, in addition to its role in reflex withdrawal, the gill of *Aplysia* participates in a spontaneous withdrawal sequence that has the properties of a fixed-action pattern: it is centrally generated and highly stereotyped, and it occurs in almost identical form in the absence of sensory input in the completely isolated ganglion. Kupfermann, Waziri and I (and more recently Bertram Peretz at the University of Kentucky Medical School) have specified parts of the wiring diagram of this behavior, and it is already apparent that it is different from the reflex withdrawal and considerably more complex; it involves one or more as yet unidentified pattern-generating neurons. An even more complex fixed-action pattern, an escape response consisting of several sequential stereotypically patterned movements, occurs in *Tritonia,* a mollusk closely related to *Aplysia.* A. O. D. Willows of the University of Washington is well on the way toward providing a wiring diagram of even this higher order of instinctive behavior. It may therefore soon be possible to contrast the neural organization of several reflexive and instinctive behaviors in closely related mollusks. If the instinctive behaviors prove to be modifiable, it may then be possible to compare the mechanisms involved in the modification of reflexive and instinctive behaviors.

It would be unrealistic to expect soon a complete set of solutions to as varied and deep a set of problems as learning presents. The pace has quickened, however, and neurobiology may soon provide increasingly complete analyses of progressively more complex instances of learning.



Exploration of Mars... with a little help from Liberty Mirror

Scientists have been studying 200 television pictures of Mars. The pictures were taken by Mariner 6 & 7 space vehicles with cameras equipped with quartz corrector lenses.

Each lens was coated with a thin, electrically conductive, vacuum-deposited film engineered by Liberty Mirror to admit light for picture-taking while keeping infrared heating in the instrument area.

The coating was designed for extreme durability, and to meet the spectral callout requirements of the mission. Also met were Jet Propulsion Laboratory specifications for moisture resistance, cleaning of the coating, and protection against effects of critical chemical atmospheres.

This IR-81-E coating has a number of applications, including defogging or de-icing of optical and instrument faces and windows, prisms, lenses and other optical elements in instruments. Most applications of IR-81-E must be specially engineered. For technical information, contact Liberty Mirror, a Division of Libbey-Owens-Ford Company, Brackenridge, Pa. 15014.

LIBERTY MIRROR

NEGATIVE VISCOSITY

In most fluid systems differences in velocity are obliterated by positive viscosity. In some rotating systems, however, nonuniform flows can be maintained by negative viscous effects due to eddies

by Victor P. Starr and Norman E. Gaut

By vigorously stirring a cup of coffee one can produce a miniature whirlpool, the center of which is rotating more rapidly than the edges. As soon as the stirring stops, the velocity in the center of the cup slows down and the velocity near the periphery tends to increase. None of this is surprising. When faster-moving particles collide with slower-moving ones, part of the momentum of the faster is transferred to the slower. When the particles of different velocity are constituents of a fluid system, we say that the fluid resists shear. We have come to expect that if shear is induced in a fluid—whether by a teaspoon or by the exhaust from a rocket—the shear will inexorably decrease with the passage of time. We do not expect momentum to be transferred from a slower-moving part of a fluid to a faster-moving part, an event that would seem as unlikely as water running uphill. And yet under certain circumstances, which are not at all rare, the amount of shear in a fluid can increase rather than decrease. In contradistinction to molecular viscosity, which is always positive, this behavior of fluids is termed negative viscosity.

The first suggestion that negative viscosity might develop temporarily in a fluid was made in 1895 by Osborne Reynolds, the British physicist who contributed so extensively to the theory of fluids in motion. More recently it has been recognized that negative viscosity can be sustained indefinitely in a wide variety of fluid systems where the motion itself is maintained by a continuous input of energy. Examples are found in the earth's atmosphere (where momentum is funneled into the jet stream that meanders around the globe), in the nonuniform rotation rate of the luminous surface of the sun and in the character of such ocean currents as the Gulf Stream. On a vaster scale there is some evidence that negative viscosity may help to explain why the arms of spiral galaxies tend to wind up in spite of forces that would otherwise keep them extended. The phenomenon of the jet stream and of ocean currents can be duplicated in quite simple laboratory experiments [*see illustration on opposite page*]. In addition, numerical solutions have been obtained for mathematical models that duplicate in varying degrees the negative-viscosity effects observed in natural fluid systems. In short, negative viscous action is now so well demonstrated it should be prominently considered in the study of any fluid system whose behavior seems puzzling or bizarre.

Hydrodynamicists distinguish between two kinds of viscosity. When the viscosity can be ascribed to the interplay among the individual particle constituents of a fluid, it is labeled molecular viscosity. On all larger scales, where viscous effects arise from the interaction of macroscopic "parcels," or masses, of a fluid, one speaks of turbulent, eddy or virtual viscosity.

The explanation of molecular viscosity had its origins in a suggestion by Isaac Newton that in ordinary fluids the drag, or stress, per unit area of one layer of a fluid on its immediate neighbor is proportional to the shear in the fluid. Shear is formally defined as the spatial rate of change in the speed of flow in a direction perpendicular to the mean fluid motion. Subsequently James Clerk Maxwell and others made the assumption that viscous effects originate in molecular motions.

They hypothesized that a fluid resists shear because its molecules do not travel precisely parallel to one another but, as a result of random collisions, are impelled at random angles across the mean direction of flow. In this way faster-moving molecules drift into slower-moving regions and vice versa [*see upper illustration on page 74*]. On balance, the number of molecules moving in one direction across a hypothetical plane dividing two regions of different average flow rates must equal the number moving in the opposite direction. The viscous effect across the plane is associated with the difference in the downstream momentum carried by the molecules moving in opposite directions. Molecules originating in the slower part of the stream tend to carry less of this momentum than those originating in the faster part. As a result the momentum tends to reach an average value and thereby lessens the shear. Maxwell, who had a keen appreciation of such statistical phenomena, computed values for drag due to molecular viscosity and obtained results that were in reasonable agreement with observed values.

Virtual viscosity, the second kind of viscosity, is associated with eddies or turbulence and thus differs fundamentally from molecular viscosity. Virtual viscosity describes the way momentum is transported when the macroscopic masses exchange places on opposite sides of a hypothetical plane that divides two regions of different mean flow rates [*see lower illustration on page 74*]. Such exchanges normally carry far more momentum than the molecular exchanges (which, however, are still present).

The macroscopic exchanges are usually identifiable as currents, vortexes or undulations large enough to be easily recognizable in the motion of the fluid system as a whole. The earth's atmosphere is such a system; the currents and eddies are of the scale familiar in the daily weather map. Often the mean flow (that is, the west-east drift of the atmosphere as a whole) is almost obscured by the day-to-day turbulence.

Hydrodynamicists have had great difficulty in trying to derive equations for turbulent viscosity that would approach the simplicity of the well-established equations for molecular viscosity. One approach has been to ascribe to the eddying masses a "mixing length," corresponding roughly to the mean free path of a molecule. The problem, of course, is that an eddying mass lacks the neat boundaries of a molecule. The second difficulty is that the mixing length cannot correspond precisely to a mean free path simply because the path of an eddy is not free. The eddying mass interacts with others in a continuous manner, so that one cannot escape from continuum mechanics as one can for molecules.

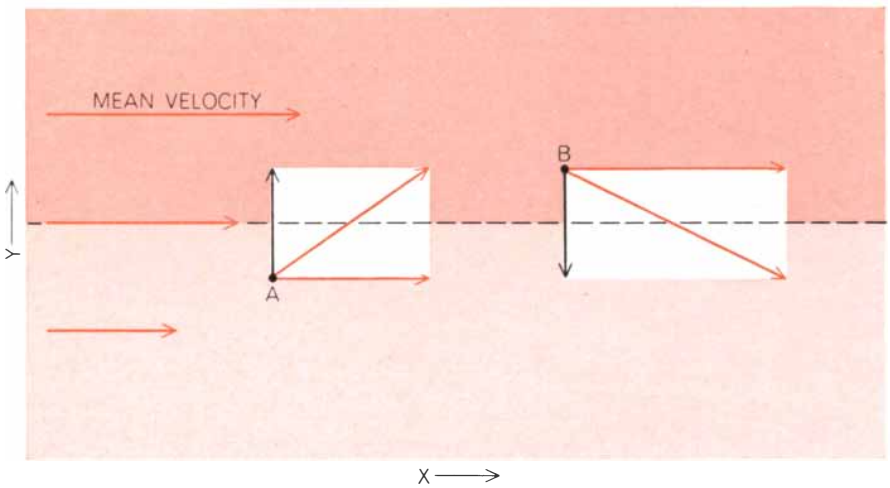
Where the analogy between molecular viscosity and eddy viscosity breaks down most glaringly is in the observation of actual fluid systems. Molecular viscosity is always positive, meaning that any kinetic energy added to the fluid system as a whole will in the end appear in the kinetic energy of the molecules themselves.

Eddy viscosity, however, can be either positive or negative, as Reynolds was the first to perceive. If it is positive, the

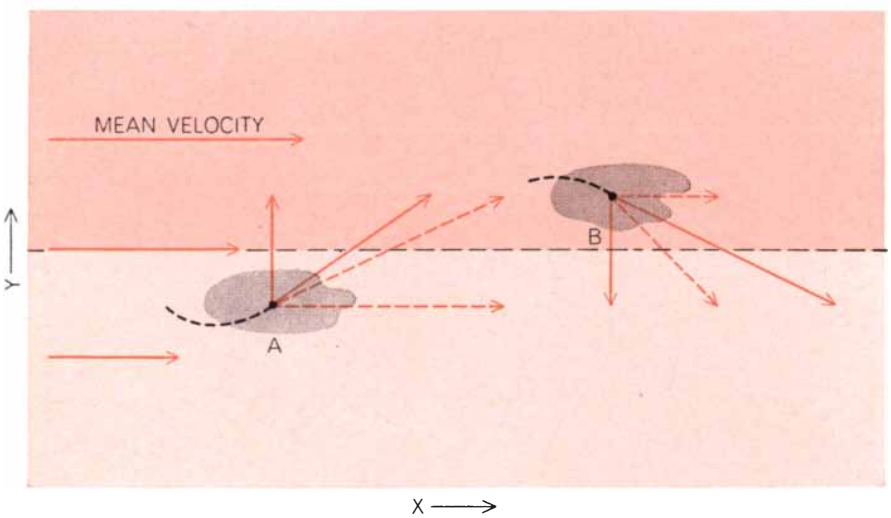
TURBULENCE IN ROTATING "DISHPAN" mimics the flow patterns and transport of momentum observed in the earth's atmosphere. By analyzing local velocities and the structure of the streamlines, made visible by white particles, one can show that angular momentum is being transported away from the periphery of the pan, rotating at a constant rate, to a point nearer the center of rotation. Thus the flow of highest velocity is shifted away from the perimeter of the pan, where it would otherwise be expected. To obtain this result the rim of the pan is being heated (in analogy with the solar heating of the earth's atmosphere) while the center of the pan is being cooled (in analogy with the loss of heat by radiation at the poles). Under these circumstances turbulent flow patterns develop that create negative eddy viscosity: the agent for transporting angular momentum from a slower-moving region to a faster-moving region. The counterpart of this momentum transport in the earth's atmosphere sustains the prevailing westerly winds in middle latitudes. The experiment was photographed in the hydrodynamics laboratory of Dave Fultz at the University of Chicago.

kinetic energy of the eddying mass (like that of a molecule) increases at the expense of the mean kinetic energy of the total system. If the viscosity is negative, however, the eddying mass *loses* some of its energy, which reappears as an increase in the kinetic energy of the limited flow for the total system. It is clear that under conditions where the viscosity is negative the eddy motions will ultimately die out unless there is an input of energy, such as heat, to replenish the kinetic energy that is being lost by the eddies.

How is negative viscosity achieved? How does the transfer of energy from eddy to mean flow take place? Let us imagine fluid flow in a frictionless reentrant channel. Reentrant means that the fluid is flowing in a closed loop so as to preclude the existence of a net pressure force along the channel. Now imagine that in the center of the channel a particular kind of sawtooth flow pattern develops [*see top illustration on opposite page*]. The streamlines in this idealized pattern are parallel to the fluid flow and have equal mass flowing between them. The important characteristic of the pattern is that the rising part of each wave cuts across the middle of the channel at a shallow angle whereas the falling part cuts across virtually at right angles. There is, however, no net transfer of mass from one half of the channel to the other.



POSITIVE VISCOSITY, or resistance to shearing motion, is normally expected in a fluid that contains regions moving at different velocities (*colored arrows at left*). If two molecules, *A* and *B*, traveling at the mean speeds characteristic of their regions of origin were to exchange places, *A* would carry less horizontal momentum upward than *B* carries downward, as shown by the lengths of the attached colored arrows. Countless such random exchanges of momentum actually occur in this fashion and account for positive viscosity.



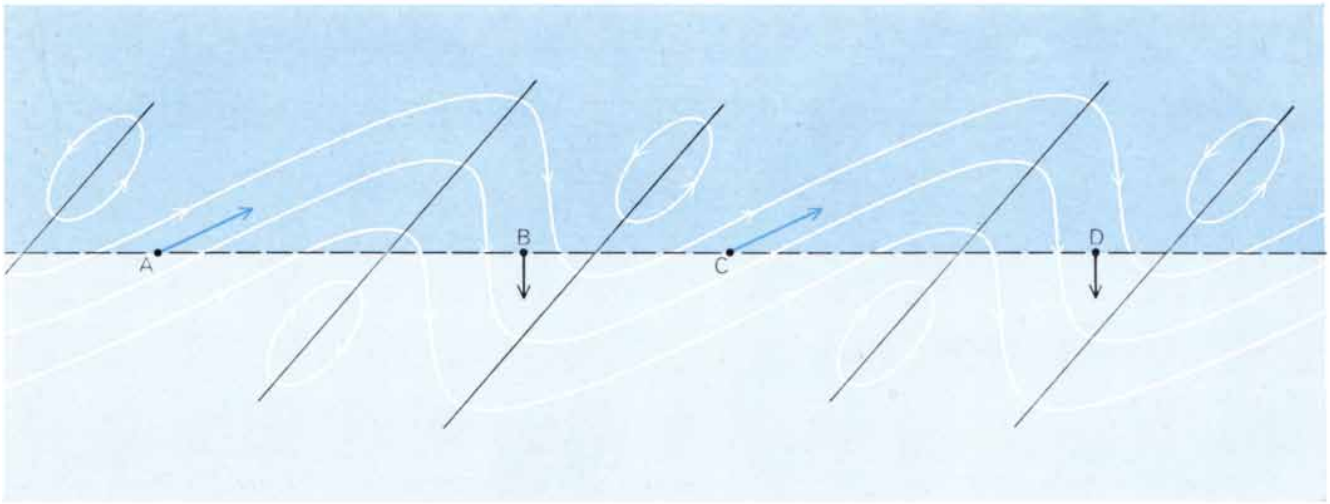
POSITIVE OR NEGATIVE VISCOSITY may result when macroscopic masses of fluid are carried along by eddies. These masses of fluid may act in a much less random way than molecules do. Positive viscosity will again result (*as shown by the attached colored arrows*) if the masses *A* and *B* simply duplicate the behavior of molecules *A* and *B* in the upper illustration on this page. If mass *A*, however, is carried upward with a horizontal momentum not characteristic of its region of origin but larger (*as shown by the broken arrows*), while *B* is carried downward with a horizontal momentum smaller than the momentum characteristic of its region of origin, there can be a net transport of forward momentum from the lower half of the channel to the upper half. This would be an example of negative eddy viscosity, the effect of which is to increase the shear in a fluid. It should be noted that in this case the action is caused by selective nonrandom effect of eddy motions.

Under these circumstances (which can be reproduced in the laboratory) the amount of momentum being carried "northward" greatly exceeds the amount being carried "southward" [*see middle illustration on opposite page*]. Since linear momentum is a conserved quantity in this model, there will be an accumulation of momentum parallel to the channel in the upper half of the system and a corresponding depletion of momentum in the lower half. As a consequence the mean speed of the fluid in the upper half of the channel will tend to increase and the speed in the lower half will tend to decrease. Such a transport of momentum in a direction opposite to the transport due to the shearing force of molecular viscosity, which would ordinarily tend to equalize the flow rate in the two halves of the channel, is what is meant by negative eddy viscosity.

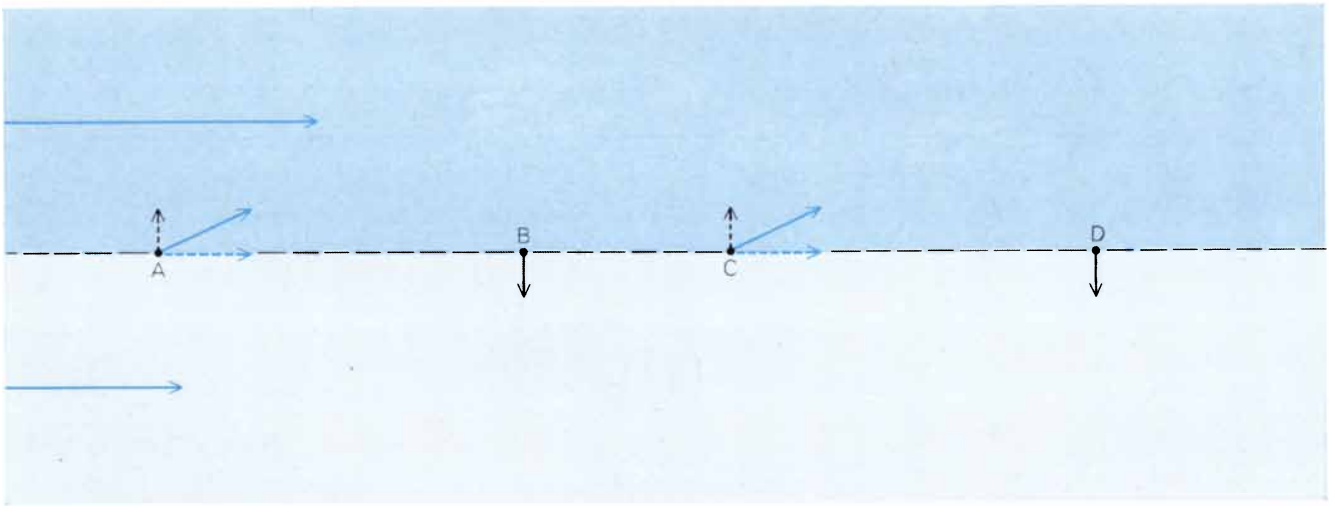
The undulating streamline pattern described in this model is often seen in fluids that exhibit negative viscosity. It is not the only pattern, however, that will produce this result. A flow pattern consisting of similarly tipped ellipses will also transfer momentum from one half of a channel to another [*see bottom illustration on opposite page*]. An interesting feature of this pattern is that the circulation in the ellipses can be reversed without affecting the direction in which momentum is transferred. Ettore Bellomo of the University of Padua has recently developed mathematical models of such elliptical flow patterns that exhibit negative viscosity.

When negative viscosity occurs in natural systems, the flow patterns do not usually have the clearly defined symmetry of the models just described. All that is required is for macroscopic masses of fluid, carried in eddies, to transfer momentum systematically and preferentially from the slower-moving part of a channel to the faster-moving part. The pattern of transfer need not be immediately obvious; it merely has to have a dominant direction in a statistical sense.

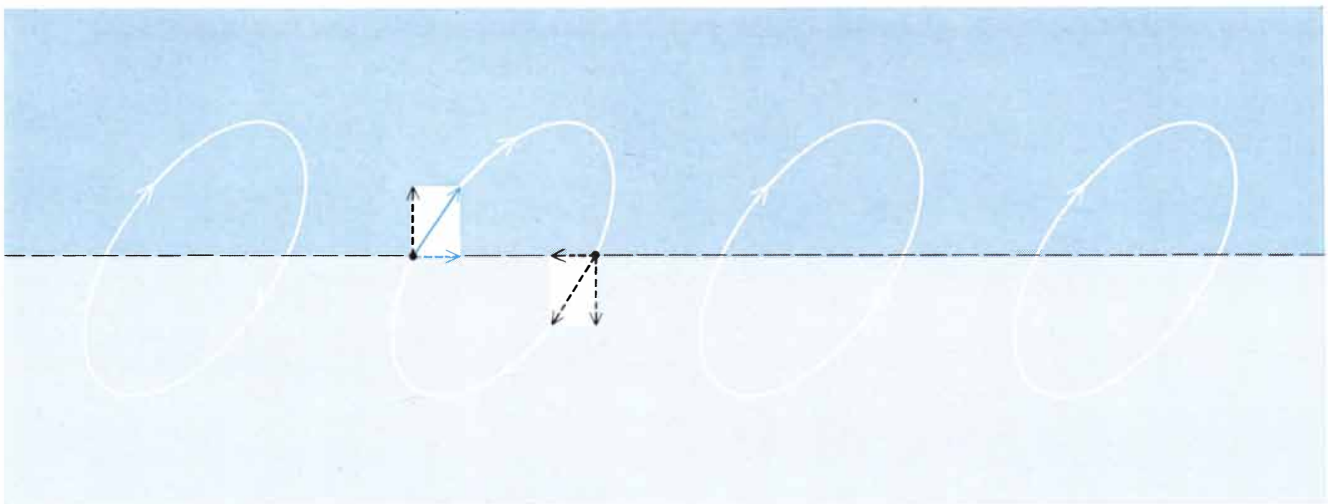
In the channel described above the loss of momentum in the lower half is



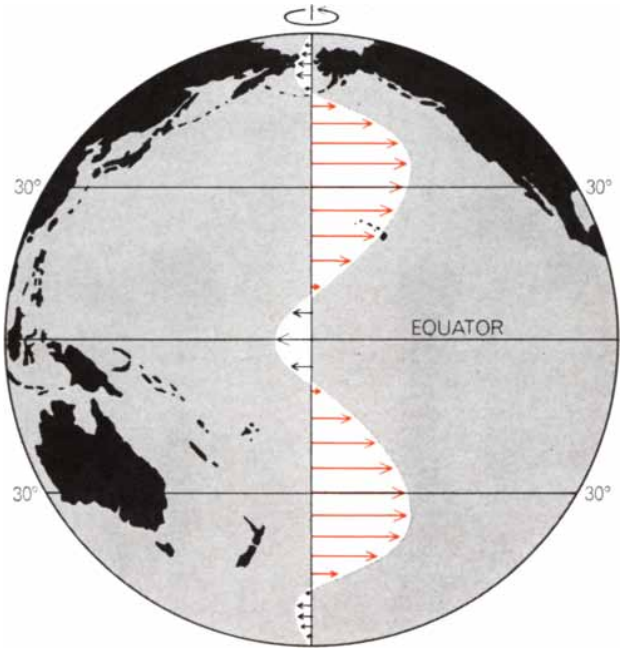
NEGATIVE VISCOUS ACTION is produced by a flow pattern in which the ascending waves (*A*, *C*) tend to follow a shallower angle than the descending waves (*B*, *D*). Such an eddy pattern can develop in a "reentrant" channel, one in which the fluid makes a continuous loop. Momentum transported by the eddies will make the fluid travel faster in the upper half of the channel than in lower.

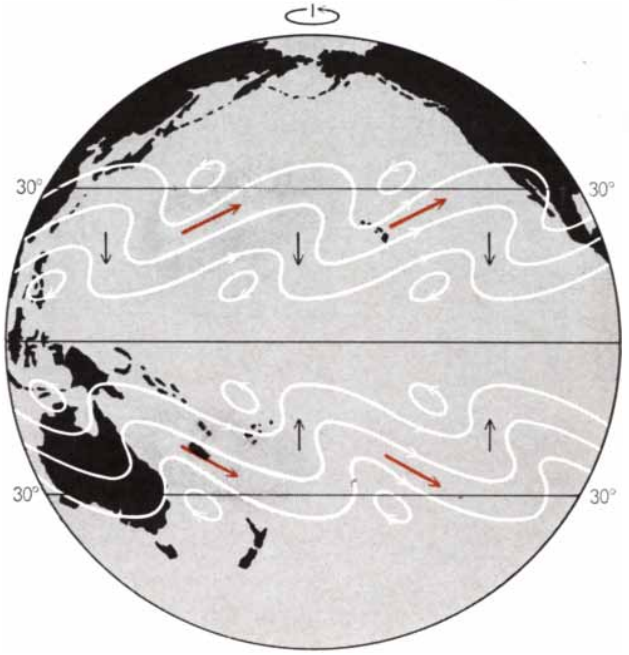


VELOCITY VECTORS drawn for an undulating wave pattern of the kind depicted in the illustration at the top of the page clarify how horizontal momentum is carried upward at *A* and *C* and is not replaced by the masses moving downward at *B* and *D*. As a consequence a negative viscous process is set up, which maintains (or can generate) a shear between upper and lower halves of channel.



TILTED ELLIPTICAL FLOW PATTERNS are also capable of transferring momentum continuously from bottom half of a reentrant channel to top half. Arrows at left show that positive momentum is carried up, and arrows at right show that negative momentum is carried down, so that two actions add up. Ellipses tilted in opposite direction would transfer momentum downward. Amount of negative viscosity generated by such a pattern depends only on tilt of ellipse and the velocity of circulation and not on its direction.

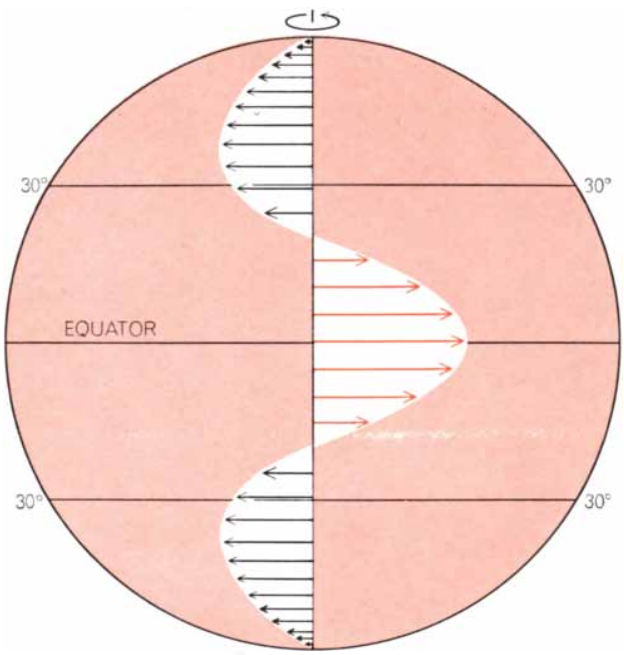


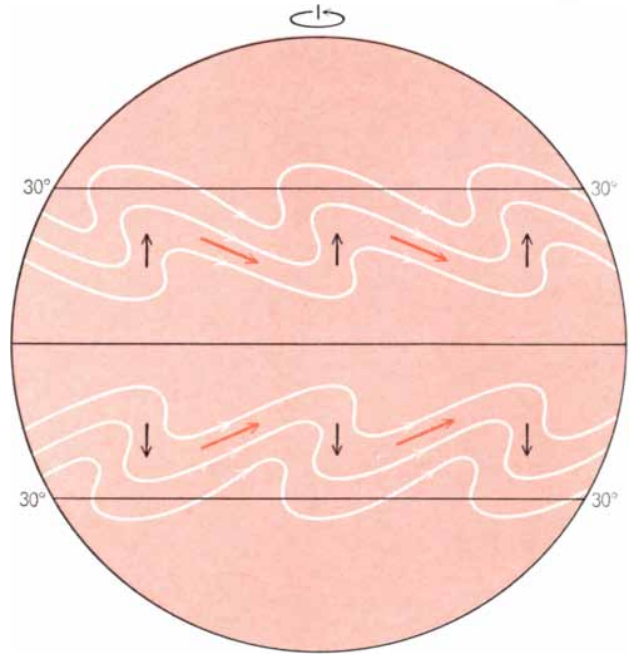


AVERAGE WIND SPEEDS IN THE ATMOSPHERE at a height of about 10 kilometers show a pronounced westerly bulge at mid-latitudes in both hemispheres (*left*), indicative of negative viscous action. Momentum needed to produce these bulges is obtained chiefly from atmospheric circulation near Equator, where prevailing winds are from east. There are also easterlies of lower strength near the two poles. Horizontal momentum transfer that sustains mid-latitude westerlies is provided by turbulent winds in entire lower atmosphere, which tend to follow streamline pattern shown at right. The colored arrows show the direction of momentum transport.

equal to the gain in the upper half. This is represented by equal and opposite changes of the mean flow velocities in the two halves. Since the mean flow kinetic energy involves the squares of the new velocities, however, the total mean flow kinetic energy for the system increases in this process. The only source for the increase is the kinetic energy of the eddy motions.

The natural fluid system that has been most thoroughly studied and that has most clearly exhibited negative-viscosity effects is the earth's atmosphere. The most characteristic feature of atmospheric motions is the presence of an irregular turbulent flow in middle latitudes (in both hemispheres) consisting of eddies in a spectrum of sizes superimposed on an average flow from west to east. These





AVERAGE ROTATIONAL SPEED OF SUN varies with latitude (*left*), showing that the photosphere, or visible surface, does not rotate as a rigid shell. If the rotation rate at about 20 degrees north and south latitude is taken as the mean, regions toward the pole lag behind whereas regions toward the equator rotate more rapidly. This is the reverse of the rotation pattern in the earth's atmosphere, but it too depends on the presence of negative viscosity. A study of sunspot motion strongly suggests a long-term flow pattern like that at right, which is the reverse of the pattern in the earth's atmosphere. Such a pattern would transport momentum toward equator.

eddies account for the familiar and constantly changing patterns seen in the daily weather map.

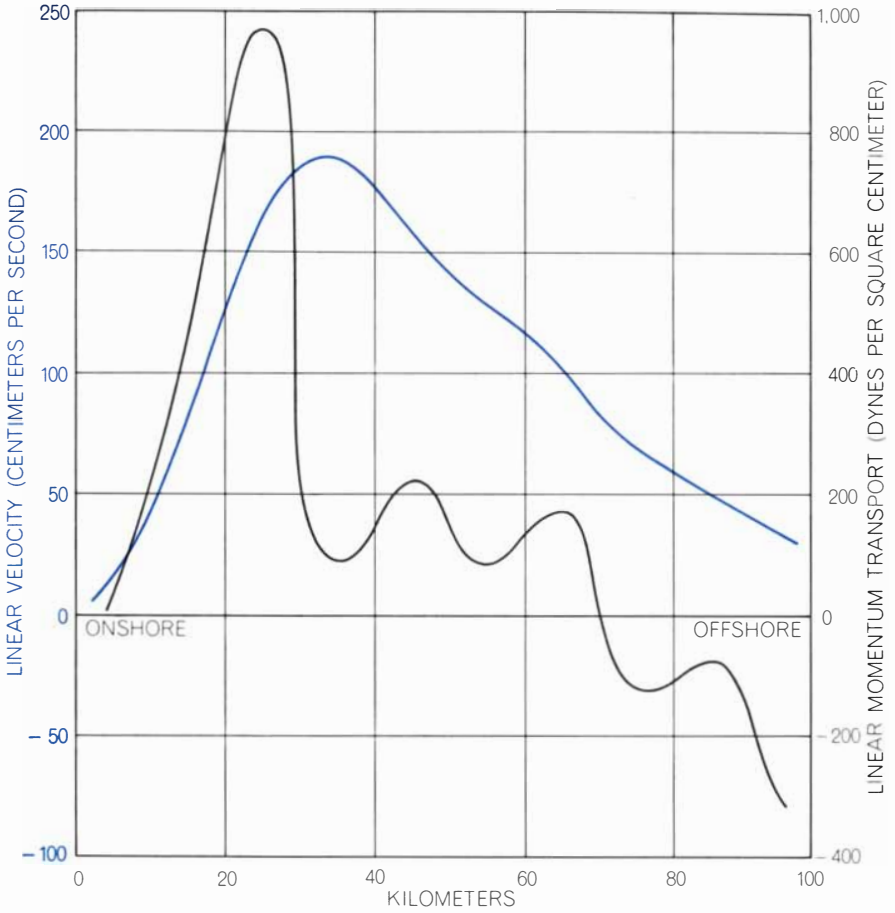
If the wind flow at a height of nine or 10 kilometers is averaged along circles of latitude, it can be seen that the prevailing westerlies build up to a maximum where the jet streams are found, that is, between 30 and 45 degrees latitude in both hemispheres [*see illustration at left at top of opposite page*]. Prevailing easterlies of lower maximum velocity are centered on or near the Equator. There is also evidence for a low-velocity flow from east to west near the poles.

The jet stream itself, which often reaches a velocity of 150 miles per hour at altitudes between 35,000 and 40,000 feet, is the most obvious evidence for negative viscosity in the atmosphere. If there were no mechanisms for channeling momentum into this swift-flowing river of air, it would be quickly slowed down by the frictional effects of positive viscosity feeding momentum vertically into ground and ocean surfaces.

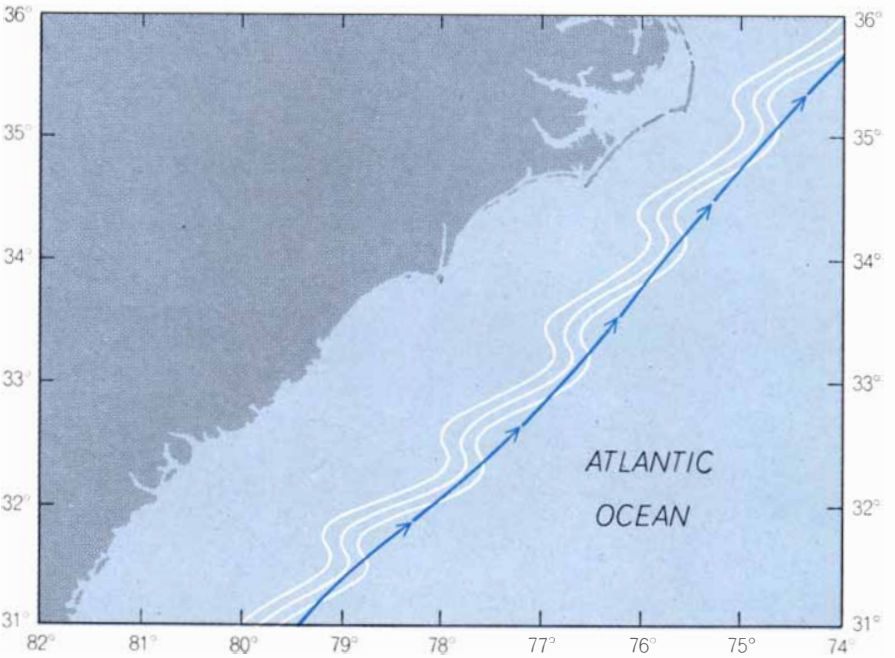
The jet stream and the prevailing westerlies are evidently produced and sustained by the horizontal patterns of turbulence that meteorologists call cyclones and anticyclones. Cyclones (not to be confused with tornadoes or hurricanes) are circular low-pressure systems from several hundred to a few thousand miles in diameter that rotate slowly counterclockwise in the Northern Hemisphere. Anticyclones are the corresponding high-pressure systems that rotate clockwise. These large-scale systems possess the necessary velocity and momentum correlations to create the negative eddy viscosity needed to maintain the westerlies of the middle latitudes.

If the wind streamlines at an altitude of 10 kilometers at these latitudes are plotted and smoothed, they exhibit tilted-trough patterns in each hemisphere capable of transporting momentum from a belt near the Equator to the vicinity of latitude 35 degrees, north and south [*see illustration at right at top of opposite page*]. The kinetic energy of the eddies is derived from large-scale convective processes within the mid-latitude cyclones, our familiar traveling storms. The cyclones in turn derive their energy ultimately from the sun.

A much larger fluid system that exhibits negative viscous effects is the photosphere, or visible surface, of the sun. The evidence is supplied by carefully plotting the movement of sunspots, which were first observed in the 17th century by Galileo. It was subsequently noted that sunspots near the equator



GULF STREAM VELOCITY AND MOMENTUM TRANSPORT are superposed for a cross section of the stream near Cape Hatteras. The peak flow rate (*colored curve*) occurs in the left, or western, half of the northbound current. Dynamic processes within the Gulf Stream transport momentum (*black curve*) into the region of highest velocity from regions still farther to the left, indicating negative viscosity. The curves were made by T. F. Webster of the Woods Hole Oceanographic Institution from ocean-current velocity measurements.



IDEALIZED FLOW PATTERN IN GULF STREAM resembles the wave patterns that produce negative-viscosity effects in the earth's atmosphere and in the sun's photosphere. Momentum is transported away from the shore into the mean current maximum, represented by the colored arrows. The oscillations, or "meanders," in the Gulf Stream have a north-south wavelength of about 100 kilometers and a side-to-side amplitude of about 15 kilometers.

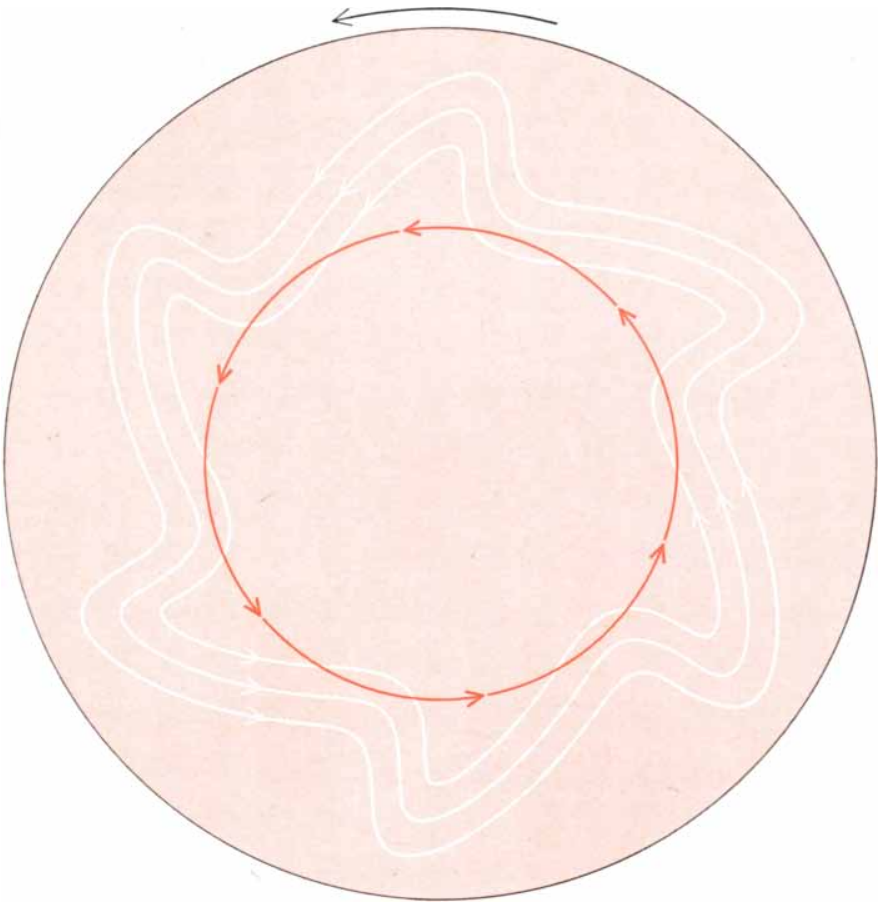
traveled more rapidly across the face of the sun than sunspots farther to the north or south did. This led to recognition that the photosphere does not rotate at a uniform rate but rotates faster at the equator—a phenomenon now known as the equatorial acceleration. A velocity profile of the sun's mean rate of rotation shows a pronounced bulge at the equator and a reverse bulge in each hemisphere [*see illustration at left at bottom of page 76*]. Confirmation of this rotational pattern has come from Doppler measurements of line-of-sight velocities at the edge of the solar disk.

The differences in the physical characteristics of the sun's atmosphere and of the earth's atmosphere are clearly enormous. The circumference of the sun is about 100 times that of the earth. The photosphere is some 400 kilometers thick and ranges in temperature from 4,500 degrees Kelvin at the top, where it joins the overlying chromosphere, to perhaps 8,000 degrees K. at the bottom. Because of the high temperature much of the gas in the photosphere is dissociated into free electrons and free nuclei of atoms, chiefly nuclei of hydrogen. This electrically charged plasma interacts strongly with the sun's magnetic field and produces kinds of turbulence that can be observed only indirectly and that are difficult to treat theoretically.

The main evidence for eddies in the mean circulation of the photosphere comes chiefly from careful observation of sunspot motions. Worldwide observations of sunspots extending back more than 75 years were collected and analyzed under the guidance of Fred Ward at the Air Force Cambridge Research Laboratories. A particular effort was made to correlate the north-south and east-west movement of spots to see if any systematic transport of momentum could be discovered. The analysis showed that enough momentum was indeed being carried toward the equator to produce the higher rotation rate of the sun's equatorial region.

There is theoretical work suggesting that patterns in the large-scale magnetic features of the sun should correlate with fluid motions on the same scale. These studies, together with experiments on numerical models, indicate that fluid flow in the solar atmosphere should exhibit, at least statistically, streamline patterns that are mirror images of the patterns present in the earth's atmosphere [*see illustration at right at bottom of page 76*]. Because the sun's patterns show tilts that are the reverse of those in the earth's atmosphere, momentum on the sun is carried toward the equator rather than away from it, thus providing another example of negative eddy viscosity.

The ultimate source of the energy that sustains the transfer of momentum to the equator is the sun's internal nuclear furnace. We do not know if there is a temperature difference between the sun's poles and the equator. If there is, it could drive the solar circulation much as the earth's atmosphere is driven by the pole-to-equator temperature gradient. Other mechanisms are conceivable, however, for feeding energy into the eddies. It also seems probable that the sun's magnetic field plays an important role in restraining large-scale motions.

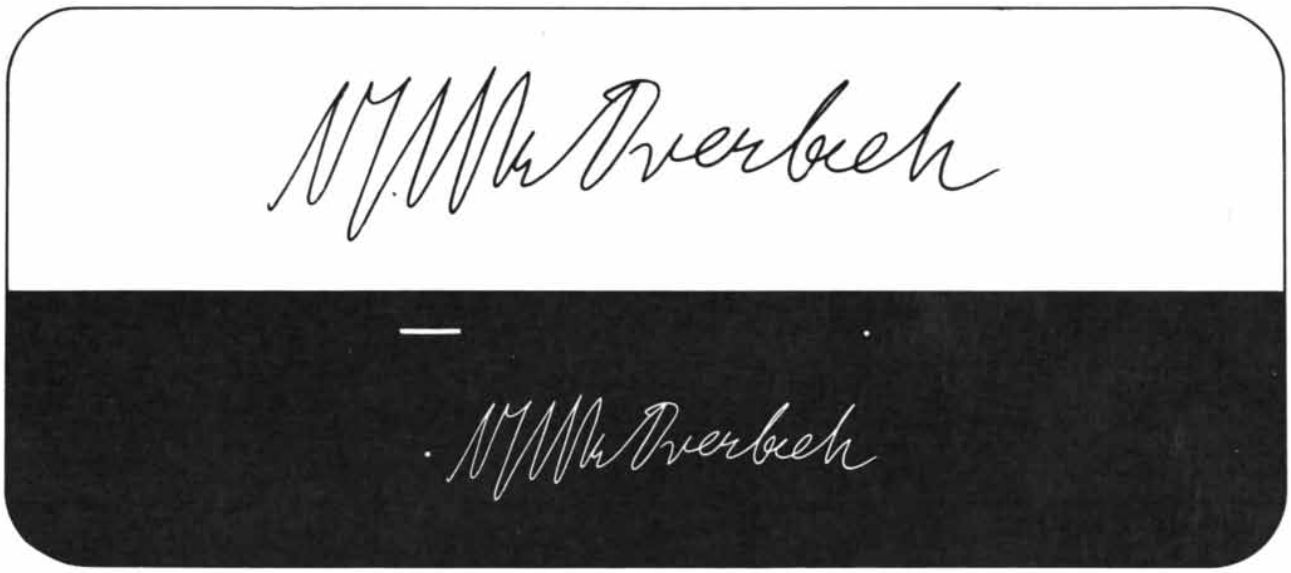


STREAMLINES IN ROTATING DISHPAN, similar to the one in photograph on page 73, have configuration needed to create negative viscosity. Eddies in flow transport momentum away from rim to zone of maximum rotation velocity (*colored circle*). Actual streamline patterns are usually more irregular, but over a period they have the requisite properties.

The Gulf Stream is a third, and again quite different, example of a fluid system that exhibits negative-viscosity effects. The Gulf Stream flows northward from 30 to 160 kilometers off the Atlantic coast of the U.S. from the Florida Strait to Cape Hatteras, where it turns more seaward. The stream axis is generally parallel to the shoreline over this distance with a maximum surface speed of about two meters per second, or about four miles per hour. If one plots the average transport of momentum, one sees that momentum is carried from the left (shoreward) side of the stream into the region where its velocity is a maximum [*see top illustration on preceding page*]. This is again clear evidence of negative eddy viscosity.

The eddies in the Gulf Stream take the form of irregular but semiperiodic oscillations known as meanders. These excursions in the mean flow have time scales of a few days when observed from a fixed reference point. This implies a north-south wavelength on the order of 100 kilometers and a side-to-side amplitude of about 15 kilometers. An idealized meander consists of a brisk current running offshore for perhaps three or four days, followed by a slower, less organized flow moving onshore for two or three days, after which the cycle is repeated [*see bottom illustration on preceding page*]. The whole pattern has the tilted troughs and ridges that bespeak the presence of negative eddy viscosity.

The transformations of energy that



Checking your signature by telephone

Most of the information transmitted in a television picture is non-essential. This is not a damning comment on the quality of your favourite programme, but a technical one. Television pictures are reproduced cinematographically. This calls for a far wider bandwidth than is actually necessary to transmit essential information. If the picture was first "translated" into symbols, the bandwidth could be reduced appreciably and still convey the message.

What, you may ask, has this to do with checking signatures by telephone?

In the course of work on methods of classifying hand-written characters, Mr. A. J. W. M. van Overbeek and Mr. L. G. Krul of Philips Research Laboratories, Eindhoven, the Netherlands, experimented with the picture-to-symbol translation technique. They observed that the vertical line is rarely crossed more than four times in handwritten characters. Consequently words or signatures could be transformed into four continuous or discontinuous curved traces, running roughly horizontally. In order to transmit one trace, all you need to know is its height above a fixed horizontal baseline, as a function of horizontal distance. Thus it can be reproduced on an oscilloscope with a linear time base, by feeding the vertical deflection amplifier with a voltage representing this height. If 4 traces are written one after the other, the 4 "pictures" together display the signature.

On the pick-up side, the problem was how to derive the trace signals. Van Overbeek and Krul solved this by using TV techniques and some logic circuitry. A TV camera, with its scanning lines vertical scans the picture. A sawtooth voltage generator runs synchronously with the line deflection. During a first raster scan, the lowest trace is produced by sampling the voltage of the sawtooth generator at the instant that a TV scanning line crosses a black line. The combined samples obtained during the raster scan form the desired lowest trace signal. During subsequent TV scanning rasters, the second, third and fourth traces are obtained by sampling the sawtooth voltage at the corresponding points where the black line is crossed. Employing somewhat different logic, the first two and the last two traces can be obtained, instead of the first four traces. This may be better for more complicated signatures. For practical purposes the equipment used in the experiments was based on European TV standards. This means that $12\frac{1}{2}$ complete signatures, with 312 samples per trace, are given every second, i.e. a bandwidth of 7.8 kHz is used. By omitting every second or second and third sample the bandwidth can be further reduced, and still produce recognizable signatures. The signals described can easily be sent over direct lines, where the d.c. component is retained. For a telephone network a subsequent modulation is, of course, required.

A feature of this method of transmitting signatures is that the beam really writes the curves, making them appear more natural than those of an ordinary slow-scan-television system with a small number of scanning lines.

At present, signature transmission is only feasible for local area calls. For distances of, say, 300 miles, special modulation techniques still have to be developed and the problems of cross-talk and phase distortion overcome.

In the Research Laboratories of the Philips group of companies, scientists work together in many fields of science. Among these are: Acoustics, Cryogenics, Information Processing, Mechanics, Nuclear Physics, Perception, Solid State, Telecommunication and Television.

PHILIPS

PHILIPS

	EARTH'S ATMOSPHERE	SOLAR ATMOSPHERE	GULF STREAM	LABORATORY EXPERIMENTS	GALAXIES
SIZE (KILOMETERS)	1.3×10^4	1.4×10^6	1×10^3	5×10^{-4}	1×10^{17}
IMPORTANT FORCES	HYDRODYNAMIC	HYDRODYNAMIC, MAGNETIC	HYDRODYNAMIC	HYDRODYNAMIC	MUTUAL GRAVITATION, HYDRODYNAMIC
TYPE OF FLUID	GASEOUS	PLASMA	LIQUID	LIQUID	GAS, STARS AND DUST
GEOMETRY	SPHERICAL SHELL	SPHERICAL SHELL	APPROXIMATE PLANE	PLANE	PLANE
APPROXIMATE PERIOD OF EDDY MOTION IN RELATIVE COORDINATES	3-5 DAYS	100 DAYS	3-5 DAYS	~1 MINUTE	10^8 YEARS
SOURCE OF ENERGY	SOLAR RADIATION	NUCLEAR	WINDS, RADIATION	ELECTRIC HEATING	NUCLEAR, GRAVITATIONAL
DISSIPATION PROCESSES	MOLECULAR FRICTION	MOLECULAR FRICTION, JOULE HEATING	MOLECULAR FRICTION	MOLECULAR FRICTION	FRICTION?
TYPICAL SPEEDS OF PARTICLES IN RELATIVE COORDINATES (CENTIMETERS PER SECOND)	10^3	10^5	10^2	1	10^6

SYSTEMS EXHIBITING NEGATIVE VISCOSITY range in size from laboratory dishpans to galaxies. The evidence for negative viscosity in the latter, however, is still tentative. The "fluids" range from ordinary water and air to stellar and galactic materials. The common characteristic of all these systems is rotation. The necessity of rotation for negative viscosity has not been proved, however.

sustain the meanders in the Gulf Stream are not well understood. Some light has been shed on this subject recently by numerical models of Gulf Stream dynamics. Isidoro Orlanski of the Geophysical Fluid Dynamics Laboratory operated by the Environmental Science Services Administration has shown that the meanders, or eddies, convert heat energy into kinetic energy somewhere below the ocean surface. Thus the eddies achieve an energy conversion similar to the conversion achieved by the eddies that maintain the general circulation of the earth's atmosphere.

Another class of fluid systems exhibiting negative viscosity can be produced quite simply in the laboratory. Some of the earliest analogues of atmospheric flow were devised by rotating dishpans of water. To simulate solar heating at the equator and radiation cooling at the poles, the dishpans were heated at the perimeter and cooled at the center. A wide variety of regimes of flow can be obtained by varying the amount of differential heating and the rate of rotation. The flow patterns that develop are made visible by introducing tracer substances, such as colored flakes of dye or powdered aluminum. Streamlines similar to those found in the earth's atmosphere, which are needed for negative viscosity, are readily produced in the experiments.

The irregular eddies that continuously form and disappear have properties and structures amazingly like the mid-latitude cyclones and anticyclones in the earth's lower atmosphere. The storm analogues move northeastward toward the center of the dishpan (if it is rotating counterclockwise, in analogy with the earth) and are dissipated in a highly realistic manner. Finally, structure in the earth's upper atmosphere near the jet stream is also reproduced in dishpan experiments. Wave forms with troughs and ridges appear that have the characteristic tilt needed to transport momentum from regions of low angular speeds near the rim (equator) inward to regions where angular speeds are higher. The phenomenon of negative eddy viscosity is unmistakably present [*see illustration on page 73*].

These four examples of negative viscosity—in the earth's atmosphere, in the sun's photosphere, in the Gulf Stream and in laboratory dishpans—are well verified by direct observation. There are other examples, less obvious but highly plausible. For instance, the hydrodynamic forces that account for the distribution of stars in spiral galaxies must be a good deal more exotic than the simple forces one can demonstrate by pouring a thin jet of cream into a swirling cup of coffee. (This analogy is sometimes presented in popular accounts of astronomy.) Analytical and numerical studies performed at the Massachusetts Institute of Technology have shown that the peculiar spiral distribution of matter—consisting of stars, dust and gas—in a spiral galaxy leads to somewhat unexpected velocity fields. Matter flows through the spiral arms. On closer examination it appears that the material flowing through them acquires a tilted wavelike streamline pattern of the type we have learned to associate with negative eddy viscosity. The negative viscosity evidently transports angular momentum from the more slowly rotating outer parts of the galaxy to the more rapidly rotating central regions.

Negative-viscosity effects can also be invoked to help explain the unexpectedly slow measured rotation rates of certain stars. It has been suggested that they have a core that is rotating much faster than their visible surface layers. Conceivably these stars have a deep outer shell in which special convection patterns could give rise to negative eddy viscosity and so transport angular momentum from the exterior part of the star inward.

The table above lists the wide diversity of fluid systems that clearly exhibit effects attributable to negative eddy viscosity. The table brings out the extremes in size, the differences in geometry, the variety of fluids, the range of forces and the scale of time over which the effects appear. The forces at work in these systems are hydrodynamic, magnetohydrodynamic and gravitational; the materials present in the fluids include gases, liquids, plasmas, stars and interstellar dust. The one feature common to all these systems is rotation. It might seem, therefore, that rotation is required for the appearance of negative eddy viscosity. Although this seems plausible, theory has not gone far enough to prove that rotation is indispensable.

Unite your computers.

Unite every computer in your business, and put them to work where they deliver the most profit. *Solving management problems.* Working together as one, powerful, companywide system through General Electric United Systems.

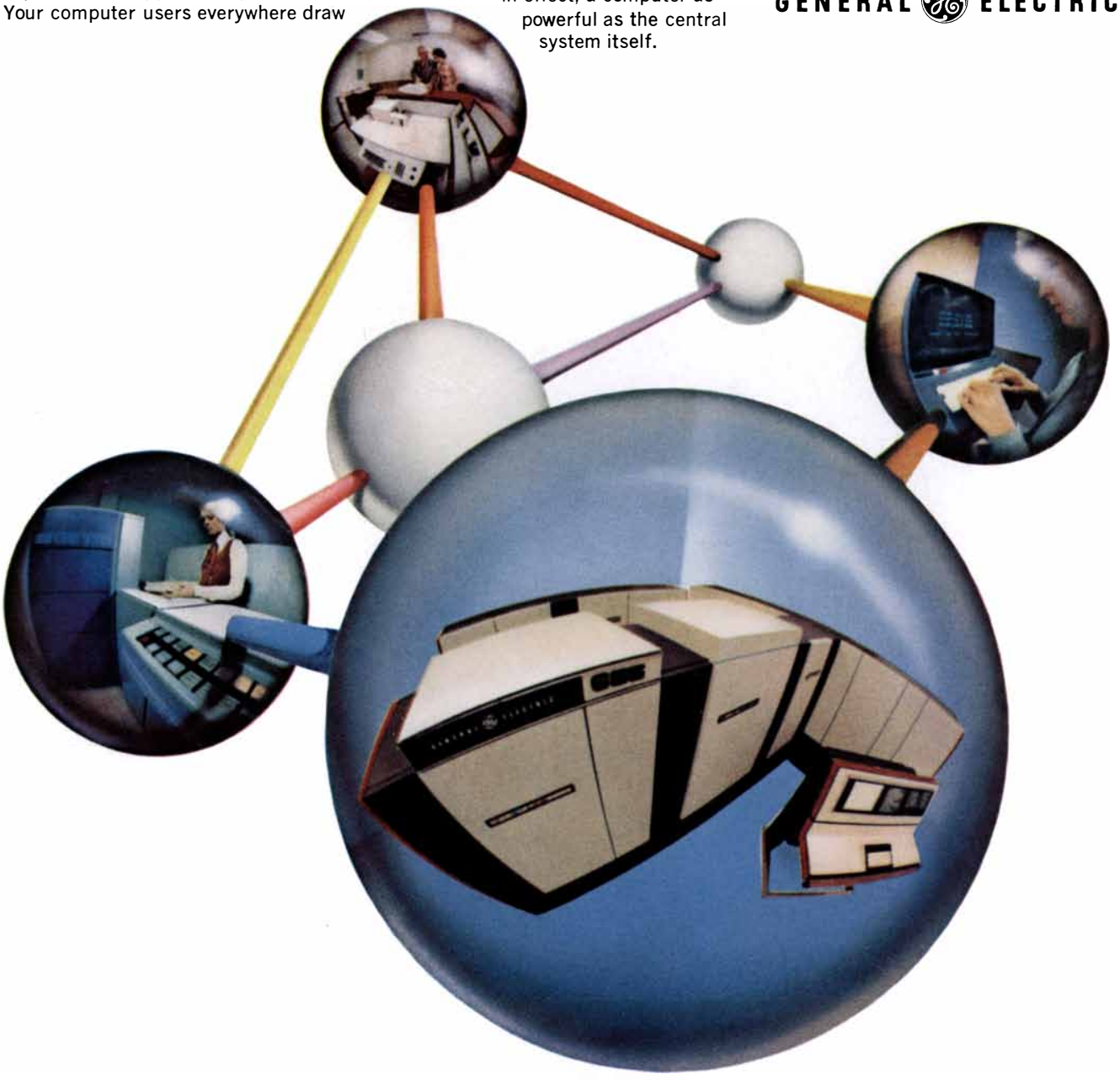
The management information systems of the 70's.

You control the system . . . establish its objectives . . . approve its priorities. Your computer users everywhere draw from a single data base of information.

No longer will remote plants and offices harbor "colonies of computers" to serve only their specialized needs. No longer will they duplicate each other's computer efforts. Each operation manager will contribute data to the central system . . . and draw whatever information and computer power he needs, when he needs it. Each will have, in effect, a computer as powerful as the central system itself.

You can begin *now* to unite your computers, your people and your business by calling your nearest GE Information Systems Sales Representative. Or write to Section 290-76, General Electric Co., 1 River Road, Schenectady, N.Y. 12305.

GENERAL ELECTRIC



General Electric Computers: The United Systems of the 70's

Bernard of Treves' recipe for gold: Start with 2000 eggs...

Bernard of Treves, Comte de la Marche, was a 15th century Italian alchemist with a bent for squandering money. Left a large fortune, he blew it searching for the "Philosopher's Stone"—a handy item for changing dross metal into silver and gold.

Among the ingredients Bernard experimented with were brandy, sea water, toads, alum and rutabagas. And once, he tells us, he came pretty close with a mixture of 2,000 eggs and dry water (uh-huh, dry water).

Now whether Bernard ever actually produced gold is academic.

The Dow Chemical Company, Midland, Michigan 48640.

But he was an experimenter, working with common things to uncork the uncommon. And that's what we do, too.

From inoculated eggs we produce life-saving biologicals, including the most widely used live-virus measles vaccine.

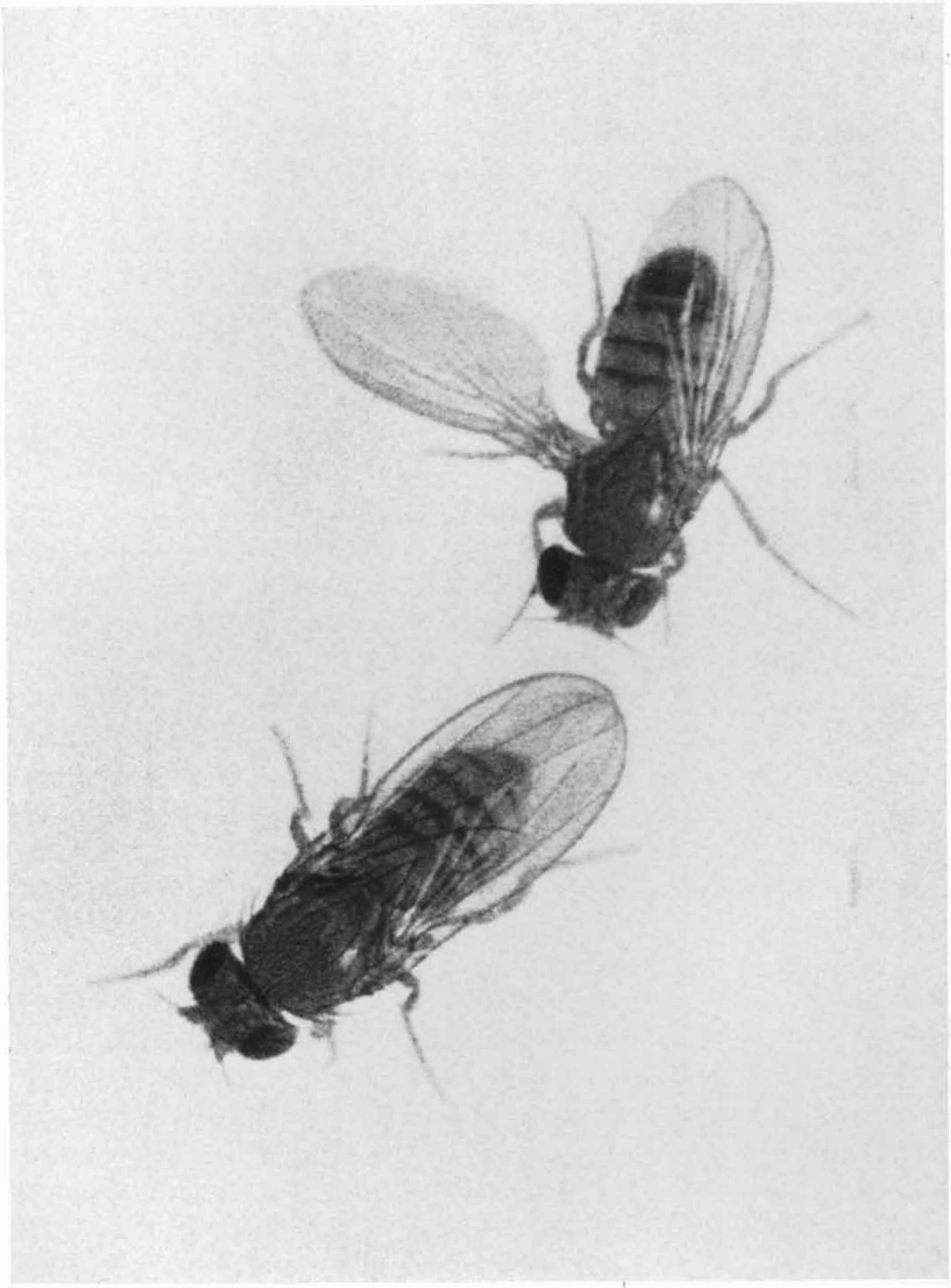
From common salts we make caustic in its new beaded form, brominated fire retardants, chlorine for cleaning, bleaching and disinfecting.

From hydrocarbons we process automotive chemicals, aspirin, antimicrobials and the versatile new ethylenimine.

Chemistry has come a long way since men sought precious metals in giant omelets. Man's needs may change, but in our view, one thing remains constant.

The search to satisfy them.

DOW



SERENADING MALE FLY stands close to the female and, with right wing extended, prepares to go through the motions that make a sound characteristic of all the males of his species. If the female fly belongs to another species, she will be unreceptive to his song.

The Love Song of the Fruit Fly

Most species of Drosophila have overlapping ranges, yet hybrid fruit flies are rare. One reason is that in courtship the females respond to a unique signal that is emitted only by males of the same species

by H. C. Bennet-Clark and A. W. Ewing

Unlike the strident proclamations of the cricket and the cicada, the love song of the tiny fruit fly is a quiet and private affair. The courting male approaches to within two millimeters of the female and engages in a wing display. He extends one wing or both and assumes a posture that is identical among all males of the species; the wing vibrations that follow generate a characteristic sound. The female senses this sound through her aristae: feathery projections of the antennae. If the female is receptive, mating ensues. If she is unreceptive (which is to say that she is immature, already mated or of a different species), she produces her own song, a buzz with a fundamental frequency near 300 cycles per second. The sound causes the male to turn away and stop courting. Recording the love songs of fruit flies in our laboratory at the University of Edinburgh, we have found that, whereas the female's repulsion signal is essentially the same among all the species we have observed, the male's song differs from species to species.

The genus *Drosophila* includes some 2,000 species of flies. The flies are not easy to tell apart. Some are light and some dark, some are short-winged and some long-winged, but even viewed side by side they all look rather alike. Some species are cosmopolitan; *D. melanogaster, D. repleta* and *D. simulans,* for example, are found on all continents. Few species have a geographical range that is not shared with several other species. (In the Hawaiian Islands alone there are more than 1,000 species.) The various species have nonetheless managed to remain genetically isolated. A few species can be interbred in the laboratory, but the offspring are often partially or wholly sterile. All of this suggests that the unique love song of each species constitutes a powerful isolation mechanism that acts as a barrier to interbreeding and thus maintains the genetic integrity of each species.

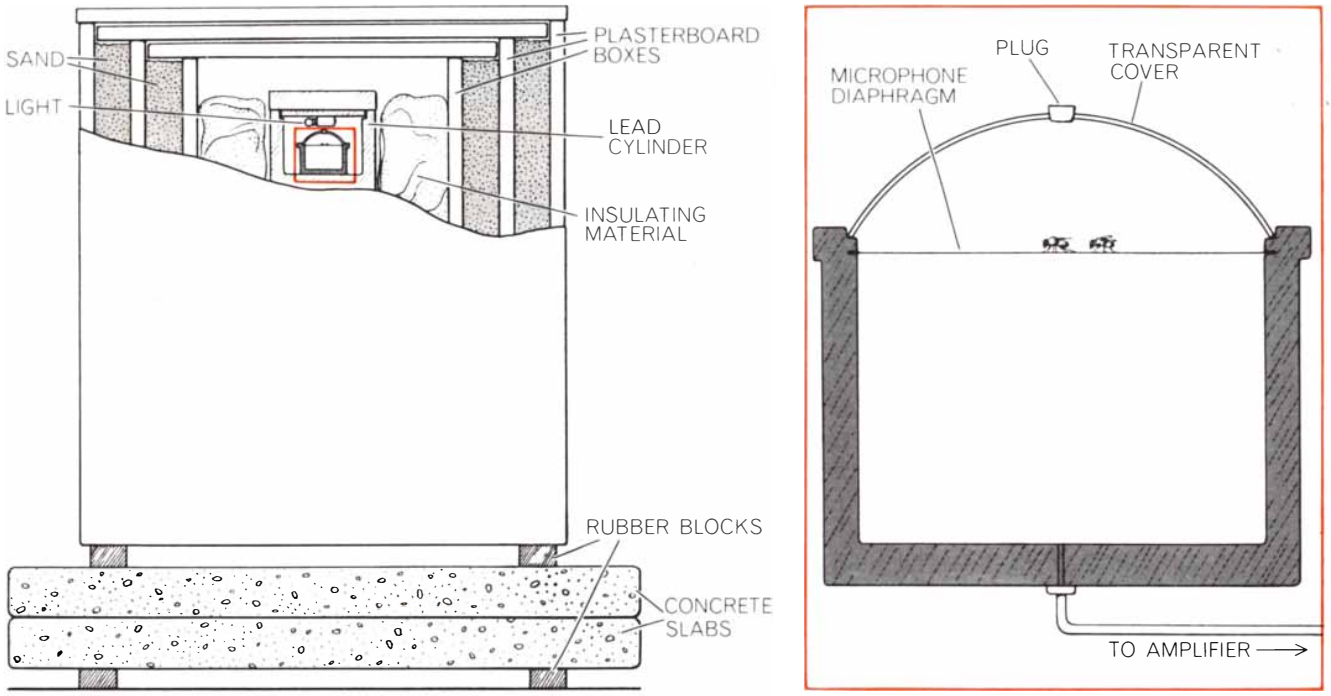
As is well known, geneticists have long favored *Drosophila* as an experimental animal. The flies are easy to maintain, their life cycle is short and comparisons can readily be made between changes in the morphology or biochemistry of the individual fly on the one hand and changes in its chromosomes on the other. Thanks both to these studies and to taxonomic investigations, so much is now known about the evolutionary history of the genus that fruit flies are becoming important for studies of the genetics underlying the evolution of behavior. We have sought answers to two such questions: How did the love song of the fruit fly evolve? What is the significance of the song?

It is not hard to demonstrate the key role that the love song plays in mating. Remove the male fly's wings and he will court with the same persistence as before, but his courtship is seldom successful. It is apparent that the male's wing display, at least, is a prerequisite to mating. Indeed, one species (*D. obscura*) sings no song and courts only by means of a silent wing display. This, however, is one of the few fruit fly species that will not mate in the dark. Since most species breed successfully at night, one must conclude that visual display alone is not enough to win female acceptance. The importance of sound to the female is also easy to demonstrate. Our colleague A. W. G. Manning has shown that when the antennae of a sexually responsive female are immobilized with glue, she ceases to be receptive.

Sound appears to be the major element in mating. Contact chemical stimuli are also evident. In the species *D. melanogaster,* for example, the male taps the female with his forelegs to "taste" her before beginning his wing display. If tasting reveals that the female is of another species, the male normally breaks off contact. The specific scent of the female does not develop, however, until about two days after she reaches the adult stage of development. Before then males will court foreign females, and differences in song may be the only clue for discriminating between species. Moreover, the male tastes the female only at the onset of courtship. Once the male is sexually aroused, he may, in a mixed population of flies, switch his attention from a female of his own species to a female of another. In such circumstances the female would again be able to distinguish a male of the same species from an alien suitor by means of the love song.

Our first step in studying the fruit flies' love songs was to record the sounds produced by males of various species. This presented difficulties because the song is so quiet. We have calculated (on the basis of amplitude of wing movement and the thrust produced in free flight) that in terms of power output the signal does not exceed a ten-billionth of a watt. This minuscule output can easily be drowned out by other sounds in the environment. In order to record the songs we had to install a sensitive crystal microphone, which fed into a low-noise amplifier, in an elaborately soundproofed box. We then placed the flies directly on the microphone diaphragm. Even under these circumstances we found it best to record the songs in the middle of the night or on the weekend, when the rest of the laboratory was quiet.

We discovered that a song usually



EXPERIMENTAL ENCLOSURE for recording the fruit flies' songs consists of a series of nesting boxes, each insulated from its neighbor by sand (*left*). The innermost box contains a cylindrical lead casting that houses a microphone and a small illuminating lamp. The female fly and the courting male are confined in a small cell that is placed directly on the diaphragm of the microphone (*right*).

consists of a train of sound pulses, with each short pulse normally followed by a period of silence. The intervals are usually the same throughout the song, but the pattern of sound and silence varies considerably from species to species. At one extreme is the song pulse of *D. melanogaster,* which consists of a single cycle that lasts only three thousandths of a second. At the other extreme is the pulse of *D. micromelanica:* a continuous buzz that may last as long as 10 seconds. Between these extremes are song pulses that continue for two to four cycles (*D. ananassae* and *D. athabasca*) or for three to seven cycles (*D. pseudoobscura* and *D. persimilis*).

Of the 22 species whose songs we have recorded, six had song pulses that consisted of a single cycle: *D. melanogaster, D. paulistorum* and four others. In all the single-cycle singers the pulse is of about the same duration, three thousandths of a second, and the tone is roughly 330 cycles per second. The flies with multiple-cycle songs, however, produce sound pulses with a variety of tones. The highest tone is generated by *D. bipectinata* and has a frequency of 575 c.p.s. The tone of *D. athabasca* is 440 c.p.s. and the tone of *D. algonquin* is 225 c.p.s.

The love songs of the fruit flies that are continuous buzzers are not monotonous, to use that word in its pure sense. The song of *D. paramelanica,* for example, is one of those that continue for several seconds. For the first tenth of a second the tone of the buzz is 440 c.p.s. The tone then shifts to 265 c.p.s. for a slightly longer interval; the two tones are then repeated at the same intervals until the song ends. The 10-second song of *D. micromelanica,* another continuous buzzer, is basically a soft 440 c.p.s., combined with an intermittent louder and lower 385 c.p.s.

In addition to possessing variations in duration and tone, the love songs of fruit flies also vary with respect to the interval between sound pulses. The briefest interval, eight thousandths of a second, is found in *D. bipectinata. D. melanogaster* repeats its song 30 times a second, and its near relative *D. simulans* does so 20 times a second. At the far end of the range is *D. melanica;* nearly four tenths of a second may pass before it repeats its song.

With such a wide range of variation the possible number of fruit fly songs is very large. We could not judge which of the many variables were significant in terms of signal content without acquiring some knowledge of the mechanisms that produce the songs and of the physiology of the female receptor organ. We began by examining the mechanics of song production. We had already recorded the buzz produced by *D. melanogaster* in flight, and we also knew that the shape of the sound pulses produced by the courting fly was similar to the shape of a portion of the flight buzz. This suggested to us how the love song might be generated.

By making a series of multiple-flash photographs of *D. melanogaster* in flight we were able to establish its normal wingbeat cycle. We next used the sounds produced in flight to trigger the multiple flashes. The resulting photographs showed that the flight sound resembling the love song is produced by the upstroke of the wing, whereas the downstroke is fairly noiseless.

We then made multiple-flash photographs of a male during courtship, using a prearranged time interval between flashes. These photographs showed that the courting male lifts his extended wing for a single beat about every three hundredths of a second; the upstroke lasts for only three thousandths of a second, whereas the downstroke lasts for more than four thousandths of a second. Since our courtship recordings had already revealed that the love song of *D. melanogaster* was a single pulse with a duration of three thousandths of a second, it was probable that the wing upstroke in courtship produces the song pulse just as the same upstroke in flight produces the characteristic buzz. The activation of

the system responsible for the sound, however, differs in the two instances. In flight it is a resonant oscillation of the fly's thoracic mechanism that produces the continuous train of sound pulses. In courtship the pulses are not produced continuously but intermittently.

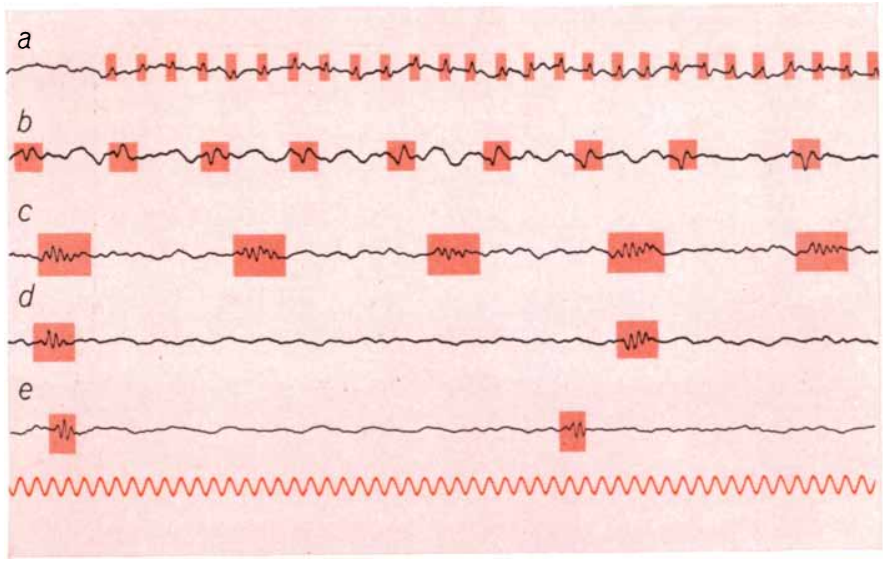
Using the same technique, we went on to examine the courtship wingbeat of *D. persimilis,* whose song consists of a five-cycle pulse with a tone of about 540 c.p.s. [*see upper illustration at right*]. During courtship this fly's wings beat about 270 times per second. Because the wingbeat frequency is only half the frequency of the pulse tone we concluded that in this fly's courtship display the wing movement produces sound on both the upstroke and the downstroke and that a short burst of movement is involved. This means that its thorax is activated in a way different from the activation of the thorax in *D. melanogaster.*

Since the oscillation that produces the love song is mechanically similar to the oscillation in flight, it appears that courtship for the fruit fly means using in a special way a mechanism that already exists. This also appears to be true in the case of the female's receptor organ. Studies of other kinds of fly have shown that the primary function of the insects' antennae is to monitor and control the act of flying. Using very loud sounds, we have excited the antennae of fruit flies and have found that the frequencies that cause the aristae to resonate are much the same as the frequency of the flies' wingbeat. The aristae's secondary function as a receptor of the male love song evidently arose from this coincidence.

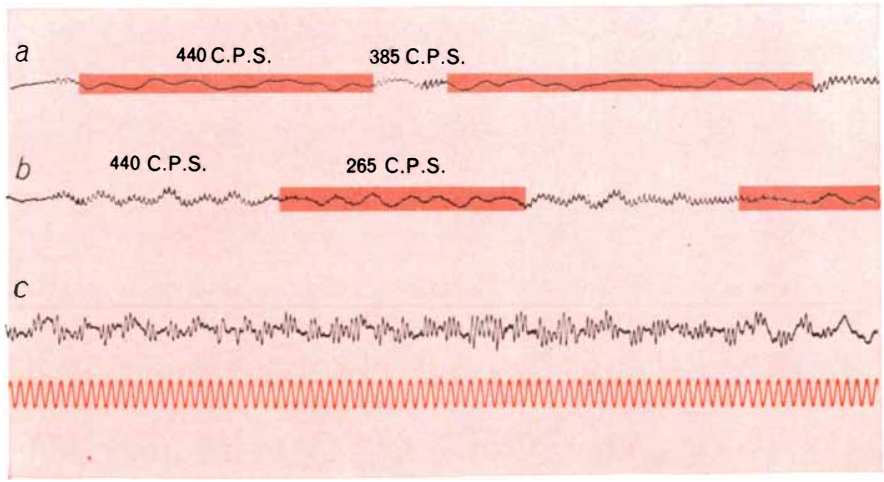
Having learned something about the ways songs are generated and about the organs that receive them, we thought it would be instructive to test the effects of synthetic songs that we could vary with respect to length of pulse and interval between pulses. As we have noted, wingless males court with the same perseverance as winged ones, but the females rarely accept their advances. This meant that we could assess the efficacy of an artificial song by observing whether or not it contributed to a wingless male's success in courtship. The species of fly studied in the experiment was *D. melanogaster.*

We used a sine-wave oscillator driven by a pulse generator to obtain single sinusoidal voltage pulses, but it was harder to reproduce the pulses than to generate them. Each pulse consists of a harmonic series of frequencies, and if the transducer is to be accurate, it must have a "flat" output for all frequencies in the series. We solved the problem by modifying a high-fidelity loudspeaker, making its diaphragm rigid and aiming it into a pipe containing the cell that housed the flies. The sound pulses we reproduced in this way were faithful imitations of the real thing.

Our first test involved spacing the song pulses at different intervals. We not only used the interval between pulses that is normal for the species—three hundredths of a second—but also piped to the flies sound pulses separated by intervals half or twice that long. Only the song with the normal interval induced receptivity in the female flies. This result interested us because many simple animal "clocks" and "counters" can be deceived by presenting signals that are half or double the normal value. The fe-



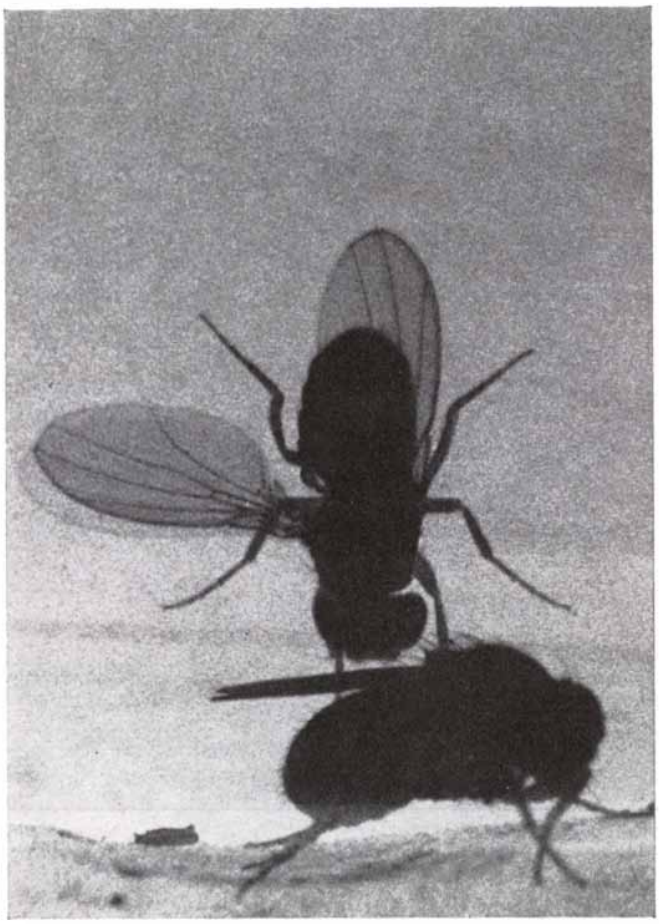
SIMPLE SONGS of five species produce quite different oscilloscope records. The first (*a*), the song of *Drosophila bipectinata,* consists of rapidly repeated pulses, each made up of a single cycle of tone (*colored area*). The interval between pulses is .0085 second. The song of *D. affinis* (*b*) consists of less frequent pulses, spaced .028 second apart and having a double cycle of tone. The pulses of the next three songs may have from three to seven cycles of tone. *D. persimilis* (*c*) and *D. pseudoobscura* (*d*) are New World species with overlapping natural ranges; interbreeding is evidently prevented by the three-to-one difference in their song intervals. The difference in interval between *D. pseudoobscura* and *D. ambigua* (*e*) is slight, but the chance of interbreeding in the wild is remote because *D. ambigua* is an Old World species. The frequency used for calibration (*color*) is 200 cycles per second.

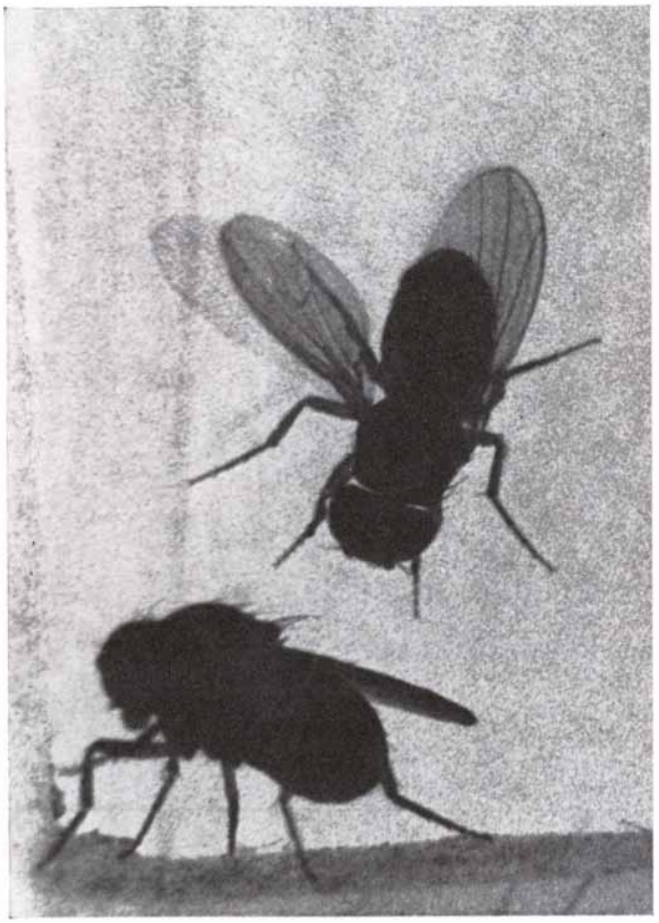


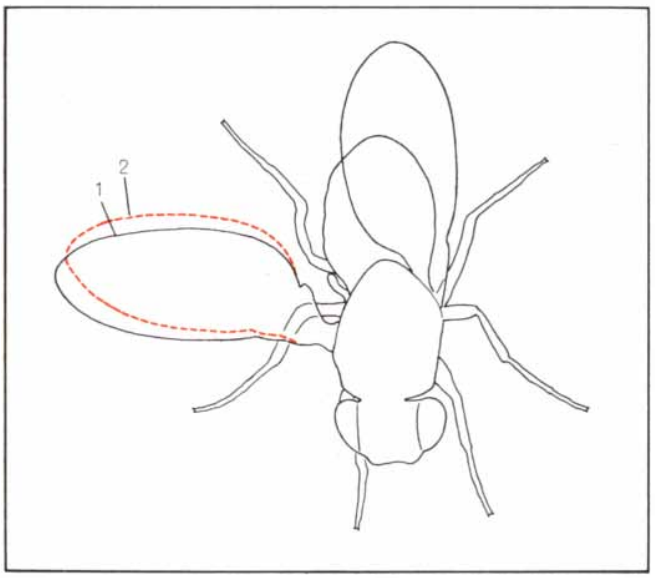
MORE COMPLEX SONGS at first appear to be almost continuous tones. In the case of *D. micromelanica* (*a*) a quiet song (*colored area*) is irregularly interrupted by a loud buzz. In the song of *D. paramelanica* (*b*) the quiet song regularly gives way to a louder portion. The song of *D. athabasca* (*c*) shows amplitude variations but a steady frequency of 440 c.p.s.

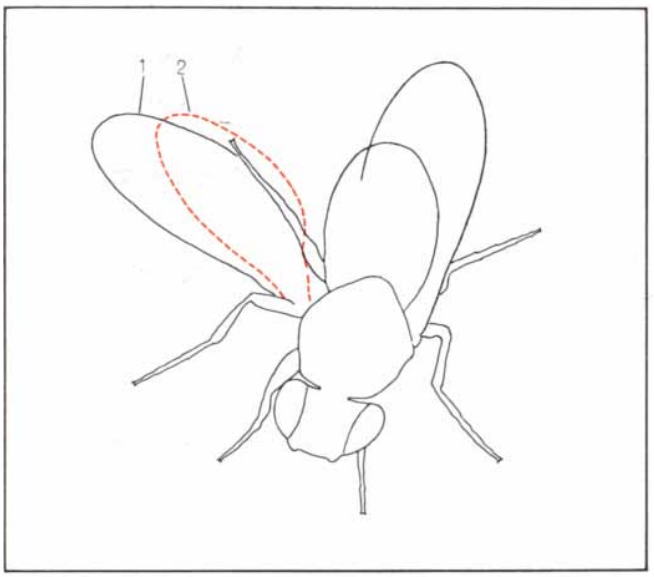
male fly's success in discrimination suggests that her analysis of the courtship signal is a fairly complex process. We also presented pulses separated by an interval of nearly five hundredths of a second, which is the normal interval for the song of a closely related species, *D. simulans*. This song too failed to induce receptivity in the female flies, leading us to the conclusion that the difference in interval is one effective isolating mechanism between these two species.

We next tested the effects of pulses of longer and shorter duration than normal. The pulse of the *D. melanogaster* love song lasts for three thousandths of a second. We presented the flies with pulses half or twice that long; the females proved equally responsive to all three signals. Since any one of the three will excite a single damped vibration of





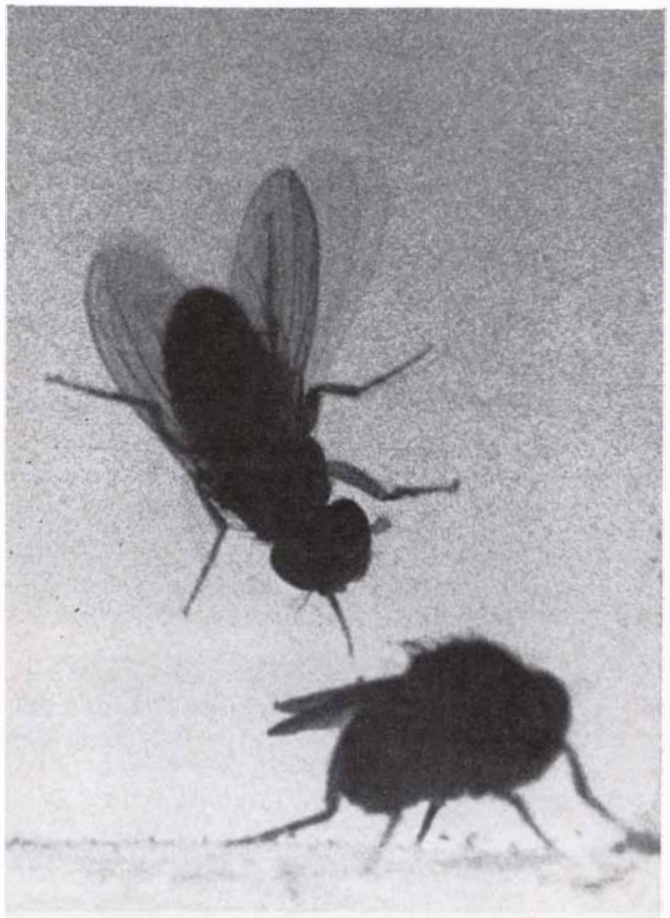


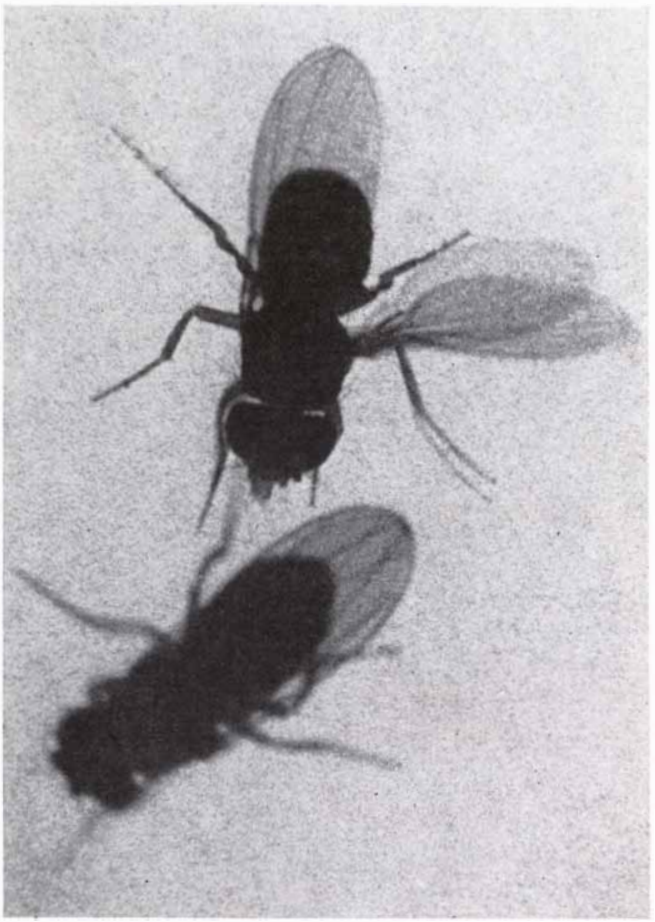


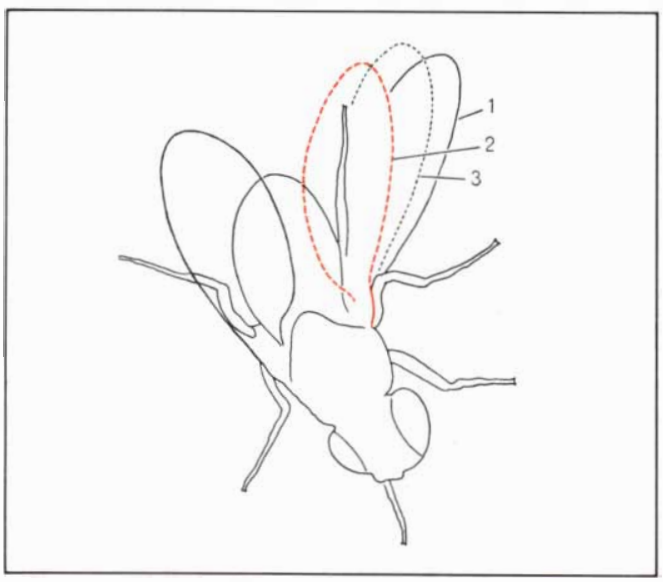
MULTIPLE-FLASH PHOTOGRAPHS of courting *D. melanogaster* males reveal the wing motions responsible for the flies' courtship song. In the first of the four photographs a dim flash is followed three thousandths of a second later by a brighter flash. The drawing below the photograph outlines changes in the right wing's position: it has been lifted up and to the rear. The single sound pulse characteristic of the species' song is produced by the movement. In the second photograph the flashes are two thousandths of a second

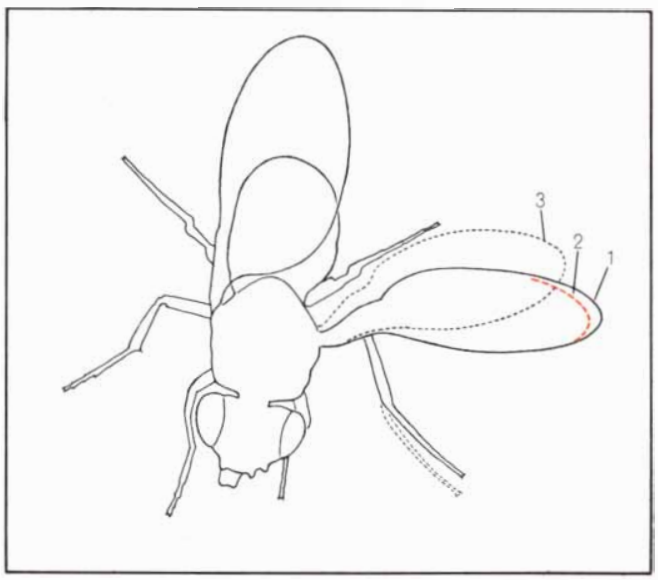
the female's aristae, we did not find the result surprising. In the case of love songs that continue the pulse for more than one cycle, however, the possibility exists that the female is perceiving not only the interval between pulses but also the frequency of the tone and the number of cycles in the pulse. If so, the vocabulary of the love song is elaborate and the range of meaningful song variation between species is great.

We found that the sound pressure needed to stimulate the female was about 115 decibels, which in human terms is roughly equivalent to the climax of Tchaikovsky's "1812 Overture." This led us to the calculation that when both the male and the female fly are standing on a flat surface, the male must approach within two millimeters of the female if his song is to deliver that much sound

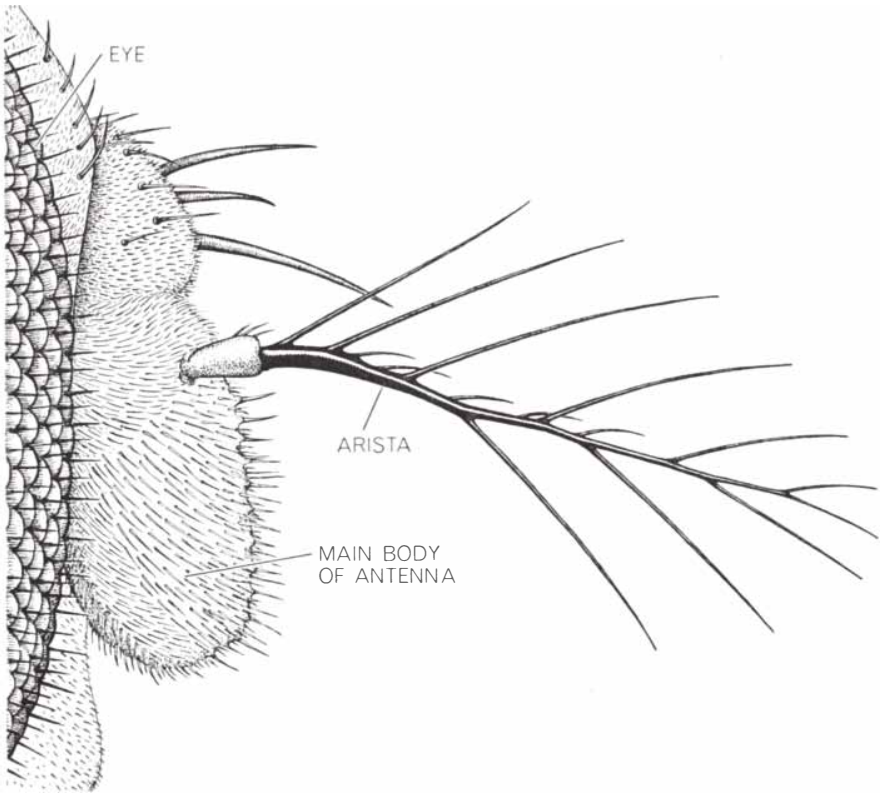




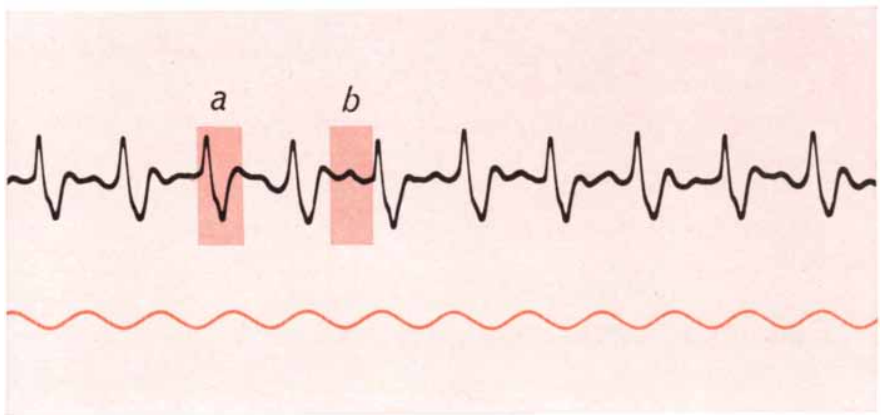




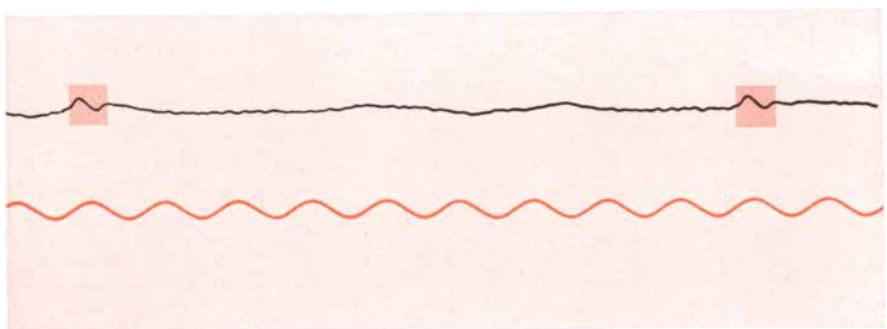
apart; the same upward and backward wing movement is evident. The male's left wing is extended in the third photograph; its motion is recorded by a sequence of dim, bright and dim flashes three thousandths of a second apart. The last of the flashes shows the wing as it moves down after completing an upstroke. The fourth photograph was also made with three flashes at the same interval; the instant recorded is evidently the interval between upstrokes. It is not until the third flash that the male's wing rises again.



FEMALE ORGAN that is vibrated by the song of the male fruit fly is the arista, a feathery extension of the fly's antenna. The organ is shown here 360 times larger than natural size.



FLIGHT SOUND of *D. melanogaster* is caused by resonant oscillations of the fly's thoracic mechanism associated with wing motion. The upstroke (*a*) is loud, the downstroke (*b*) quiet.



COURTSHIP SONG of *D. melanogaster* closely resembles the sound produced by the upstroke of the insect's wings in flight. The pulses (*colored areas*) are .034 second apart.

pressure to her aristae. By getting that close two flies can meet and identify each other even in noisy surroundings.

Good circumstantial evidence supports our conclusion that courtship songs are important in maintaining sexual isolation between various fruit fly species. The closely related species *D. pseudoobscura* and *D. persimilis* can scarcely be told apart on morphological grounds, but it is simple to separate them on the basis of their completely different love songs. The same is true of the closely similar species *D. algonquin* and *D. affinis*. On the other hand, *D. ananassae* and *D. athabasca*, species that are only distantly related, have songs that are quite alike.

Two closely related species of fruit fly have similar songs; they are *D. pseudoobscura* and *D. ambigua*. A real possibility of hybridization might exist here, but a meeting of the species outside the laboratory is extremely unlikely. In nature *D. pseudoobscura* is confined to North America and *D. ambigua* to the Old World. All of this suggests that different song patterns evolve only when there is a possibility that the integrity of the species might break down.

How did the various song patterns evolve? One of us (Ewing) has examined this question, working with flies of closely related races within the superspecies *D. paulistorum*. This is a group that has been studied from the genetic standpoint by Theodosius Dobzhansky and his co-workers; it is believed to be in the process of splitting into several true species. There are six distinct races of *D. paulistorum*, and until recently the species *D. pavlovskiana* was also included in the superspecies. Crosses between the races are partially sterile. The races also show some behavioral isolation: in mixed populations the individual flies mate preferentially with members of their own race.

The songs of the group follow the pattern of *D. melanogaster:* each pulse consists of a single cycle of a single tone. The intervals between pulses, however, are variable. The songs of five of the *D. paulistorum* races are quite similar but the song of the sixth is unique. *D. pavlovskiana* also has a unique song. Judging by other criteria, the race with the unique song is also the most distinctive of the six. Taking all seven populations into consideration, it is evident that the most distinctive are the ones with the unique songs. This suggests that the original barrier to interbreeding of the races is not a difference in song but rath-

Or, for that matter, is Los Angeles? Or Chicago? Or Philadelphia? Or Dallas?

Or any other city groping its way to an uninhabitable anachronism.

A curious situation has developed in America. Eighty per cent of the people in this country live on less than ten per cent of the land area.

There used to be a good reason for this.

At the time of the industrial revolution, we congregated in cities because that's where the sources of energy were. Coal. Water. Electricity.

And our communications network was so limited that we had to be in close proximity to each other for business and social purposes.

No more.

There are no longer any good reasons to continue this hopelessly outmoded life style.

With the advent of the whole spectrum of new communications available to us (wide-band communications, laser beams), we will have the opportunity to live in significantly less dense population centers.

This is no idle prophecy.

The concept is quite realistic and well within the bounds of engineering capabilities which we already have.

Not only do we have the tools to provide the means for new styles in human settlements, but also to rebuild, in a sociological sense, the crowded inner core of our major cities.

The combination of international satellites and cable will provide the means of bringing individuals all the information they need or want without interference or control.

And without the need to be in any specific place.

(Think for a moment about the Apollo 11 moon landing in July, 1969. 500 million people around the world saw, via television, *precisely the same thing at the same time.* Being in New York or Los Angeles held no advantage over being in Keokuk or Harrisburg.)

Historically, we've been preoccupied with moving people and objects. Thus, our intricate network of highways and railroads and airlines — all of which have become enormously inefficient (not inherently, but in application).

The future will see us moving *information,* not, by necessity, people and things.

Your home will be the absolute center of your life.

You will work from home, shop from home, "visit" with family and friends from home, receive in your home any intellectual or cultural achievement known to man.

Fantastic, yes. Fantasy, no.

It is quite within reason to expect these changes by the 1980's.

If we want them.

If we want to change. If we want a better life for ourselves.

Technology has advanced to such an extent, that man is now, literally, capable of changing his world.

Yet, today, a certain gap has developed between the potential of technology and its use by mankind.

There is an obvious contradiction in a method which can land a man on the moon, yet tolerates, perhaps even accepts as inevitable, poverty and ignorance here on earth.

There is a contradiction in a method which affords the best of everything for some, and next to nothing for others.

So we must, in a sense, catch up with the technological potential and apply it for the benefit of all mankind.

All we need sacrifice are the antiquated work practices and our anachronistic traditions.

At RCA, through research and product development, we are committed to closing the technology gap and cancelling the contradictions.

This is the age of the engineer. Nobody understands this better than RCA.

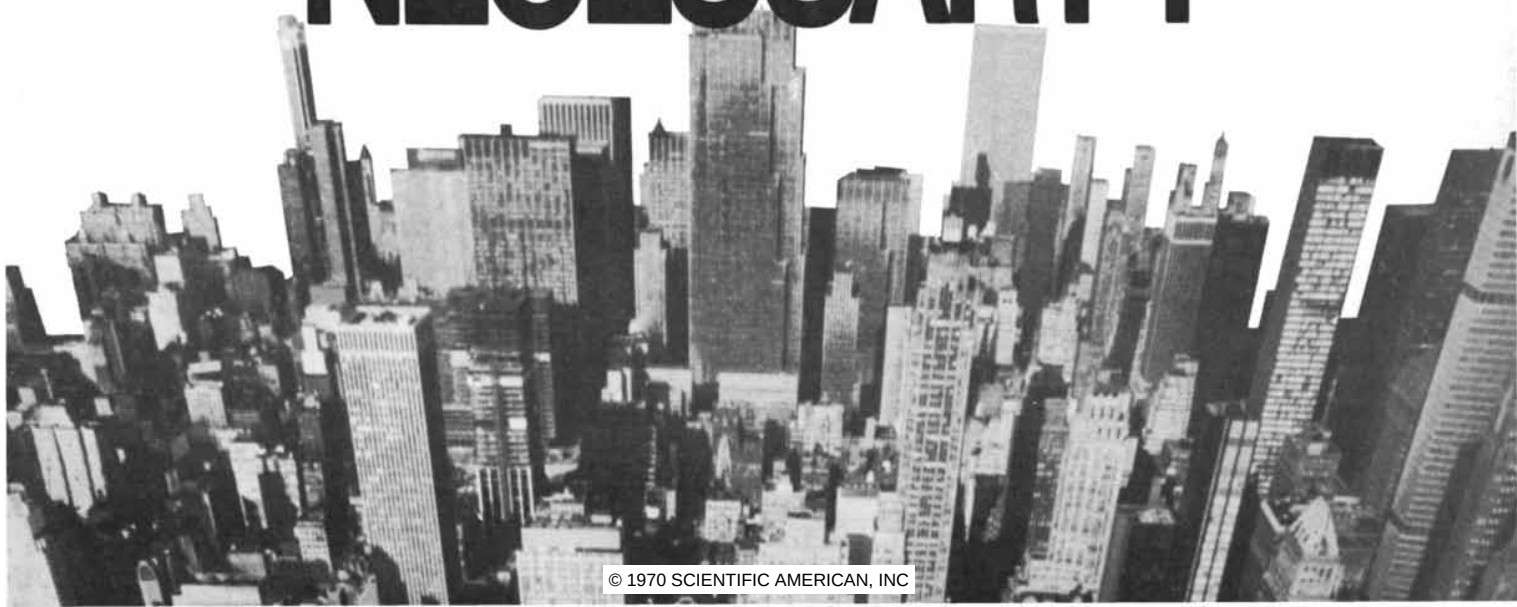
If you're an engineer, scientist or systems programmer, and want to be part of RCA's vision of the future, we invite inquiries.

If you are interested in a comprehensive index of over 1100 technical papers published by RCA scientists and engineers during 1969, let us know.

Write to: Mr. A. C. Bennett, RCA, Bldg. 2-2, Camden, New Jersey 08102. We are an equal opportunity employer.

RCA

IS NEW YORK REALLY NECESSARY?



er a mechanism that evolved in some other way. The unique songs evidently evolve later as a secondary mechanism reinforcing an existing sexual isolation.

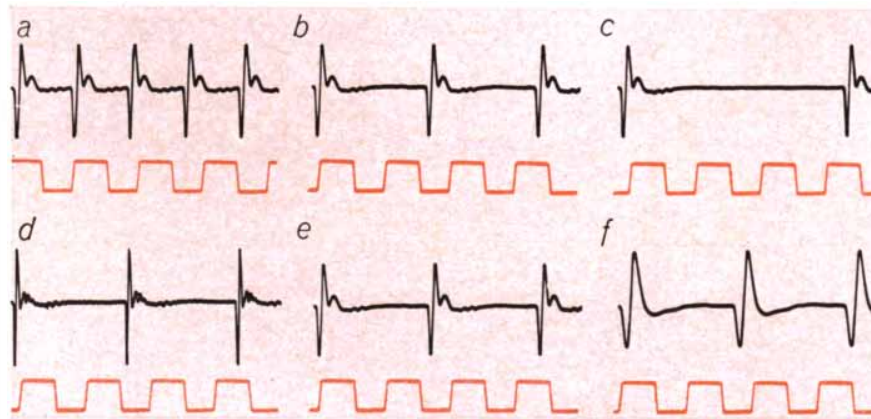
Once the song patterns have evolved, how are they inherited? One of us (Ewing) was able to examine this question by hybridizing *D. persimilis* and *D. pseudoobscura* in the laboratory. The hybrid males are sterile but they still court, so that it is possible to record their songs. The hybrid females fortunately are fertile and so can be backcrossed with either parent species.

The songs of the two species are not at all alike. *D. pseudoobscura* begins its song by emitting a series of widely spaced pulses with a tone of 540 c.p.s. The interval between pulses is nearly two tenths of a second. When the part of the song with the low repetition rate is concluded, there follows a part with a high repetition rate: now the pulse tone is 260 c.p.s. and the interval between pulses is four hundredths of a second. In contrast, the song of *D. persimilis* seldom includes a part with a low repetition rate; when on occasion it does, the pulses are few in number and together last for no more than half a second. Usually the song consists only of a signal with a high repetition rate; the pulse tone is 540 c.p.s. and the interval between pulses is about two hundredths of a second.

The type of song the hybrids inherited proved to be the type of the maternal species. Furthermore, when first-generation females are backcrossed to males of either species, half of the male progeny perform songs of the first species and half perform songs of the second. It follows that the inheritance of song type is sex-linked and that the genes are borne on the X chromosome rather than on the Y chromosome. The inheritance of pulse interval is more complex and may be partly controlled by genes on chromosomes other than sex chromosomes (the X and Y chromosomes).

The outcome of the hybridization experiment suggests an answer to one of the knottier evolutionary questions that is raised by the very existence of elaborate, species-specific signals. How, one wonders, have the signals been able to change with the passage of time? Any change that occurred in the signal or the receiver would obviously be disadvantageous. This is borne out by the fact, established by our studies, that the songs of a single fruit fly species are extremely durable, showing no measurable variability even when the populations of the species are geographically isolated. We have recorded *D. melanogaster* males that were native to Edinburgh, to San Francisco and to the Amazon basin; all have identical songs.

The fact that inheritance of song type is sex-linked suggests a means for new songs to evolve. The first-generation hybrid females from the crossbreeding of *D. pseudoobscura* and *D. persimilis* were receptive to males of both species. The same could well be true in the case of a gene change that affected the qualitative aspect of a song. The result would be first-generation males that performed a new song pattern and first-generation females that were receptive to it. Selection for sex-linked characteristics is much more rigorous than selection for non-sex-linked ones. Under suitable circumstances such a genetic change would rapidly become fixed in a population.



ARTIFICIAL SONGS were played to *D. melanogaster* females to determine if the interval between pulses was a key factor in song recognition. The interval normal to the species is .034 second (*b*); the females were responsive to this signal. When pulses were presented at twice the rate (*a*) and half the rate (*c*), the females did not respond. The normal duration of each song pulse is .003 second (*e*). Artificial pulses of half the duration (*d*) and twice the duration (*f*) but retaining the normal interval between pulses were played to the females; they responded to all three signals. The calibration frequency (*color*) is 50 c.p.s.

MGB/GT '70

It's like the man who drives it. The exception rather than the rule.

Here's the one GT that isn't aimed at the ordinary man. Its new mag-style wheels, radial-ply tires, and sophisticated recessed grillework reflect someone who intuitively understands the grand touring spirit. Someone who looks for luxuries like molded bucket seats and pile carpeting. And driving features that make the rules. Like the 1798 cc. MG engine, rack-and-pinion steering, and race-proven suspension. That's the MGB/GT '70. The exceptional car for the exceptional man.

For overseas delivery information, write British Leyland Motors Inc., Leonia N.J. 07605.

NETWORK ANALYSIS

This new planning method can be applied to power grids, computer networks, gas pipelines and other flow-handling systems in order to increase their effectiveness and reliability and reduce costs

by Howard Frank and Ivan T. Frisch

A modern society is to a large extent a system of networks for communication, transportation and the distribution of energy and goods. The complexity and cost of these networks demand that existing networks be effectively used and that new networks be rationally designed. To meet this demand there has evolved a new discipline called network analysis.

Network analysis is the study of connected lines and points. The lines, termed branches, can represent roads, power lines, telephone wires, railroad tracks, airline routes, water mains or the generalized channels through which commodities flow. The points, called nodes, can represent communities, highway intersections, power stations, telephone exchanges, railroad yards, airline terminals, water reservoirs and computers; in general, a node can represent any point where a flow originates, is relayed or terminates. The natural limitations and capabilities of the network's nodes and branches can be described by numbers. These numbers can be fixed or they can vary with time. They can even be random numbers whose values cannot be precisely predicted.

A power system might be represented by a network in which the branches are transmission lines and the nodes are generator stations, substations and customers. Thus a node representing a power station might be characterized by numbers representing the maximum power output, the number of generators at the station, the reliability of each generator and the cost of electricity per kilowatt-hour. A typical branch might have three numbers corresponding to its maximum power-handling capacity, its reliability and its cost.

The origins of network analysis are old and diverse. Network analysts rely heavily on graph theory, a branch of mathematics that was founded with Leonhard Euler's formulation and solution of the first graph-theory problem in 1736. More than a century later James Clerk Maxwell and Gustav Robert Kirchhoff discovered certain basic principles of network analysis in the course of their studies of electrical circuits. Early in the 20th century telephone engineers in Europe and the U.S. devised methods for determining the best capacity of telephone trunk lines and switching centers in order to guarantee specified levels of customer service. In the 1940's operations research yielded a number of techniques for the mathematical study of large-scale systems. Most of these studies took up problems of economics or logistics with such methods as linear programming.

Concepts of this kind, together with queuing theory, probability theory, statistics and computer programming, are the armamentarium of network analysis. This form of systems analysis has been applied to network problems of such complexity that they were considered virtually unsolvable. One such application of network analysis was made by a consulting group of electrical engineers and mathematicians, of which we were members, attached to the Office of Emergency Preparedness in Washington. This group was asked to develop a rational, economic method for planning offshore natural-gas pipeline networks. The method the group developed ultimately saved many millions of dollars. Last year we and other members of the group formed the Network Analysis Corporation.

We and our colleagues have analyzed and designed large-scale irrigation systems, computer networks and cable television networks. Among the many other areas where network analysis is making significant contributions are transportation, ground and satellite communications, the warehousing and distribution of goods, industrial scheduling and energy transmission. Here we shall outline the scope of network analysis by discussing how it can be utilized to find maximum flows in networks, to determine the vulnerability of a network to disruption, to select pipe diameters for gas pipeline networks and to design the network itself.

In considering the performance of a network it is often necessary to calculate the maximum flow between a source and a terminal. This calculation is sometimes called the one-commodity flow problem because the flows correspond to a single commodity such as electric power, water, information or air traffic. Flows of a single commodity behave in certain predictable ways. For all nodes other than the source and the terminal, the total flow into the node equals the total flow out of that node. This property is called conservation of flow. Each branch can accommodate a maximum amount of flow, which is called the capacity of the branch. It follows that the total flow of any node is limited by the capacity of the branches that must handle this flow.

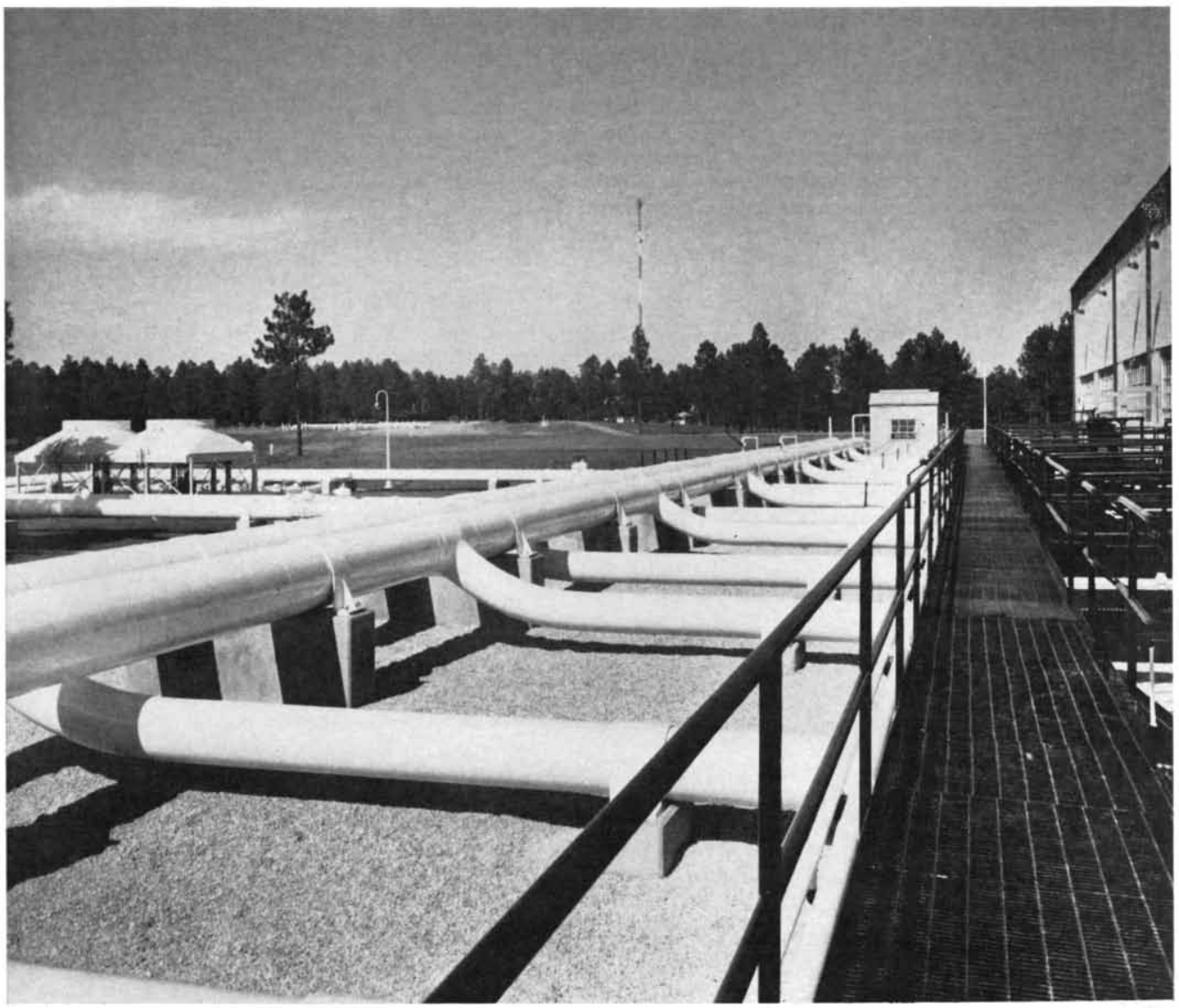
The maximum flow is determined by a fundamental structural property of the network called a cut. A cut consists of a set of branches that connect one group of nodes to all remaining nodes [*see illustration on page 96*]. The capacity of a cut, that is, its ability to handle flow, is defined as the sum of the capacities of its branches. If a cut consists of two branches, each with a capacity of one unit of flow, the cut's capacity is two units.

An important conclusion can be drawn

from this set of facts. If the source and the terminal are in different groups of nodes, all flow from one to the other must pass through the cut that connects them. Therefore the maximum flow from source to terminal cannot exceed the capacity of the cut. In particular it cannot exceed the capacity of the minimum cut, that is, the cut of smallest capacity [*see illustration on page 97*]. One of the most important results in the theory of flows, the max-flow min-cut theorem (proved by Lester R. Ford, Jr., of the General Research Corporation and Delbert R. Fulkerson of the Rand Corporation), can now be stated: The maximum flow from source to terminal is equal to the capacity of the minimum cut between them. Furthermore, if certain nodes between the source and the terminal also have limited capacity, then the maximum flow is limited by the capacities of these nodes. The max-cut min-flow theorem can relate flows to capacities of nodes as well as to capacities of branches.

The max-flow min-cut theorem is satisfying because it can be simply stated and confirms our intuitive notions of the behavior of flow. One would expect that variations of the theorem would hold in more general situations. In particular let us consider the flow problem in which there are two or more different commodities being sent through a network. In this problem each commodity has a single source and a single terminal. A given node may, however, be a source or a terminal for more than one commodity. The flow of each commodity is conserved at each node except for its own source or terminal. The sum of the magnitudes of the flows of the various commodities in a branch must not exceed the capacity of that branch. We then want to find the maximum of the sum of the flows of the various commodities from their sources.

The max-flow min-cut theorem tells us that for one commodity the minimum cut will determine the maximum flow. The proposition can be stated in more familiar terms. A flow follows a path from a source to a terminal. This path is an alternating sequence of nodes and branches, beginning with the source and ending with the terminal, such that no nodes or branches are repeated and each branch is connected to the nodes immediately preceding and following it in



BRANCHING PIPES at a natural-gas pumping station are an example of the kind of system that can be subjected to network analysis. The large pipe in the foreground at left feeds gas through branches to compressors in the building at right. The pipe behind it returns the compressed gas to the pipeline. The pumping station, which is operated by the Trunkline Gas Company, is located at Pollock, La.

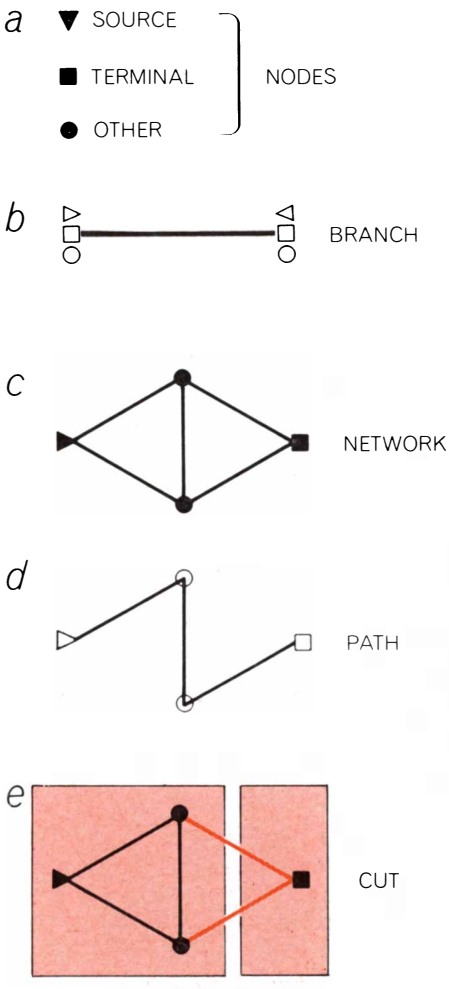
the sequence. All paths from the source and the corresponding terminal can be broken by the removal of one or more branches. In the one-commodity case it is clear that the flow cannot exceed the minimum capacity among such sets of branches.

For a time it was thought that maximum flow equaled capacity for all flows of more than one commodity. This conjecture, however, does not hold true in all cases. To indicate the elusiveness of the problem let us describe three situations, two where flow equals capacity and one where flow is less than capacity.

We have already seen that flow equals capacity for the one-commodity case. It has been shown by S. Louis Hakimi of Northwestern University and by T. C. Hu of the University of Wisconsin that the maximum flow equals capacity when there are only two commodities in the network. In the case of three commodities Bill Rothfarb of the Network Analysis Corporation and one of us (Frisch) have shown that if every node is the source or terminal for exactly one commodity and each commodity has only a single source, flow will equal capacity.

It is possible, however, to construct a network with three commodities for which maximum flow is always less than capacity [*see top illustration on page 98*]. Each branch in this network has a capacity of two. Again there are three sources and three terminals—one source and one terminal for each of three commodities. The minimum number of branches that must be removed to separate the sources from the corresponding terminals (the equivalent of the minimum cut) is four. Since the capacity of each branch is two, we know that the capacity (C) of this set of branches is eight. We might therefore expect the maximum flow to be equal to eight.

It is, however, less than eight. This conclusion is easily shown using the interesting fact that in this network every path between a source and its terminal consists of exactly three branches. Thus every unit of flow of any commodity between its source and its terminal must occupy three units of branch capacity. Three times the quantity of flow (V) of a commodity cannot exceed 20 units of capacity, the total of all the branch capacities. Stated algebraically, $3V \leq 20$, therefore $V \leq 20/3$ and $V \leq 6\frac{2}{3}$, but $C = 8$. Hence V is less than C.



GLOSSARY of network analysis terms lists five basic concepts: node, branch, network, path and cut. A node (*a*) may be the source where flow originates, a terminal where it ends or a point through which it passes. A branch (*b*) conducts flow between nodes. A network (*c*) consists of nodes connected to one another by branches. A path (*d*) is a sequence of nodes and branches that conducts flow between source and terminal. There are node disjoint paths and branch disjoint paths. Node disjoint paths are paths that share no nodes with one another. Branch disjoint paths are paths that share no branches with one another. A cut (*e*) consists of the branches (*color*) that connect a set of nodes with the remaining nodes.

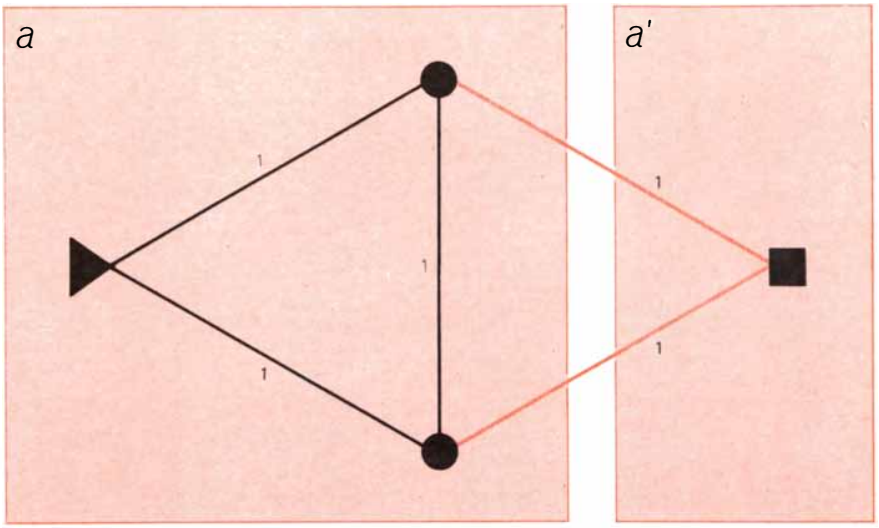
The max-flow min-cut theorem and related rules enable network analysts to understand and solve many other important problems. One of these is the problem of vulnerability and reliability. The larger and more complex a network is, the less likely it is that all its components will operate perfectly at all times. Accordingly one of the most important objectives in designing a large network is to guarantee that it will function effectively even after some of its elements fail. The study of the possible failure of network elements and the subsequent overall degradation of network performance is called vulnerability analysis or reliability analysis. The term reliability is used when element failures are the consequence of factors such as natural disturbances and the aging of equipment. Vulnerability concerns the network's resistance to major damage.

The first step in studying the reliability or the vulnerability of a system is to formulate precise criteria for its failure. The choice of criteria depends on the type of network being considered and its purpose. A communications network might be said to fail if (1) a message from a sender to a receiver must be relayed through a large number of intermediate nodes to reach its destination, (2) a specified set of senders cannot communicate with a specified set of receivers or (3) at least one sender cannot communicate with at least one receiver. In order to understand some of the aspects of reliability and vulnerability suppose a network fails if at least one sender and one receiver are disconnected. It will be assumed that a network is connected if there is at least one direct path between each pair of nodes; otherwise it will be considered to be disconnected.

For a given network affected by outside factors rather than by disruption or aging a fundamental problem is to find the minimum number of nodes or branches that must fail in order to disconnect the network. It may be very difficult to answer this question. For example, it may not be at all obvious that the removal of a single node or a single branch from a network will disconnect it. Yet it is essential to be able to identify such branches efficiently for two reasons. First, it may be necessary to determine the reliability of an existing network. Second, it may often be necessary to analyze and modify a proposed network repeatedly before a reliable design is achieved.

Here the max-flow min-cut theorem can be used to find the minimum number of nodes and branches that must be removed in order to disconnect a network. Let us illustrate this point in more detail by indicating how one can find the minimum number of branches. If each branch is assigned a capacity of one unit, then the maximum flow between any two nodes is equal to the minimum number of branches that must be removed in order to break all paths from one of the nodes to the other. It can also be shown that this flow is equal to the maximum number of paths between the nodes that have no branches in common. Such paths are called disjoint paths. We can now repeat this calculation to find the maximum flow between all pairs of

nodes in the network. The number of branches is then given by the smallest of these maximum flows. Similarly, in order to find the minimum number of nodes that must be removed to break all paths we can assign capacities equal to one to all nodes other than the source and the terminal. These capacities represent the upper bounds on the flows into the nodes. We can then find the maximum flow from the source to the terminal by an algorithm developed by one of us (Frisch). This number is equal to the maximum number of paths that have no nodes in common except the source and terminal. These paths are called node disjoint paths. The calculation is repeated for all pairs of nodes. The number of nodes that must be disconnected is then revealed by the smallest of the maximum flows.



MAX-FLOW MIN-CUT THEOREM states that the maximum flow in this network from the source (*triangular node at left*) to the terminal (*square node at right*) cannot exceed the capacity of the minimum cut, which consists of the two branches (*color*) that join the nodes in *a* (*shaded area at left*) to the node in *a'* (*shaded area at right*). If both branches in cut were broken, there could be no flow from nodes in *a* to the nodes in *a'*. Therefore the flow cannot exceed capacity of the minimum cut, that is, two-unit capacity of these two branches.

Significant simplifications can be made in this kind of calculation. A typical problem is to determine whether or not a network of 1,000 nodes will be disconnected if as many as four arbitrarily chosen nodes fail. To solve the problem the number of disjoint paths connecting each pair of nodes could be counted. Such a procedure might involve almost 500,000 sets of the complicated calculations called flow maximizations, because each node would have to be examined in conjunction with every other node. It also involves far more work than is actually necessary. Daniel J. Kleitman of the Massachusetts Institute of Technology has shown that the answer can be derived by performing the following calculations.

Choose any node and verify that there are at least four node disjoint paths from this node to every other node. Remove the node and all the branches connected to it. From the resulting network choose another node and verify that there are at least three node disjoint paths connecting this node to every other node. Remove this node and its branches, choose another and then verify that there are at least two node disjoint paths from it to every other node. Choose a fourth node. Remove it and its branches. Verify that there is at least one path from it to every other node. These are the only calculations that are required. If the number of disjoint paths between each pair of nodes had been only three, then at the last stage of repetition it would have been impossible to verify that the required number of paths existed [*see bottom illustration on page 99*].

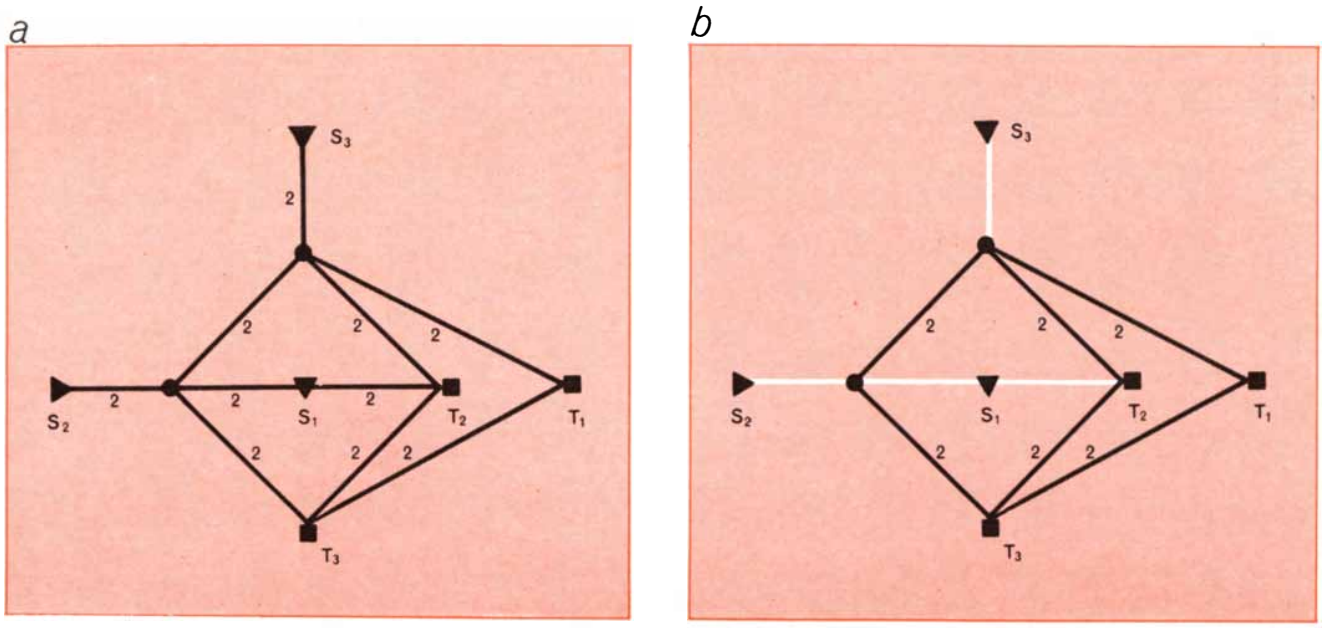
The Kleitman solution for the problem of finding the number of nodes involves 3,999 flow maximizations for a network of 1,000 nodes, a number smaller by a factor of 1,000 than the total number of maximizations that must be made in a straightforward counting procedure. Given any number of nodes and any number of required node disjoint paths, this method makes it unnecessary to examine all possible combinations of nodes in the network. It is also more efficient because with each repetition one node and its associated branches are removed, so that the difficulty of executing the next step is reduced.

A measure of performance for a network affected by random factors such as aging and equipment failure is the probability that it is connected. The closer the probability to one, the more reliable the network. Such networks have peculiar properties that are not easily deduced by studying the analogous properties in networks not subject to random disturbances. We have already discussed the problem of finding the minimum number of nodes that must be removed from a connected network in order to disconnect it. If nodes fail randomly, the corresponding problem is to compute the probability of the network's remaining connected.

Both of the networks shown at the top of page 99 have eight nodes and 12 branches. Suppose we attempt to disconnect both by removing the smallest possible number of nodes. We can disconnect the first network by removing Node 1 and Node 2. In fact, after these nodes are removed, along with the branches connected to them, there are no branches left in the network, and so all the other nodes are isolated from one another. The minimum number of nodes that must be removed from the second network, however, is three. The removal of nodes 6, 8 and 3 will separate Node 7 from the rest of the network.

Now, suppose the nodes fail randomly, so that the failure of one does not affect the operation of any other. The measure of the reliability of the network will be the probability that at least two nodes do not fail and that the network consisting of the operating nodes remains connected. In general it is extremely difficult to compute this probability for a network with many nodes and branches. Because these two networks are simple, however, the calculations are easily done. The first network will be connected unless nodes 1 and 2 fail simultaneously or nodes 3, 4, 5, 6, 7 and 8 fail simultaneously. In the second network any combination of two nodes can fail and the network will remain connected. By enumerating the ways nodes in this network can fail we discover that there are eight combinations of three nodes whose failure will disconnect the network.

Because the first network can be disconnected by removing two nodes and the second cannot be disconnected by removing any pair of nodes, it might be expected that the first is less reliable than the second. Actually it is by no means

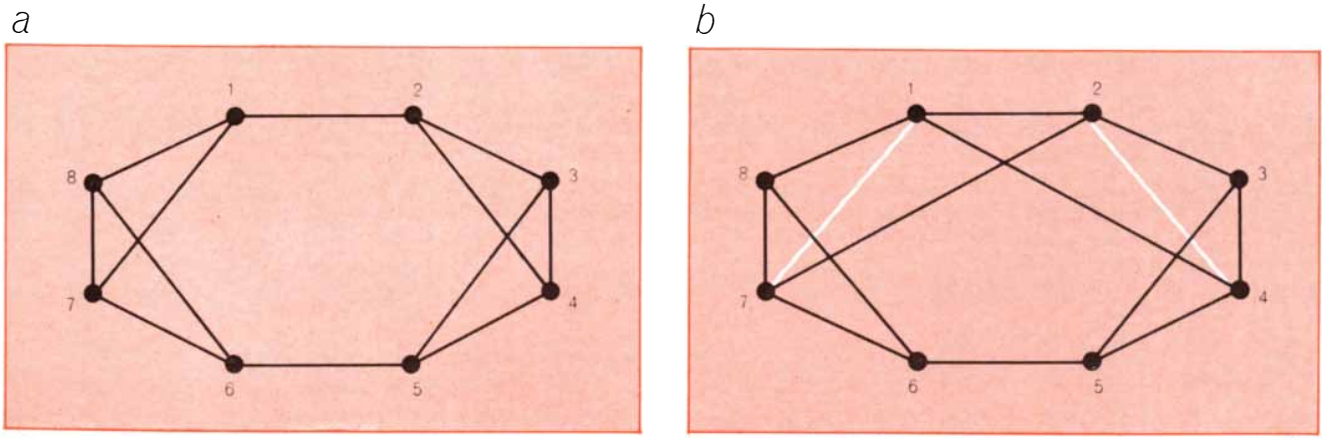


THREE-COMMODITY SITUATION where maximum flow is less than capacity exists in this network. In *a* each source (S_1, S_2 *and* S_3) corresponds to a terminal (T_1, T_2 *and* T_3). A branch can carry two units of flow; each path connecting a source with its terminal has three branches. A unit of flow between a source and a terminal therefore uses three units of capacity. Yet in *b* it can be seen that the minimum cut has a capacity of eight. (Four branches, shown in white, must be removed in order to isolate each source from its terminal.) The total available capacity of the network is equal to 20 units of capacity. Since each unit of flow uses three units of capacity, however, the maximum flow in the network cannot exceed 20/3, or 6⅔, units, a value smaller than that of the minimum cut.

clear which network is "better." If nodes 1 and 2 in the first network do not fail, all six remaining nodes must fail before it is disconnected. On the other hand, the second network has eight distinct combinations of three nodes whose removal will disconnect it.

One of us (Frank) has developed a general theory of networks with randomly failing elements. This theory shows that if the probability that any node is functioning properly is small enough, the probability that the first network is connected is always greater than the probability that the second is connected. If, however, the probability of correct functioning is large, then the second network is more reliable. This relation means that it is not always possible to find a best network. A situation may arise in which a network is attacked. The nodes may operate 99 percent of the time if there is no attack, whereas each node may have only an 80 percent chance of survival when the network is attacked. In this situation the first network will be less vulnerable but more unreliable than the second one.

Thus far we have considered examples in which we were examining structural properties of networks. Network analysis also provides methods for solving complex design problems involving many different considerations such as choosing the most economic plan for underwater gas pipelines. Economies are important because the installation of underwater pipelines is a major cost in de-



MAXIMIZING THE MINIMUM CUT is a method for determining the best arrangement of branches in a network. In *a* the minimum cut contains two branches. In *b* branches connecting nodes 1 and 7 and nodes 2 and 4 are removed and replaced with new branches connecting nodes 1 and 4 and nodes 2 and 7. The minimum cut now joining nodes 1–5 with nodes 6, 7 and 8 consists of three branches.

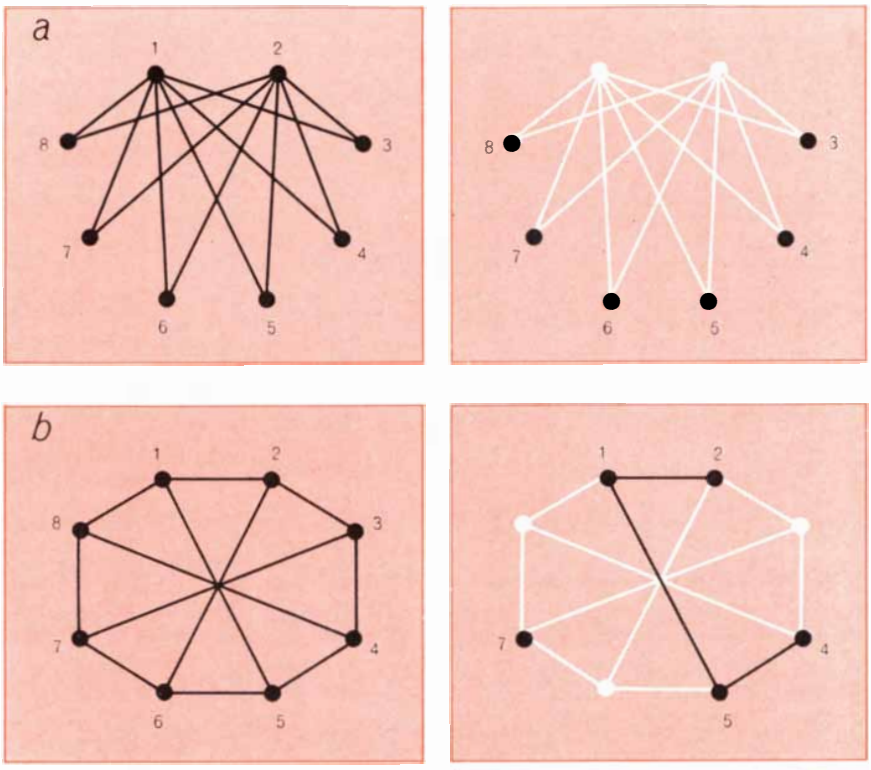
veloping the enormous reserves of natural gas located in the Gulf of Mexico and other offshore regions. A pipeline network is costly to buy and install. A pipeline in the Gulf of Mexico, for instance, can cost from $65,000 to $325,000 per mile, depending on pipe diameter and pipeline depth.

Each pipe is characterized by its internal and external diameter, its type of material and fabrication and its length. The flow of gas through a pipe depends on internal diameter, type of material, fabrication, length, the pressures at the end points, the grade, the temperature and the physical characteristics of the gas. The equation most often used for gas flow is $Q = K[(P_1^2 - P_2^2/L)]^{.5394} \times D^{2.6182}$. In this equation Q is the flow in volume in cubic feet per day, K is an empirically derived constant, P_1 is the input pressure in pounds per square inch of area, P_2 is the output pressure in pounds per square inch of area, L is the length of pipe in miles and D is the internal pipe diameter in inches.

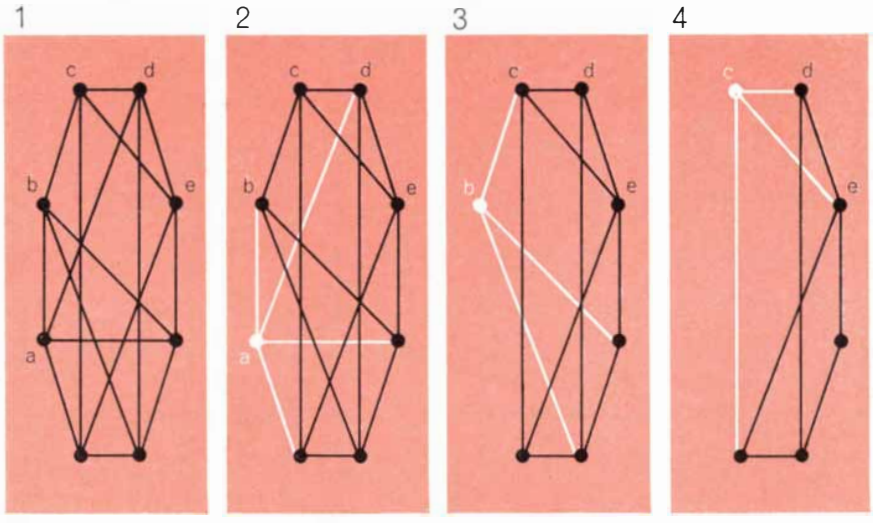
There are constraints on the maximum pressure that can be applied at any point in the system and on the minimum pressure at which gas can be delivered ashore. Clearly the pressure in a pipe cannot exceed a specified limit determined by factors such as safety conditions. In addition gas must be delivered to the shore at a pressure greater than or equal to a specified lower bound determined by the capabilities of the onshore processing plants and network.

The flows in each branch of a network are determined by the flows out of each gas field. Specifying the diameter of a branch also specifies, by means of the flow equation, the difference in the squares of the pressures across that branch. It follows that once diameters have been assigned to all the branches, the node where the pressure is greatest can be determined. Since there is a maximum allowable node pressure, the pressure at the delivery point must be set low enough so that the necessary pressure-squared differences can be maintained without exceeding the pressure limit anywhere in the network.

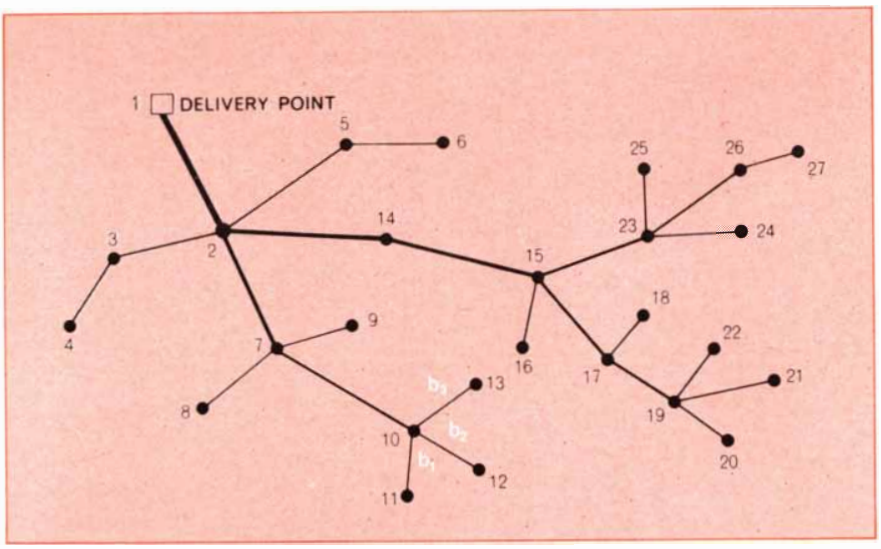
A typical gas field may require a network that connects 25 wellheads with an onshore delivery point. The problem is to select a set of pipe diameters that minimizes the sum of the annual operating and investment costs. It can be seen that, even with a computer, rational selection by the examination of every conceivable combination of pipe diameters for the network is impossible. For seven pipe diameters there are 7^{26} possible



RELIABILITY ANALYSIS requires determination of which of two networks is more reliable. It appears that Network *a* is less reliable than Network *b* because Network *a* can be disconnected by the removal of nodes 1 and 2 and their related branches as shown at upper right. Network *b*, however, can be disconnected if some combination of three nodes such as nodes 3, 6 and 8 and their branches are removed, as shown at lower right. Actually, however, Network *a* may prove more reliable than Network *b*. If nodes 1 and 2 do not fail, then six other nodes must fail at once to disconnect Network *a* by separating Node 1 from Node 2.



RELIABILITY PROBLEM consists of determining how many nodes need to be removed in order to disconnect a network. (A network may be disconnected if a node is isolated from all other nodes.) Here the problem is to find out if the removal of fewer than four nodes will disconnect the network shown in *1*. According to an algorithm developed by Daniel J. Kleitman of the Massachusetts Institute of Technology, the number of nodes that must be removed can be determined in the following way. In *1* select any node, such as Node *a*, and verify that there are at least four node disjoint paths from it to every other node. Remove this node and its branches. The result of this operation is the network in *2*. (*White indicates nodes and branches removed.*) The operation is repeated on this network and then on the resulting networks in *3* and *4*. Each time the required number of node disjoint paths is one fewer than it was before. In *4*, after Node *c* has been removed, at least one connects the nodes. It can therefore be concluded that at least four nodes must fail to disconnect the network.

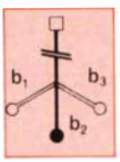


"TREE" NETWORK of pipelines connects natural-gas wellheads in Gulf of Mexico with onshore delivery point (*open square*). In order to reduce the cost of building and operating the network optimal combinations of pipe diameters must be assigned to all the branches.

PRESSURE DIFFERENCE

PIPE DIAMETER (INCHES)	11¾	12¾	16	20	24	26	30
b_1 LIST	120	111	92	66	54	40	31
b_2 LIST	150	139	118	87	75	70	67
b_3 LIST	94	86	80	61	55	48	32

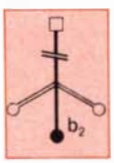
a



PRESSURE DIFFERENCE

b_1	b_2	b_3
120	**150**	94

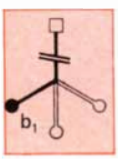
b



PRESSURE DIFFERENCE

b_1	b_2	b_3
120	**139**	94

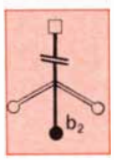
c



PRESSURE DIFFERENCE

b_1	b_2	b_3
111	87	94

d



PRESSURE DIFFERENCE

b_1	b_2	b_3
66	**67**	61

PARTIAL-ASSIGNMENT METHOD is utilized in order to derive the best combinations of pipe diameters for the branches b_1, b_2 and b_3 in the gas pipeline in illustration at top of page. There are seven pipe diameters for each branch. The object is to find the combination that will reduce the difference in pressure along the critical path, which connects delivery point with the wellhead producing the highest pressure. In *a*, when smallest diameter is assigned to b_2, the pressure difference is greatest along this branch (*black*). Therefore neither b_1 nor b_3 is on the critical path. Smallest diameter is also assigned to b_1 and b_3 because larger diameter would only increase pipe cost without reducing pressure difference on critical path. In *b* the pressure difference is reduced by assigning the next-largest pipe diameter to b_2. In *c*, after several recalculations, the pressure differences have been further improved by assigning the 12¾ inch pipe to b_1. Therefore the critical path cannot lie along the other branches. In *d* pressure difference is optimized by assigning largest diameter to b_2.

combinations of diameters. If a computer could evaluate a million combinations per second, it would need millions of hours to examine all the possible combinations.

Through network analysis the uneconomic diameter combinations can be eliminated without even being examined. First the types of network under consideration are simplified. Among the configurations that can be used are those called "trees," which consist of the smallest possible number of pipes that can connect the gas fields and delivery points. One method for finding the best set of pipe diameters for a specified tree structure was developed by one of us (Frank) in collaboration with Rothfarb, Kenneth Steiglitz of Princeton University and Kleitman.

In this approach diameter combinations are examined for one set of branches at a time. An assignment of diameters to such a set is called a partial assignment. This method reduces computation time by discarding partial assignments that cannot be in the optimal assignment for the entire network, without discarding partial assignments that might be in the optimal assignment.

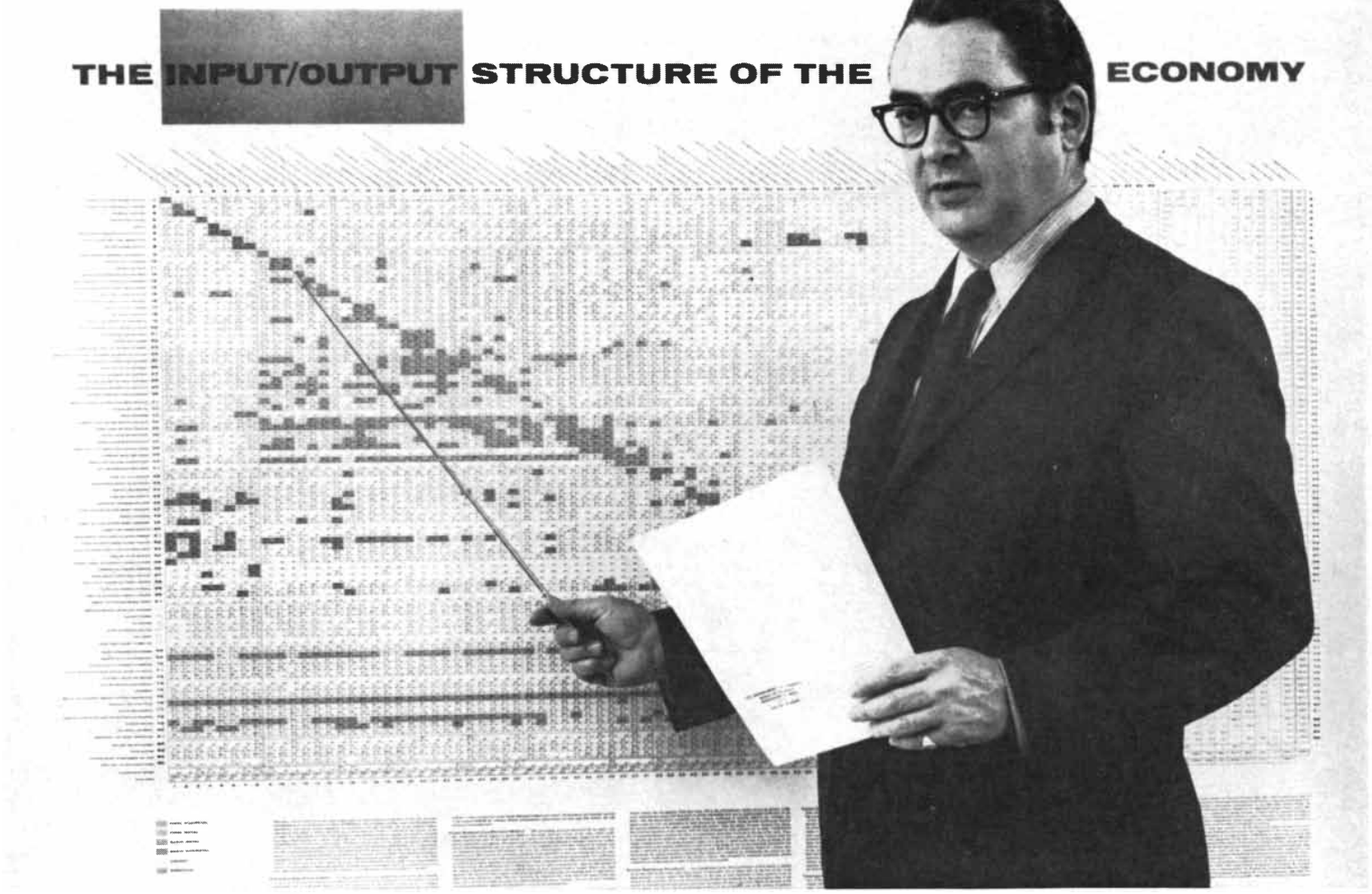
In order to understand how the partial-assignment method works it will be helpful to consider a specific example such as the selection of diameter combinations for the set of three branches (b_1, b_2 and b_3) that are part of the gas-pipeline network in the top illustration at the left.

The path in the tree from the delivery point to the node of greatest pressure is termed the critical path. The sum of the pressure-squared differences along the critical path determines the delivery-point pressure and hence the cost of compression needed to meet the minimum required delivery-point pressure.

There are seven diameters of pipe that can be assigned to each branch. The first step is to make a list of pressure-squared differences and pipe costs for each possible diameter of the branches b_1, b_2 and b_3. When the list is drawn up, it appears that the pressure-squared difference is highest across b_2 when the smallest diameter is used for this branch. Therefore if b_2 has the smallest diameter, neither b_1 nor b_3 could be on the critical path.

The next step is to eliminate diameter combinations for b_1 and b_3. Since neither branch can be on the critical path with b_1 at its smallest diameter, these branches should also be assigned minimum diameters. Larger diameters on b_1 and b_3 would only increase the pipe cost without reducing the compression cost.

Can the diameter assignment on the



WHAT DOES IT TAKE TO PRODUCE A $1000-BILLION GNP?

The 1970 input/output table, just published by SCIENTIFIC AMERICAN, displays the interindustry flows of raw materials, intermediate products and business services required to carry the U.S. economy to the benchmark Gross National Product of $1000 billion.

Input/output tables provide management, government administrators, economists and market analysts with a powerful new tool for forecasting and measuring the indirect as well as the direct interindustry relationships that structure our industrial economy.

This handsome and informative wall chart (70" x 46", in eight colors) offers a unique entry into the rapidly developing discipline of interindustry (or input/output) analysis. Based upon input/output tables issued by the Office of Business Economics of the U.S. Department of Commerce, the chart can be used as a teaching tool and for study of practical and theoretical questions about the U.S. economy.

The chart presents an interindustry matrix of 99 rows and 99 columns; each of the nearly 10,000 cells in the matrix shows (1) the direct input/output coefficient, (2) the "inverse" coefficient and (3) the interindustry dollar flow for a $1000-billion Gross National Product. The input/output coefficients as published by OBE have been recomputed by the Harvard Economic Research Project to reflect gross domestic output. The 370 sectors of the detailed tabulations have been selectively aggregated to 99 sectors to provide maximum feasible detail for the wall chart. Where the ratio of input to output exceeds 1/100, the cell is tinted in the color-code of the industrial bloc from which the input comes. This device, combined with triangulation of the matrix, brings the structure of interindustry transactions into graphic visibility.

Offprints of five SCIENTIFIC AMERICAN articles on the technique of input/output analysis accompany the chart. The articles are:

Input/Output Economics
by Wassily W. Leontief

The Economic Effects of Disarmament
by Wassily W. Leontief and Marvin Hoffenberg

The Structure of Development
by Wassily W. Leontief

The Structure of the U.S. Economy
by Wassily W. Leontief

The Economics of Technological Change
by Anne P. Carter

W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, California 94104
58 Kings Road, Reading RG1 3AA, Berkshire, England

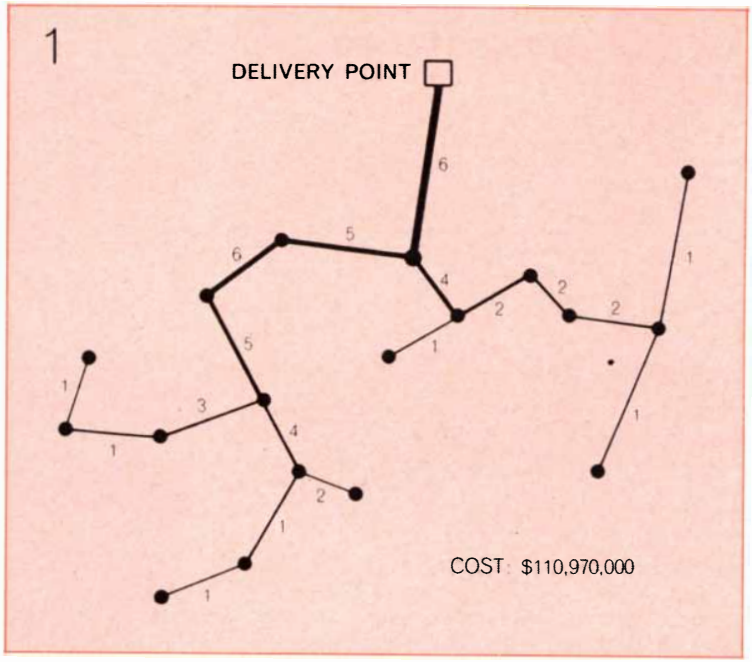
I enclose $_____. (California residents please add sales tax.)
Send me, postpaid, _____ input/output wall charts at $10 each, plus the offprints listed.

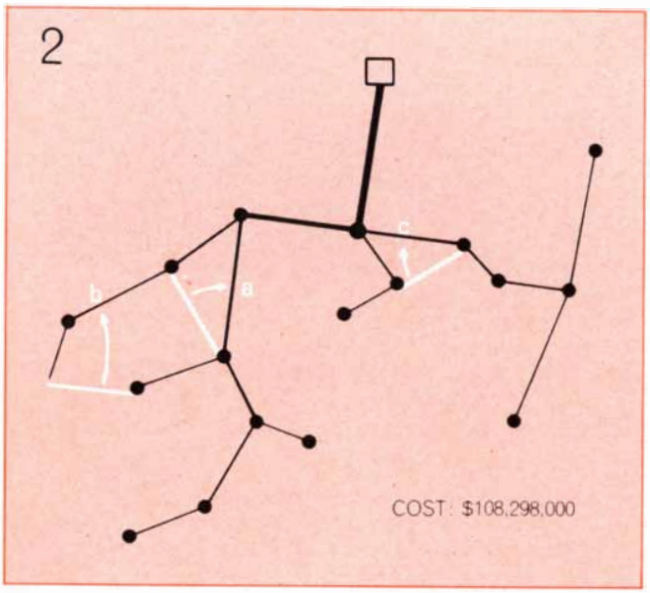
Name_____

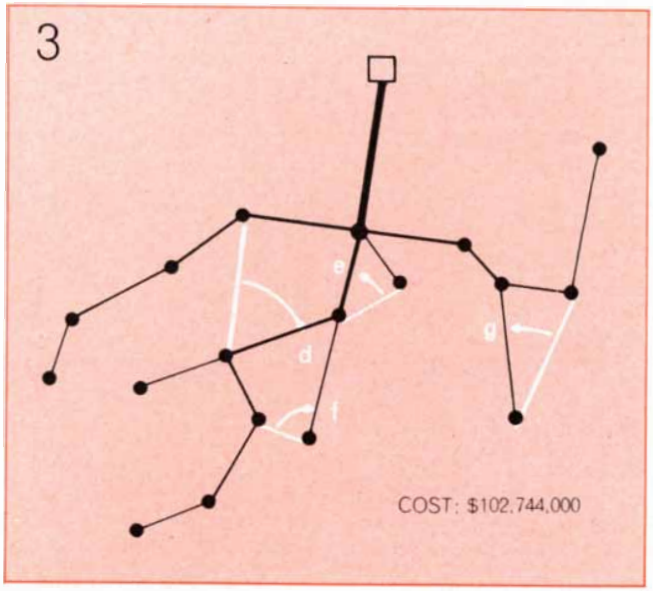
Company_____

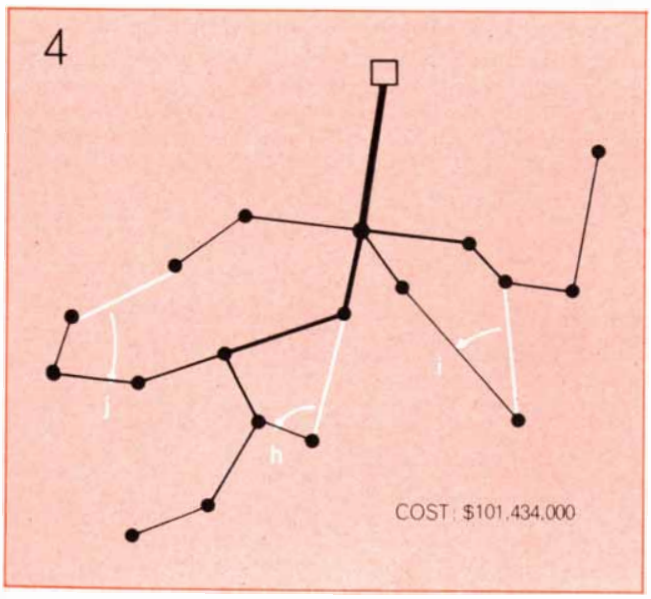
Address_____

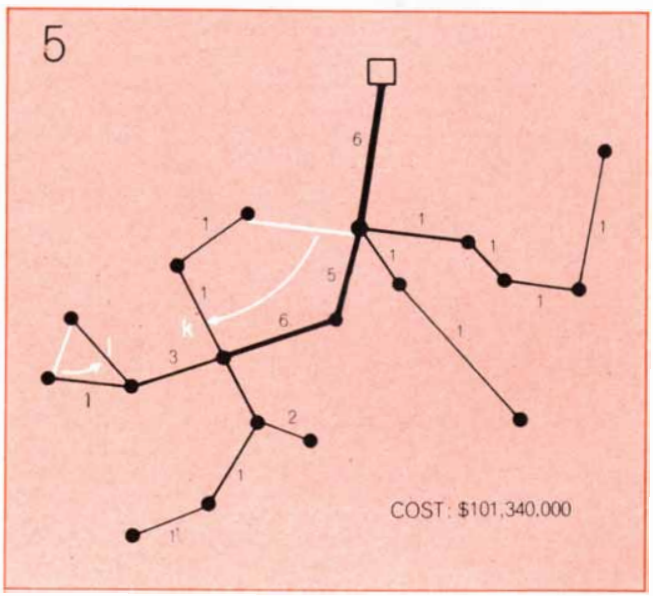
_____











critical path be improved so that the pressure-squared difference will be reduced? It turns out that it can, by assigning the next-largest pipe diameter to b_2. Since this pipe is larger, the pressure-squared difference falls. It turns out, however, that b_1 and b_3 are still the most economic at the smallest diameter, since they cannot be on the critical path.

This calculation is continued. At each stage we find the pipe with the largest pressure-squared difference and increase its diameter. The process terminates when the largest pressure-squared difference occurs across a pipe with the largest available diameter. Further consideration of partial assignments is unnecessary, since such consideration could lead only to greater pipe costs and no possible saving in compression cost at the delivery point. Each set of diameter assignments to the three branches represents an entry on a new list. This list will have no more than 19 entries, and no other partial assignments need be considered.

The advantages of this method are made clear by looking at the alternative, which is the examination of each of the 343 possible partial assignments for the three branches. Our method avoids the necessity of looking at every possible combination. No assignments are overlooked that could be optimal, and those that are not economic are eliminated from consideration without being examined at all. The assignments in the new list are thus candidates for inclusion in the optimal assignment for the entire network. Together these assignments can be regarded as the list of an "equivalent branch" that replaces those branches whose lists were examined.

An analogous method can be developed for creating an equivalent branch for two "series branches" such as b_5 and b_6. These two methods can be combined and written in the form of a computer program so that the most economic diameter assignments for an entire network can be determined in less than a second.

BRANCH EXCHANGE is a method for selecting the most economic network, which involves substituting a new branch for an old one and recomputing pipe diameters (*numbers indicate relative sizes*) throughout the network to see if cost has improved. In *1* the network would have cost $111 million to build. In *2* the cost is reduced as new branches *a*, *b* and *c* replace old branches (*white*). In *3* pipe *d* replaces pipe *a*, which was added in *2*. In *4* more changes are made, and in *5*, after two final alterations, the most economic form of the network is achieved. The numbers beside branches indicate final relative pipe diameters for the network.

Thus far we have not discussed the problem of obtaining a good layout for the gas-pipeline network. Here is an example of a general problem in network analysis. Given a set of nodes and their locations, it is important to connect them by branches in order to achieve a specified design goal. In view of the many different types of network that may exist, it is remarkable that there are some powerful design techniques that are almost universally applicable. One of the simplest of these methods is called a branch exchange. The method consists in modifying an initial network design by repetitively removing branches and replacing them with the same number of new branches in new locations.

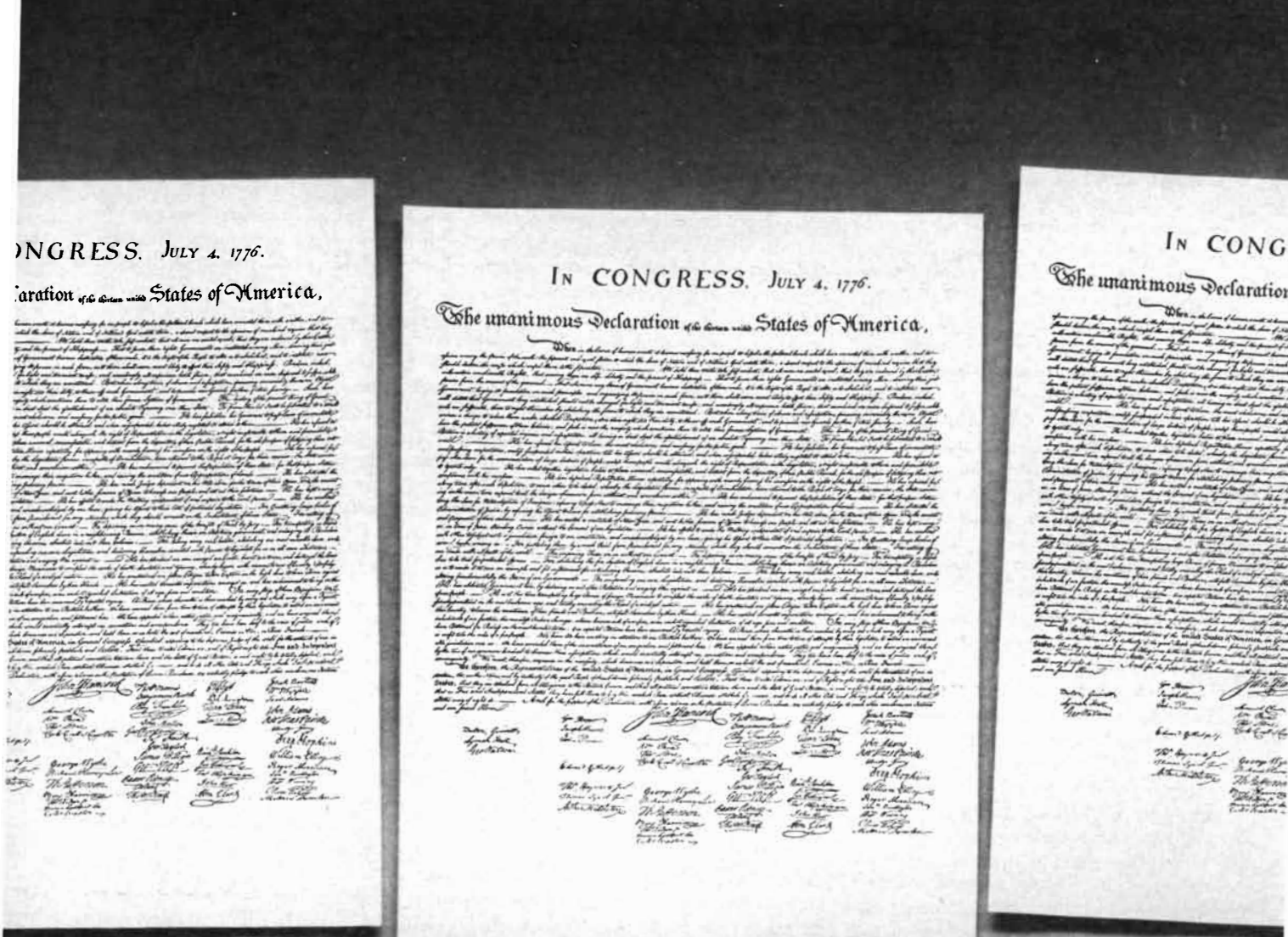
To illustrate this method let us first reconsider the vulnerability problem. Suppose we are given a set of nodes and a fixed number of branches. We want to construct a network in which the number of branches in the smallest cut is as large as possible. We begin with a network in which the smallest number of branches at any node is as large as possible. Such networks are easily constructed. C. Jack Edmonds of the University of Waterloo in Canada has shown that the size of the smallest cut can be maximized by successively removing the branches between two sets of nodes and then reconnecting them so that the new branches are added to the existing smallest cut. In another application of this kind of transformation Hakimi has shown that for a given set of nodes one can generate all possible trees from a given tree by repeatedly removing a single branch and replacing it with a new branch such that each resulting network is still a tree.

The problem of designing gas-pipeline networks is difficult because of the many configurations that can be used to connect a set of gas fields to a delivery point. If pipe cost were the only significant factor, one could select a network that used the smallest total length of pipe. Diameter combinations, however, must also be calculated in order to save compressor costs. The total cost of one such network would have been $110.9 million, even with the pipe diameters optimized. It was still possible to reduce the cost of this network by generating new networks from the old ones and then recalculating the optimal combinations of pipe diameters.

The computer program used in this instance begins by taking any node in the tree and connecting it directly to one of its three nearest neighbors by a new branch. The program then in effect ignores each of the original branches in the tree, one at a time, and recalculates the optimal diameters throughout the new network. This process is repeated for the other two nearest neighbors. The new branch yielding the best reduction in total costs is retained and replaces the old branch. The entire process is repeated for all nodes in the tree. Using this method, the computer is capable of making a number of surprising improvements. In this particular instance a shorter (and therefore cheaper) pipe was replaced during one iteration by a longer (and more expensive) one, but the replacement yielded a net saving of more than $2 million in compression costs. A new line that was added to produce a saving was also replaced in a later and even more economic version of the network. The final network generated by this program cost $10 million less than the original one would have, and yet it has the identical performance.

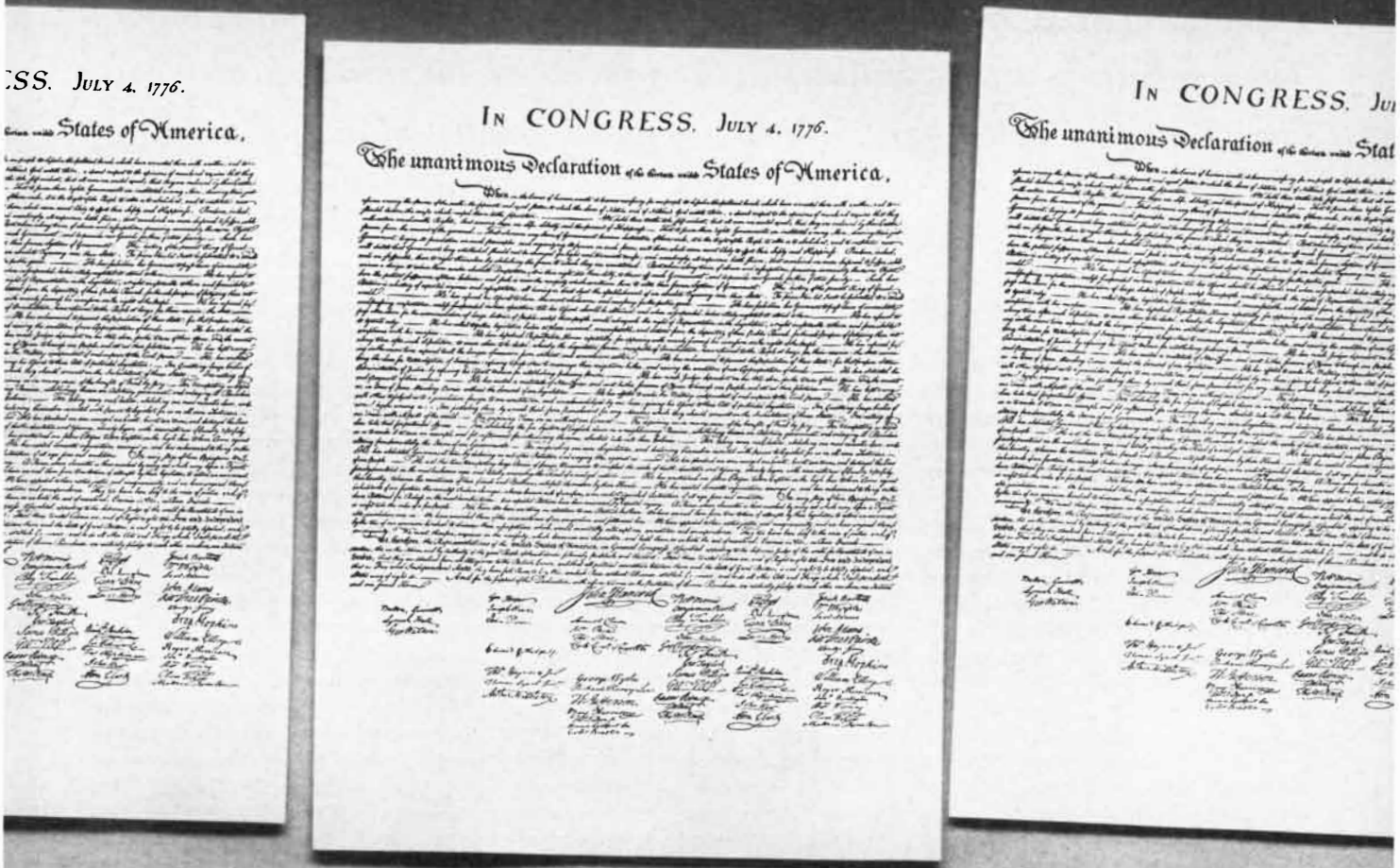
When our computer program and similar ones were applied to proposed gas-pipeline networks for the Gulf of Mexico, more than $200 million of construction was examined and redesigned so that a saving of between 25 and 50 percent was made. One proposed system that originally cost $82 million was redesigned to double the gas reserves it will tap, yet its cost was reduced by $7 million. The natural-gas pipeline industry now uses this design method as a standard technique, and the Federal Power Commission has adopted it as a major tool in analyzing the economic efficiency of proposals brought before it.

A number of general conclusions can be drawn from the problems we have discussed. There are fundamental techniques such as the study of flows that lead to powerful results useful in the study of any network. Moreover, there are universal problems, such as vulnerability and reliability, that are common to all networks. Each problem, however, has its own peculiarities that demand new insights and novel approaches. In the study and design of systems such as gas pipelines, for instance, a great many factors such as pressure, flow and cost complicate the picture. Nevertheless, a remarkable property of the methods we have described is that they do not depend on the particular equations involved and hence can be applied to a wide range of problems.



We believe all copies

XEROX IS A REGISTERED TRADEMARK OF XEROX CORPORATION.



should be created equal.

We believe your copies—from first to last—should be as crisp and clear as the original.

With a Xerox machine, they are.

Where the original is white, the copies are white. Where the original is black, the copies are black. The "i"s stay dotted and the "t"s stay crossed.

And because the copies are made on ordinary untreated paper they'll last as long as the paper.

Thanks to this consistent quality, nobody has to waste time sorting good and not-so-good copies according to office rank. Or throwing away bad ones and trying again.

So everyone around the office gets the word faster. And can act on it that much sooner.

All things being equal.

XEROX

Population Trends in an Indian Village

Population growth is beginning to slow down in villages of the Punjab. The analysis of such variables in the process as education, health care and economic status may influence India's family-planning program

by Carl E. Taylor

There is reason for optimism about the demographic future of India. To be sure, most demographers are gloomy about India's population problems, and when I consider the overall statistical projections, I tend to share their concern. Then I come back to our research base in the Punjab, in northwestern India, and feel the stirrings of social and economic development. I note that birth rates have begun to decline in the Punjab and at the scattered sites of demonstration projects in other parts of the country. I realize that recent progress has been dramatic, and then even the vastly accelerated program that is still essential begins to seem possible.

By taking a village view of India's population problem one can look beyond the slowly changing demographic statistics and examine the basic processes of local change. To be sure, our village of Narangwal is not typical. It is in Ludhiana District, which leads the Punjab in agricultural production, and the Punjab is India's leading state economically. Above all, Narangwal is atypical in that it has begun to change. It may nevertheless serve as a prototype for what can happen and so provide a basis for predicting a pattern of development for hundreds of thousands of villages in many parts of the world.

More than two-thirds of the world's people live in rural villages. It is in the villages that the population problem is most acute; it is there that massive population increases are bound to occur during the transition from the traditional balance of high birth rates and high death rates to the much lower balance associated with economic development. It is in the villages, however, that organizers of family-planning programs have been most frustrated—by the lack of transportation and communication, by the villagers' low educational level, traditional cultural orientation and resistance to social change and by the reluctance of technicians and professionals to live and work in rural surroundings.

Some policy-makers have tended to turn their attention to the cities, where it is easier to work, in the hope that urban influences will seep out into the rural areas. On the basis of Western experience with the shift of population from rural to urban areas, however, it is unreasonable to expect that birth rates will fall spontaneously if efforts are concentrated in cities, where the population is going to be, rather than in the villages, where the people are now. The demographic realities are too urgent to allow this relaxed and essentially defeatist approach. Countries such as India start with massive rural population densities as well as cities that are already among the largest in the world. Anyone who has seen Calcutta's swarming slums realizes that to encourage migration from the countryside to the cities is to promote disaster.

Rural population problems can be solved by concentrating on the villages and by encouraging and utilizing fundamental changes in living patterns. Just as early workers in malaria control had to "learn to think like mosquitoes," those responsible for population policy must learn to think like villagers. The national and international considerations of economic planners and administrators have little relevance at the village level, where motivation is highly personal and the family largely controls the pattern of life. It is in the privacy of individual homes that the night-by-night decisions are made that really determine whether or not family planning will be practiced, and why.

We set up the Johns Hopkins University Rural Health Research Projects in Narangwal because of our feeling that too many village programs had been run from cities. The institutional base is a rural teaching health center of Christian Medical College in Ludhiana. Research began eight years ago with a five-year study of the rural orientation of physicians. We now have three studies in progress: a functional analysis of rural health centers, an investigation of the interactions of malnutrition and common infections in children during weaning and a prospective study of the impact of health care for mothers and children on the acceptance of family planning. With a research staff of 70 at work in 22 villages and with a population of some 25,000 under close observation, we are gaining a deepening insight into the dilemma of health and population.

A brief review of local village history provides perspective. Narangwal was settled about 350 years ago when the Punjab was overrun by warring military bands. In return for effective fighting in the forces of a local ruler a clan of warriors was given a tract of open land. The land was fertile and their military reputation helped them to organize thriving settlements. Other caste groups were added to form a number of self-sufficient communities, but a steady increase in population led eventually to disputes over land. On a set day about 100 years ago cousins from Narangwal and a neighboring village fought from sunrise to sunset, and the final line of battle was marked by large rocks to stabilize the village boundaries. Successive village maps since that time show the remarkable fragmentation of landholdings in succeeding generations [*see illustration on page 109*].

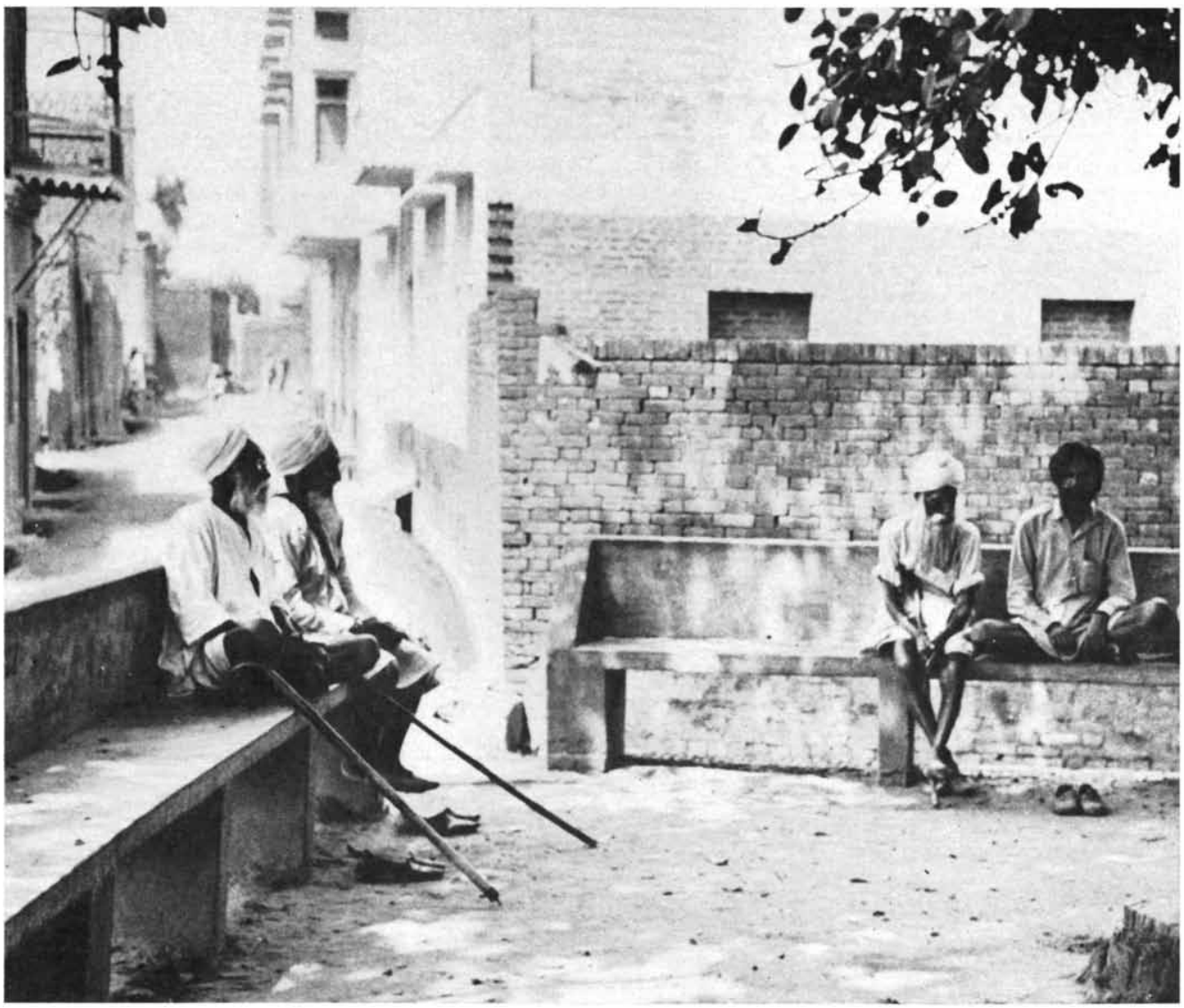
Punjabi villagers are wonderfully strong and hardworking, and the alluvial Indo-Gangetic plain is fertile. In spite of a rigorous climate that ranges from the freezing winter to the 120-degree desiccation of May and June and the monsoons of July and August, agricultural productivity was sufficient to maintain a high population density even with traditional farming methods. Birth and death rates were moderately out of balance at about 45 and 35 per 1,000 of population respectively.

Some 50 years ago the region began to undergo a more rapid demographic transition. The major epidemic diseases such as cholera and smallpox became less of a threat. (The reason is unknown; certainly the health services were not well enough developed to claim much credit.) Pandemics of plague 70 years ago and of influenza 40 years ago killed between a quarter and half of the population of many villages. In the 1950's a national malaria campaign brought under control the most important single cause of illness and mortality.

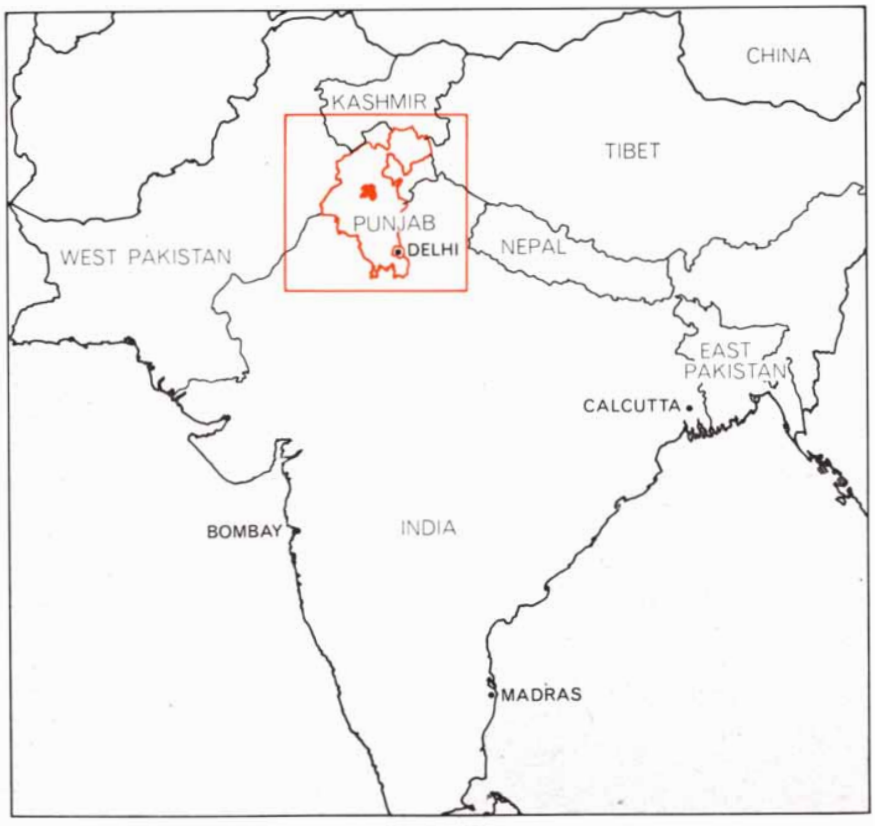
Epidemics are dramatic, but they come and go and therefore do not cause continuing high death rates. The control of children's diseases is more important demographically. In the less developed countries half of all the babies born die in childhood. The synergism of malnutrition and common infections causes not only illness and death but also irreparable stunting, both physical and mental. Particularly in the past 20 years the improvement in death rates has been mainly due to the saving of children. Organized health services cannot take as much credit for the improvement as better economic and nutritional standards and the more intangible benefits of better mothering.

Vital statistics are poor but a reasonable estimate is that 20 years ago the Punjab's rural death rate was more than 25 per 1,000 of population. It has dropped below 20 and will eventually stabilize at about 10 per 1,000. The remaining child mortality is largely among the low-caste Harijans, who constitute a third of the population. The poorer children still suffer from serious nutritional deficiencies and readily preventable infections.

We are not sure just when or why birth rates started to decline. The rural Punjabi birth rate was probably in the low 40's per 1,000 at the time of India's independence in 1947. Precise data



VILLAGE OF NARANGWAL in the Punjab was the site of the program in family planning described by the author. The scene is the Baitthak, the name for the meeting place of the village elders, who have much to discuss in these days of rapid development.



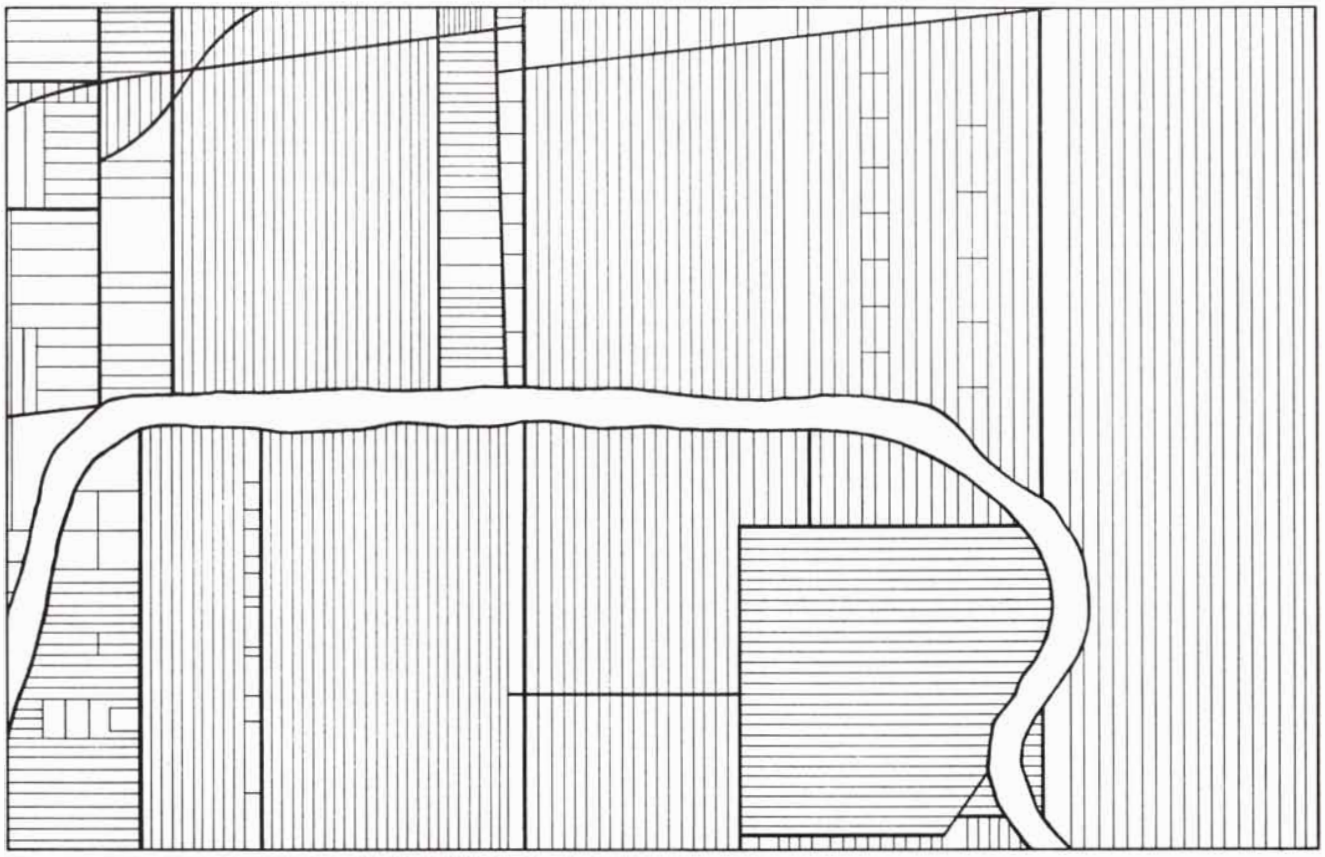
STATE OF PUNJAB in India is the nation's leading state economically. The region where the program of health care and advice on family planning is under way is outlined by the colored rectangle. The same region is shown in somewhat greater detail in the map below.

LUDHIANA DISTRICT leads the Punjab in agricultural production. Hence the state, the district and the village of Narangwal are somewhat atypical of India but suggest what may happen in terms of economic and population trends as the rest of the country develops.

collected by the India-Harvard-Ludhiana Population Study from 1955 through 1959 showed a birth rate of 39 per 1,000. Base-line surveys in Narangwal indicate that the birth rate of the upper-caste two-thirds of the population is probably now about 30, whereas low-caste birth rates still are in the middle 30's, for an overall rate of 33 in 1969. A fall of six points in nine years is more than we had dared to expect. We cannot say yet how much credit the Indian government's Family Planning Program can take for this decline. An important factor is undoubtedly the dramatic economic improvement of the Punjab in the past 15 years. Previous personal observations suggest that when economic conditions started to improve in the 1950's, people first acted on the assumption that they could afford to have all the children they wanted. That phase seems to have passed. People talk increasingly about the desirability of having two educated children rather than large numbers of farmhands.

Anyone who feels that villages are too stagnant to develop should experience the dramatic and pervasive social and economic revolution now taking place in the Punjab. The extension program of the Punjab Agricultural University has contributed to the "green revolution," with support from the U.S. Agency for International Development. At first it was only the rich farmers who profited, but now the benefits are reaching all economic levels. Wheat production has increased by two to four times in a few years. Tube wells powered by the electric-power grid from the Bhakra-Nangal Dam provide regular irrigation; the irrigation makes up to three crops a year feasible instead of one or two, and that means more diversification. Tractors and simple farm machinery are increasingly used because of the labor shortage and larger crops. Wages are three times as high as they are in neighboring states, so that migrant labor moves into and out of the area, diffusing change. Small entrepreneurs display inventive talent in setting up workshops for agricultural equipment; a single electric motor may have separate power takeoffs for operating a cotton gin, a corn-shelling mill, a flour mill, a fodder chopper and a grinder.

The rapid changes in agriculture have brought a greater readiness to modify other patterns of life and work. Industrial development in Ludhiana attracts commuter traffic as concentrated as in any American urban area (but on bicycles), and the commuter rush is reaching



FRAGMENTATION OF LANDHOLDINGS in Narangwal is one of the economic problems that retarded agriculture until recently. This map is based on one made 100 years ago; the heavy lines show the original fields, which averaged about 10 acres in size, and the lighter lines show the subsequent fragmentation. Improved agricultural efficiency resulted from a land consolidation 10 years ago.

Narangwal, 15 miles away. Many more village stores retail mass-produced items for better living. A constant stream of bicycle peddlers moves from house to house with wares such as groceries, cloth and brass utensils.

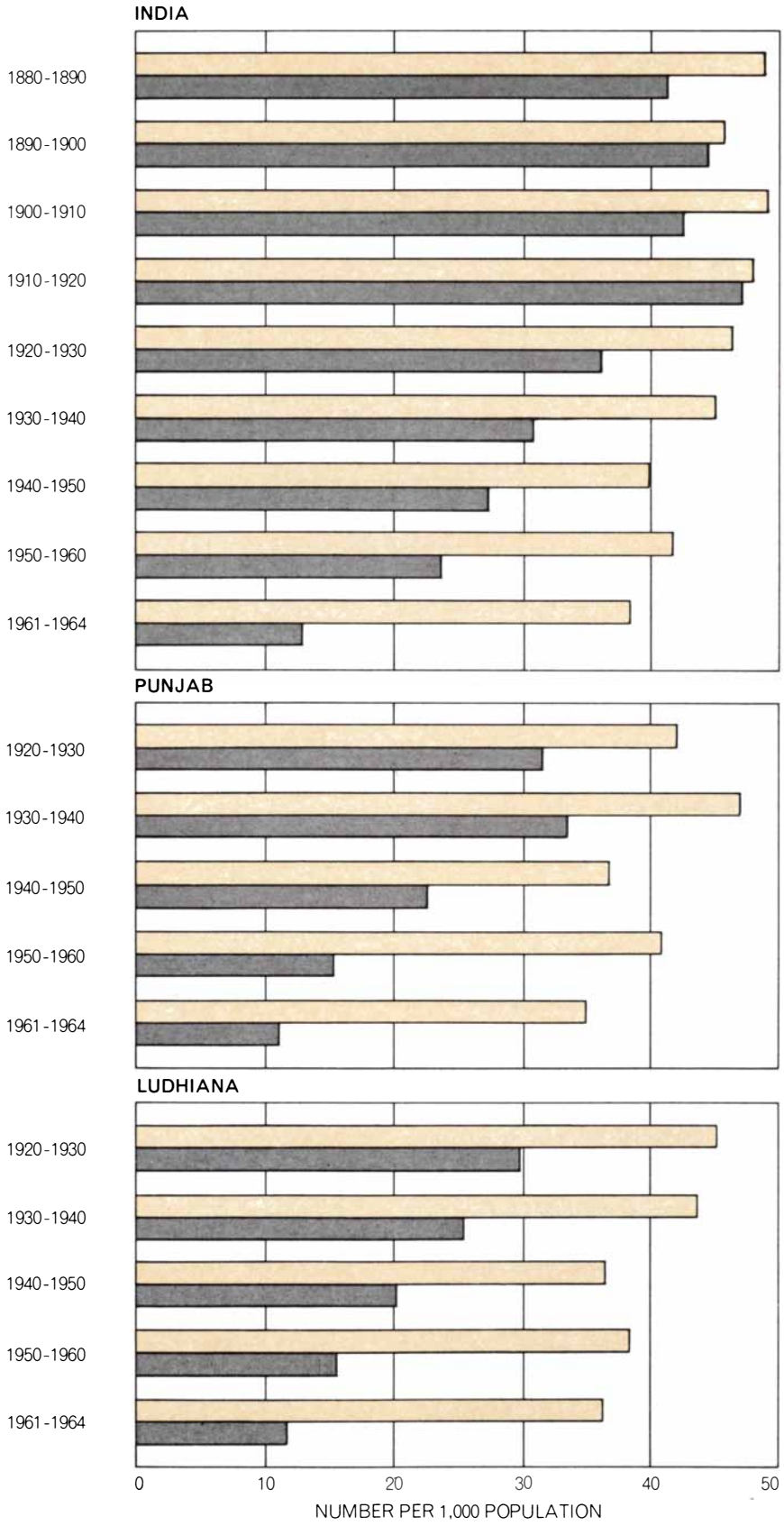
Underlying this material improvement is a general feeling of economic adventure. Villagers are beginning to view change as a good thing. Narangwal, with a population of fewer than 1,500, has both primary and high schools, and two years ago a rural college offering a B.A. degree was founded with local financing. It is becoming commoner for village boys and girls to say they expect to live here. Narangwal is still a village, but a rapidly improving one; its experience suggests that village development may in fact lead to a better pattern of life than urban development.

Although one of our tasks has been to record the changing ecology of village living as it affects family planning, we have focused our research and our action program on another hypothesis: Where death rates are high the fear of parents that a high proportion of their children will die is among the most important motivational blocks to the acceptance of family planning. This subconscious concern is based on the villagers' own childhood experience and what happened to their brothers and sisters and friends. A slow demographic transition provides time for spontaneous recognition, over one or two generations, of the increased survival rate among children. Where there is a precipitously falling death rate, however, deliberate measures combining birth control and death control are needed to reduce the lag between the change in the death rate and the change in attitude toward family size.

In any traditional society about 10 percent or more of the women share with quiet desperation a feeling that they must not bear more children—whether for health or economic reasons or because they simply do not want the "bother." They come eagerly to try any new contraceptive method that can be substituted for their traditional methods, taboos and abortions. Most women, however, have ambivalent and tentative motivations. They are greatly influenced by reports of other village women's experiences and they want methods and facilities that are safe, cheap, aesthetic, accessible and easy to use. These considerations of quality of care and convenience most influence the young women who have had few children; they have the highest potential fertility and are therefore demographically the most important. For them the integration of family planning with health services is almost essential. In these families family-planning education for men and the availability of male methods of birth control are also important.

Our research project is trying to demonstrate experimentally that improved care for children will give parents enough confidence so that they will begin to practice family planning when they have only one or two children rather than continuing the present tendency to wait until they have at least four. We are also testing the acceptability of family planning as a routine part of postpartum care for women, something that has been well documented in urban hospitals but not as yet in villages. All of this service is provided by paramedical auxiliaries in a pattern that is suitable for general application in India.

For the past several months we have



POPULATION TRENDS are charted for India (*top*), the Punjab (*middle*) and Ludhiana District (*bottom*). Growth is indicated by the difference between births per 1,000 population (*color*) and deaths per 1,000 population (*gray*). Official records are grossly underreported, but corrections based on census data have been applied to the all-India figures up to 1960. Punjab and Ludhiana figures are uncorrected; rising rates in early years are probably due to improved reporting. In Punjab and Ludhiana birth rates seem to be starting down, but best evidence for this comes from records of individual villages, as described in text.

followed two pairs of twins who illustrate some of the basic problems. They were born at about the same time on opposite sides of a dusty lane in the poorest low-caste section of one of our villages. When our expanding child-care program picked them up at the age of a year and a half, each weighed less than 10 pounds; the one girl weighed less than eight pounds. They had overt marasmus: a wasting condition characterized by wizened, shrunken features, like those of tiny old men, and sluggish matchstick extremities. The condition (which the villagers call *sukha,* or "drying up") starts with the common practice of "breast starvation." Village mothers think that as long as the breasts produce any milk they need offer no other food until children can start the coarse adult diet, and they also prefer prolonged breast-feeding because they know it is associated with decreased fertility. The synergism of the resulting undernutrition and common and normally nonfatal infections during the weaning period is our most frequent cause of death. These babies had had recurrent diarrhea and respiratory infections.

At first both mothers cooperated in our feeding program, and the responses were encouraging. Even more dramatic than the weight gain of up to a pound a week was the sudden infusion of liveliness. Then both mothers reported that they had missed menstrual periods. At that point we made a tactical mistake: as a pregnancy test, we tried a drug that precipitates menstruation in a woman who is not pregnant. Only mother *A* responded. On being told she was pregnant, mother *B* accused us of having given mother *A* a more effective aborting agent; we could not convince her that she had not been discriminated against. With another child coming, she lost all interest in the care of her twins. Their weight fell. They developed recurrent diarrhea, and with each bout their growth curves dipped lower and their susceptibility to repeated infections increased; we diagnosed primary tuberculosis in both of them. All this time the twins of family *A* were thriving. The mother enthusiastically accepted long-acting progesterone injections, a new family-planning method we are testing.

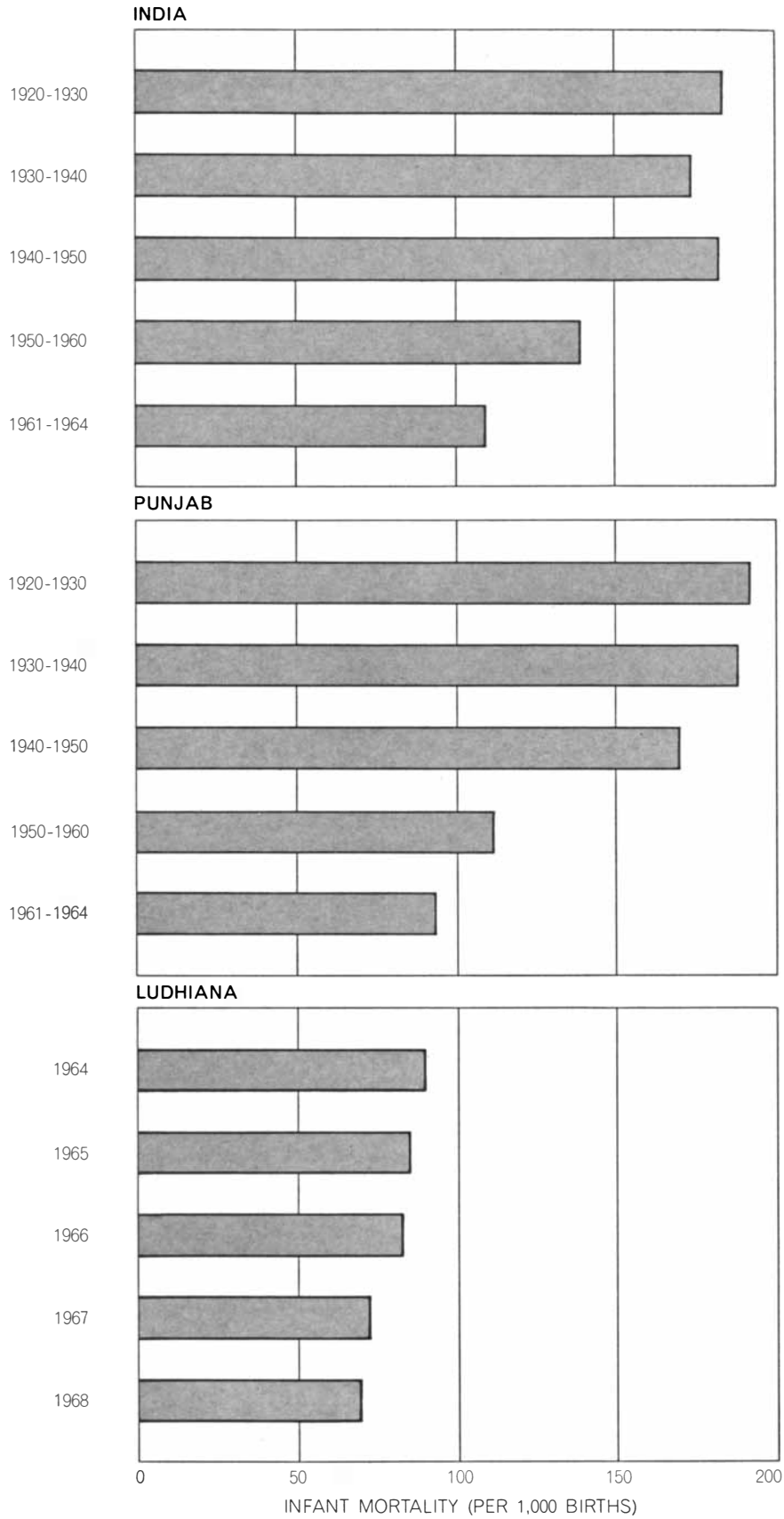
Finally we persuaded the parents of family *B* to admit the marasmic twins as inpatients in the health center, and we started an intensive-care program with one auxiliary nurse-midwife in constant attendance. The response was dramatic; the twins soon went home healthy. The mother decided that if we were willing to do so much for her babies she would

care for them. Moreover, she agreed to be sterilized after her baby was born, and some of her friends have been influenced by her example.

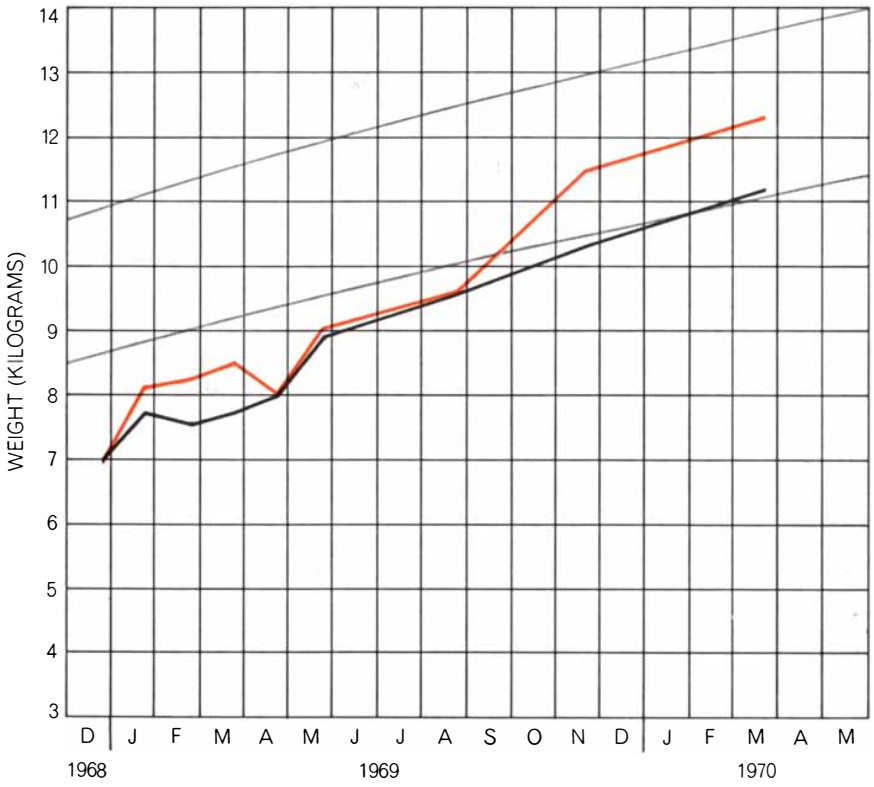
When experts engaged in international population programs meet in conferences and committee sessions, there often arises an artificial polarization of views: "family planning" v. "population control." It is now evident that a phased approach is needed. Family planning to serve the individual family is appropriate at the beginning stages of program development; social and legal pressure for population control as a national objective should be applied only after family planning is already available. Unless one has lived through the ambiguities of fieldwork, there is a natural tendency to oversimplify the issues in a search for quick answers. When I first started to work on family planning in India 20 years ago, the solutions seemed to me straightforward and simple. (I was like the enthusiastic young missionary, starting work in a centuries-old village, who wrote home that he had arrived "just in the nick of time to prevent disaster.") The already motivated women eagerly coming to our family-planning clinics led me to think that just providing contraceptives would solve the problem. It was only as we tried to get total village acceptance of family planning that we realized the inherent difficulty of changing people's living patterns.

On the other hand, it is wrong to take the opposite view: that people's responses are hopelessly slow and that some form of compulsion is needed. Any form of coercion is unwise and is unacceptable in developing countries, where politically shaky governments do not dare risk interfering too blatantly with the right to procreate. It seems particularly unreasonable for international consultants to propose economic or legislative constraints in developing countries until they can show that such provisions can be enforced in their own countries. And it is irrelevant to talk of changing tax, educational or social provisions favoring more children because most rural people in developing countries do not now get tax or other benefits for having children.

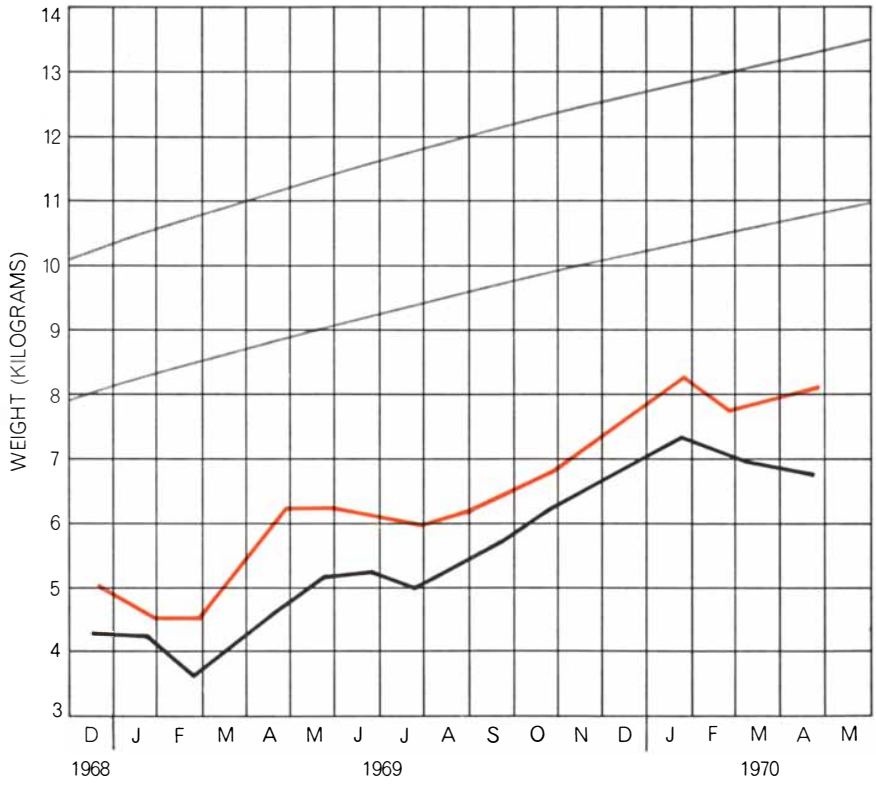
Recent shifts in foreign-assistance policy have been particularly confusing to recipient countries. What has most disturbed some developing nations is a particular sequence of events: a reduction of support for health programs and then a push for family planning. In the early 1960's, for example, the Rockefeller Foundation abdicated from its long lead-



INFANT MORTALITY has moved generally downward in India (*top*) since 1920 and steadily downward in the Punjab (*middle*) and Ludhiana District (*bottom*). Again the data are underreported, with fluctuations being at least partly due to changes in reporting. In each case infant mortality is expressed in number of deaths per 1,000 births. When larger numbers of children survive, parents presumably reduce their estimate of how many children they should have. The Johns Hopkins University Rural Health Projects try to shorten the one-generation gap in this transition by integrating child health and family planning.



GROWTH OF TWINS suffering from the effects of malnutrition and repeated infections improved steadily after they entered Johns Hopkins child-care program at age of 18 months. The twins were boys. Gray curves show upper and lower limits of normal weight gain.



ANOTHER SET OF TWINS entered the child-care program at the age of 13 months and were withdrawn at the age of 18 months, when their mother discovered she was pregnant. The weight of the twins declined once again for some months until hospitalization started a slow improvement. Tuberculosis was diagnosed in both twins at the age of 26 months. In this case the twins were a boy and a girl; the girl consistently weighed less than the boy.

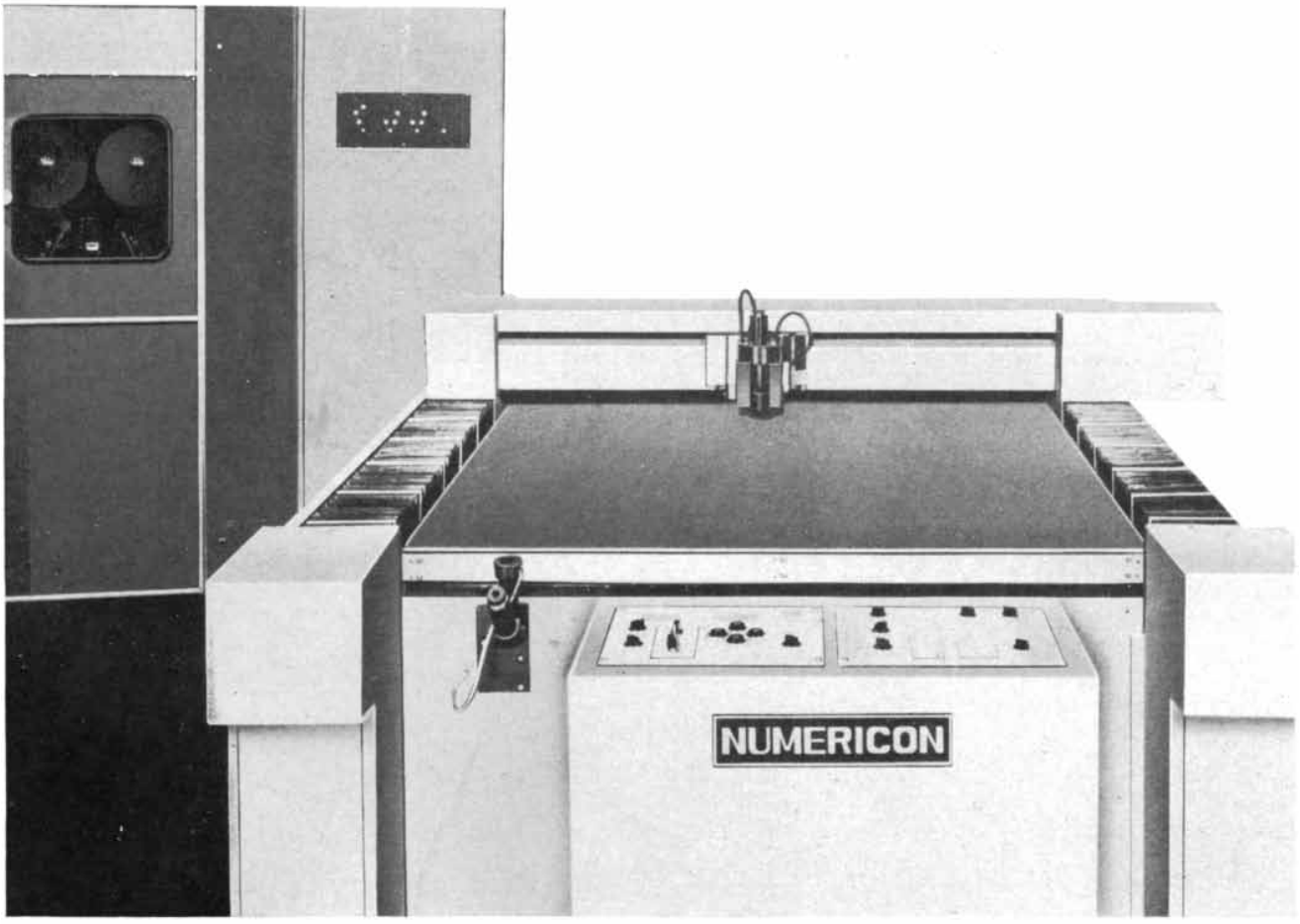
ership in health work and simultaneously announced that its money was supporting the Population Council and that its own efforts would be shifted to agriculture. At the same time, the Ford Foundation proclaimed a major investment in population programs while publicizing its determination to stay out of health work. These moves apparently were at least in part a reaction to accusations that health activities had helped to precipitate the world population crisis. It would have been much better, it is now clear, if the Rockefeller Foundation had taken advantage of its unique potential for developing integrated services for health and family planning. The earlier contributions of the Rockefeller Foundation have not yet been forgotten in developing countries, and they would provide a base for the foundation's moving again into a position of leadership.

U.S. Government assistance has followed a similar erratic course. The health activities of U.S. foreign aid agencies were among their most popular and effective efforts. Beginning in about 1960, however, a growing awareness of the world population problem coincided with a progressive reduction in foreign assistance for health activities. After a lag of a few years there were strong pressures to finance family planning. It is not surprising that some African countries voiced the same concern about "genocide" that is heard in urban ghettos. The new insistence of recipient countries on combined programs of health and family planning is politically wise, and the Agency for International Development is now moving strongly in that direction. The World Health Organization, after prolonged hesitation because of religious constraints in some of its member states (including the U.S.), is now assuming active leadership in the effort to integrate family planning and health services.

Tremendous advances have already been made in India's Family Planning Program. Awareness of the idea that the government thinks two or three children per family is enough has spread widely. Most of those who were initially ready to accept family planning have been reached by mass campaigns in which mobile units and "camps" make sterilization and the fitting of intrauterine devices (IUD's) available intermittently. Now the more ambivalent parents must be reached, and this requires the development of locally available services and more general educational programs.

Education becomes a promising approach once it is recognized that the population problem will be with us for

A total geometric data processing system for today's complex design-manufacturing operations.



Numericon

Mutoh Industry introduces a revolutionary system of geometrical data processing for design output and manufacturing input. A proven time-saving, money-saving automated method of planning and design for most every field of industry. An entirely new dimension for man's imagination.

Numericon revolves around your existing automation system, consists of a Digigrammer coordinate digitizing system capable of making two-dimensional and three-dimensional measurements, a Numericon automatic drafting system which draws two-dimension drafts, and DAP-1, the software which links them.

The Digigrammer transforms man's three-dimensional design vision into digital values, the information fed into the computer as input data for processing. Your computer, meanwhile, makes the calculations for the plans and designs—and this output data becomes the input data for any of a variety of flexible, adaptable Numericon Automatic Drafting Systems now available.

Numericon applications today are wide—and proven successful. In shipbuilding, automobile and aircraft manufacture, bridge and building construction. For precise designs and master masks of intergrated circuits and printed circuits. For complex architectural and civil engineering projects. Even for analyzing management data.

If you think that Numericon can help you solve a problem, we'd like to share our detailed specifications with you. Write or telephone today: Consul & Mutoh, Ltd., 519 Davis Street, Evanston, Illinois, 60201. Telephone (312) 869-9545.

MUTOH

MUTOH INDUSTRY LTD.

1-3, Ikejiri 3-chome, Setagaya-ku, Tokyo, Japan Telephone: (03) 460-8111 Telex: 242-2575 Cable Address: DRAFTERMUTOH TOKYO

a long time. In addition to increasing the general understanding of population dynamics, a comprehensive effort is required to prepare people to live in a crowded society in which smaller families will become the rule. Among the many social adjustments are basic changes in the economics of family work distribution. For example, who will take the buffaloes and cattle out to graze when there are no longer lots of boys who do not go to school? How can the smaller family provide someone to bring in water and fuel—tasks that once took the time of one person in five?

The government's primary health centers in the Punjab are already devoting approximately half of their effort to family planning. Each serves some 80,000 people through four to eight subcenters. They can realistically cover less than a tenth of such a population, however. A basic requirement is to increase the number of subcenters to more than 30 in each area, so that one auxiliary nurse-midwife will serve about 2,500 people rather than from 10,000 to 20,000. Work patterns will need to be changed drastically, and that means modifying the training program. The professional staff's roles should be shifted; physicians and nurses, instead of trying to do everything themselves, should provide supervision of the mass care given by auxiliaries.

A village woman said to one of our family health workers recently: "I will have a sterilization, but only because you are living right here in the village and will be available to help if I get into trouble." Intrauterine-contraceptive-device cases in particular have to be followed because minor complications are common. (Our surveys show that 65 percent of Indian women now express great fear of "the loop" because of exaggerated rumors about bleeding and other problems. In much of India the loop has become a negative symbol of the excessive pressure that was exerted to recruit women for IUD camps.) This experience indicates that personnel with appropriate skills should be readily available in the villages before "the pill" or newer methods are introduced.

A chronic problem in rural health care has been the reluctance of the medical profession to permit paramedical auxiliaries to diagnose and treat the 90 percent of illnesses that are common and simple. An auxiliary who provides care opens up many channels of cooperation with the patient, and we have demonstrated that in routine cases effective medical care can safely be reduced to standing orders or written directions giving clear-cut routines, with referral of atypical or problem cases to staff physicians. Such a shift in responsibility would enable physicians to devote time to preventive services for health and family planning and to participate in community action and education, functions that require more flexibility than routine medical care and are harder to perform well.

It is in village subcenters that the most constructive steps can be taken to solve India's population problem. In order to improve the integration of family planning and health services, more than 200,000 new subcenters are needed in India. Providing the minimum package of maternal and child care necessary to make family planning acceptable is economically feasible. And it will give the greatest benefit-to-cost payoff of any development investment.

MASS MEETING is held at the Narangwal health center to discuss health and family planning. At this meeting the Minister of Health of India and the Chief Minister of the Punjab addressed the people of the village about their participation in the planned program.

CHEVY'S NEW LITTLE CAR: IF YOU LIKE THE 1971, YOU'LL LIKE THE 1975.

If you've been with us for the past few weeks, you know we're ready with our own little car, the Vega 2300. Almost.

But first, there are a few things we've been meaning to talk to you about.

No shiny new ashtray knobs.

Before we tell you anything about the car itself, we think you should know that once it comes out, it's going to stay out. We don't plan to change it for at least four years.

We've got it just the way we want it, and we think you'll like it.

Of course, there is the possibility that we'll find ways to improve Vega from a functional standpoint. If we do, we will. We'll make you a promise, though: no change for the sake of change.

So, when you see the 1971 Vega, you'll be getting a preview of Vegas to come. That's how confident we are. Here's why.

Little, but big.

To be perfectly honest, we're pretty proud. Our little car is unlike any other little car.

For one thing, it is indeed little: it's on a tight 97-inch wheelbase, and it's built for four passengers. Yet it feels bigger, because there's as much room per passenger as there is in many big cars.

There are other things that make our Vega a lot of little car, like its zippy performance, its quiet ride, and even its taut, smooth handling. In fact, our little car is more than just maneuverable—it's plain old fun.

Naturally, all of these things are ads in themselves, so stay tuned to this magazine.

Little, but little.

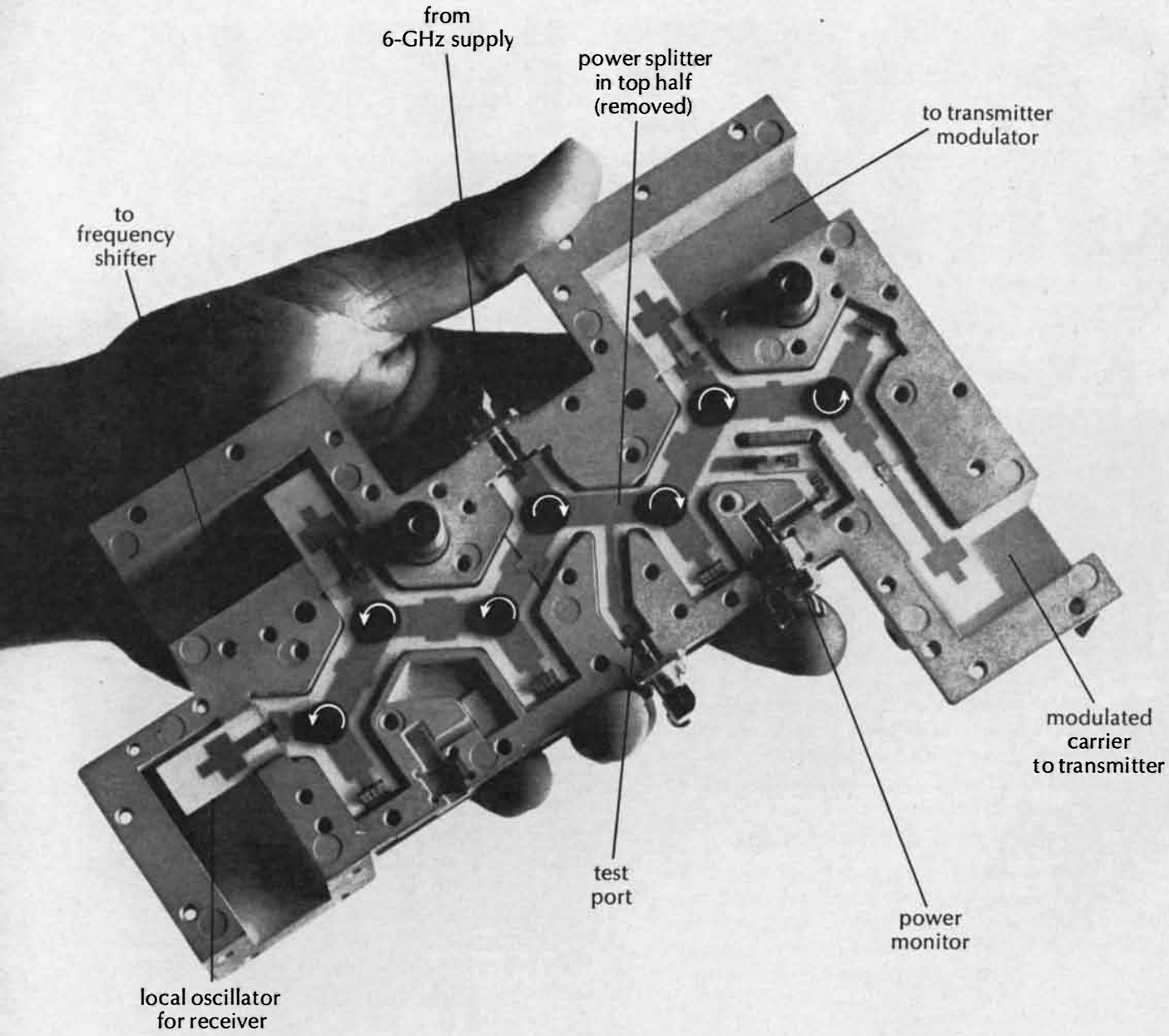
Although our little car feels and acts like a much bigger car, there are times when its littleness really shines.

Like when you pull into a gas station. We aren't ready with final figures yet, but we can say this much: Vega will get gas mileage in the same neighborhood as the best of the economy cars. And that's a pretty good neighborhood.

Our little Vega will also seem very little when you go to buy it: prices will be very close to ordinary little cars'.

GM
MARK OF EXCELLENCE

VEGA
CHEVROLET



Miniature crossroads for microwaves

At every repeater station of a microwave relay system, there's a microwave distribution network—a circuit assembly that combines, divides, and directs the signals of one transmission channel. It interconnects the waveguides with coaxial cables that distribute microwave power to frequency mixers, modulators, and amplifiers.

Now, engineers at Bell Laboratories' Allentown, Pennsylvania, location have developed an integrated-circuit version. This one structure, smaller than a cigar box, has only a tenth the weight and a fifteenth the volume of the previous assembly. And, it costs less.

The network is shown above with its top half removed. The paths for the microwave signals are "stripline"—small rectangular channels with a copper-strip center conductor, electrically much like coaxial cable. The conductor strip is plated over an evaporated thin gold film on a ceramic substrate. Terminations and resistors are made by depositing tantalum nitride on the substrate. The four cross-shaped stubs (at the ends of the stripline) are stripline-to-waveguide transducers.

The seven black disks on the center conductor are ferrite microwave circulators, three-port devices which let microwave power flow from any port to the next one in the indicated direction only. This controls signal flow and isolates circuitry. The power splitter in the conductor feeds the test port.

Bell Laboratories engineers and their colleagues at Western Electric carefully selected this combination of modern materials and the techniques for working with them—including precision aluminum die casting and tantalum and gold thin-film technology. Analytical studies defined the geometry of the various circuit components to meet the rigorous standards of long-distance communications. This resulted in a superior component for our radio relay system and, at the same time substantial reductions in cost, size, and weight.

From the Research and Development Unit of the Bell System:

Bell Labs

MATHEMATICAL GAMES

Diophantine analysis and the problem of Fermat's legendary "last theorem"

by Martin Gardner

"All right," said Simon. He took a deep breath. "My question is this: Is Fermat's Last Theorem correct?"

The devil gulped. For the first time his air of assurance weakened.

"Whose last what?" he asked in a hollow voice.

—ARTHUR PORGES,
"The Devil and Simon Flagg"

An old chestnut, common in puzzle books of the late 19th century (when prices of farm animals were much lower than today), goes like this. A farmer spent $100 to buy 100 animals of three different kinds. Each cow cost $10, each pig $3 and each sheep 50 cents. Assuming that he bought at least one cow, one pig and one sheep, how many of each animal did the farmer buy?

At first glance this looks like a problem in elementary algebra, but the would-be solver quickly discovers that he has written a pair of simultaneous equations with three unknowns, each of which must have a value that is a positive integer. Finding integral solutions for equations is today called Diophantine analysis. In earlier centuries such analysis allowed integral fractions as values, but now it is usually restricted to whole numbers, including zero and negative integers. Of course in problems such as the one I have cited the values must be positive integers. Diophantine problems abound in puzzle literature, and many have appeared in this department. The well-known problem of the monkey and the coconuts (discussed here in April, 1958) and the ancient task of finding right-angle triangles with integral sides (October, 1964) are among the classic instances of Diophantine problems.

The term "Diophantine" derives from Diophantus of Alexandria. He was a prominent Greek mathematician of his time, but to this day no one knows in what century he lived. Most authorities place him in the third century A.D. Nothing is known about him except some meager facts contained in a rhymed problem that appeared in a later collection of Greek puzzles. The verse has been quoted so often (see Oystein Ore, *Number Theory and Its History,* page 180) and its algebraic solution is so trivial that I shall not repeat it here. If its facts are correct, we know that Diophantus had a son who died in his middle years and that Diophantus lived to the age of 84. About half of his major work, *Arithmetica,* has survived. Because many of its problems call for a solution in whole numbers, the term Diophantine became the name for such analysis. Diophantus made no attempt at a systematic theory, and almost nothing is known about Diophantine analysis by earlier mathematicians.

Today Diophantine analysis is a vast, complex branch of number theory (not algebra) with an enormous literature. There is a complete theory only for linear equations. No general method is known (it may not even exist) for solving equations with powers of 2 or higher. Even the simplest nonlinear Diophantine equation may be fantastically difficult to analyze. It may have no solution, an infinity of solutions or any finite number. Scores of such equations, so simple a child can understand them, have resisted all attempts either to find a solution or to prove none is possible.

The simplest nontrivial Diophantine equation has the linear form $ax + by = c$, where x and y are two unknowns and a, b and c are given integers. Let us see how such an equation underlies the puzzle in the opening paragraph. Letting x be the number of cows, y the number of pigs and z the number of sheep, we can write two equations:

$$10x + 3y + z/2 = 100$$

$$x + y + z = 100$$

To get rid of the fraction, multiply the first equation by 2. From this result, $20x + 6y + z = 200$, subtract the second equation. This eliminates z, leaving $19x + 5y = 100$. How do we find integral values for x and y? There are many ways, but I shall give only an old algorithm that utilizes continued fractions and that applies to all equations of this form.

Put the term with the smallest coefficient on the left: $5y = 100 - 19x$. Dividing both sides by 5 gives $y = (100 - 19x)/5$. We next divide 100 and $19x$ by 5, putting the remainders (if any) over 5 to form a terminal fraction. In this way the equation is transformed to $y = 20 - 3x - 4x/5$.

It is obvious that if x and y are positive integers (as they must be), x must have a positive value that will make $4x/5$ an integer. Clearly x must be a multiple of 5. The lowest such multiple is 5 itself. This gives y a value of 1 and z (going back to either of the two original equations) a value of 94. We have found a solution: 5 cows, 1 pig, 94 sheep. Are there other solutions? If negative integers are allowed, there are an infinite number, but here we cannot allow negative animals. When x is given a value of 10, or any higher multiple of 5, y becomes negative. The problem therefore has only one solution.

In this easy example the first integral fraction obtained, $4x/5$, does not contain a y term. For equations of the same form but with larger coefficients, the procedure just described must often be repeated many times. The terminal fraction is made equal to a new unknown integer, say a, the term with the lowest coefficient is put on the left, and the procedure is repeated to obtain a new terminal fraction. Eventually you are sure to end with a fraction that has only one unknown and is simple enough so that you can see what values the unknown must have to make the fraction integral. By working backward through whatever series of equations has been necessary, the original problem is solved. (An outline of this standard algorithm can be found in Helen Abbott Merrill's *Mathematical Excursions,* a Dover paperback, Chapter 12.)

For an example of an equation similar to the one just analyzed that has *no* solution, assume that cows cost $5, pigs $2 and sheep 50 cents. The two equations are handled exactly as before. The first is doubled to eliminate the fraction and the second is subtracted, producing the Diophantine equation $9x + 3y = 100$. Using the procedure of continued fractions, you end with $y = 33 - 3x -$

1/3, which shows that if x is integral, y cannot be. In this case, however, we can tell at once that $9x + 3y = 100$ has no solution by applying the following old theorem. If the coefficients of x and y have a common factor that is not a factor of the number on the right, the equation is unsolvable in integers. In this case 9 and 3 have 3 as a common divisor, but 3 is not a factor of 100. It is easy to see why the theorem holds. If the two terms on the left are each a multiple of n, so will their sum be; therefore the term on the right also must be a multiple of n. An even simpler instance would be $4x + 8y = 101$. The left side of the equality obviously must be an even integer, so that it cannot equal the odd number on the right. It is also good to remember that if all three given numbers do have a common factor, the equation can immediately be reduced by dividing all terms by the common divisor.

As an example of a variant of the basic problem that has a finite number (more than one) of positive-integer solutions, consider the case in which cows cost \$4, pigs \$2 and sheep a third of a dollar. The solutions will be given next month.

Many geometric problems are equivalent to finding integral solutions for Diophantine equations. (An example was given last month: finding integral values for the crossed-ladders problem.) A well-known unsolved problem of this type is finding a brick-shaped solid (called a rectangular parallelepiped) whose three edges, three face diagonals and single space diagonal (from one corner to the corner on the opposite side of the solid's center) are all integers [*see illustration below*]. The problem is equivalent to finding integer solutions for the seven unknowns in the following set of equations:

$$a^2 + b^2 = c^2$$
$$a^2 + d^2 = e^2$$
$$b^2 + d^2 = f^2$$
$$b^2 + e^2 = g^2$$

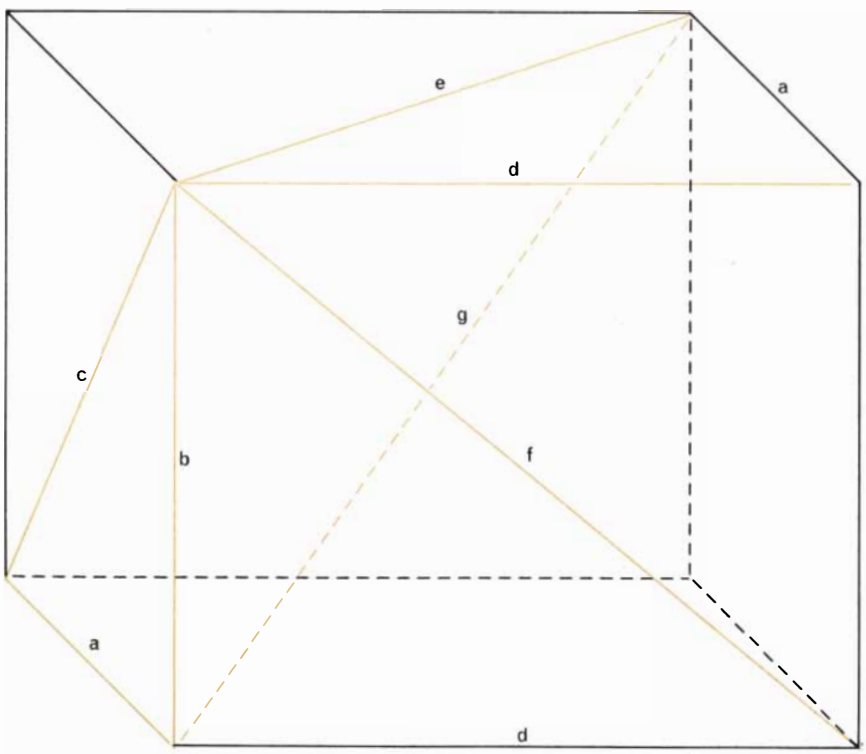
The problem has not been shown to be impossible, nor has it been solved. John Leech, a British mathematician, has been searching for a solution, and I am indebted to him for the following information. The smallest brick with integral edges and face diagonals (only the space diagonal is nonintegral) has edges of 44, 117 and 240. This was known by Leonhard Euler to be the minimum solution. If all values are integral except a face diagonal, the smallest brick has edges of 104, 153 and 672, a result also known to Euler. (The brick's space diagonal is 697.) The third case, where only an edge is nonintegral, has not, as far as Leech knows, been considered before. It too has solutions, but the numbers are, as Leech puts it, "hideous." He suspects that the smallest such brick may be one with edges of 3, 4 and the square root of 136,990,339,200. Of course the brick's volume is also irrational.

A much easier geometric problem, which I took from a puzzle book by L. H. Longley-Cook, is illustrated at the top of the opposite page. A rectangle (the term includes the square) is drawn on graph paper as shown and its border cells are shaded. In this case the shaded cells do not equal the unshaded cells of the interior rectangle. Is it possible to draw a rectangle of proportions such that the border—one cell wide—contains the same number of cells as there are within the border? If so, the task is to find *all* such solutions. The Diophantine equation that is involved can be solved easily by a factoring trick, which I shall explain next month.

In ancient times the most famous Diophantine problem, posed by Archimedes, became known as the "cattle problem." It involves eight unknowns, but the integral solutions are so huge (the smallest value contains more than 200,000 digits) that it was not fully calculated until five years ago, when a computer managed to do it. The interested reader will find a good discussion of the cattle problem in Eric Temple Bell's *The Last Problem* (1961, pages 151–157) and the final solution, by H. C. Williams and others, in the journal *Mathematics of Computation* (October, 1965).

The greatest of all Diophantine problems, which is still far from solved, is the "last theorem" of Pierre de Fermat, the 17th-century French amateur number theorist. (He was a jurist by profession.) Every mathematician knows how Fermat, reading a French translation of Diophantus' *Arithmetica*, added a note in Latin to the eighth problem of the second book, where an integral solution is asked for $x^2 + y^2 = a^2$. Fermat wrote that such an equation had no solution in integers for powers greater than 2. (When the power is 2, the solution is called a "Pythagorean triple" and there are endless numbers of solutions.) In brief, Fermat asserted that $x^n + y^n = a^n$ has no solution in integers if n is a positive integer greater than 2. "I have discovered a truly marvelous demonstration," Fermat concluded his note, "which this margin is too narrow to contain."

To this day no one knows if Fermat really had such a proof. Because the greatest mathematicians since Fermat have failed to find a proof, the consensus is that Fermat was mistaken. Linger-



The integral brick, an unsolved Diophantine problem

ing doubts arise from the fact that Fermat always *did* have a proof whenever he said he did. For example, consider the Diophantine equation $y^3 = x^2 + 2$. It is easy to find by trial and error that it has the solutions $3^3 = 5^2 + 2$ and $3^3 = -5^2 + 2$. To prove, however, that there are no other integral solutions, Bell writes in *Men of Mathematics*, "requires more innate intellectual capacity . . . than it does to grasp the theory of relativity." Fermat said he had such a proof although he did not publish it. "This time he was not guessing," Bell continues. "The problem is hard; he asserted that he had a proof; a proof was later found." Fermat did publish a relatively elementary proof that $x^4 + y^4 = a^4$ has no solution, and later mathematicians proved the impossibility of the more difficult $x^3 + y^3 = a^3$. The cases of $n = 5$ and $n = 7$ were settled early in the 19th century.

It is not hard to show that Fermat's last theorem is true for all exponents n except primes greater than 2. There is a special class of primes for which, in 1941, the theorem was shown to be true if n is no greater than 253,747,889. For all other prime values of n it was proved in 1964 that n must be at least 25,013. (See J. L. Selfridge and B. W. Pollack, *Notices of the American Mathematical Society,* Vol. 11, 1964, page 97.)

Attempts to prove Fermat's last theorem are so frustrating, in view of how easy it looks, that the theorem provides a means by which a mathematician outwits the devil in a fantasy story from which I took the quotation at the beginning of this article. (The story is reprinted in Clifton Fadiman's entertaining anthology, *Fantasia Mathematica.*) It is the deepest unsolved problem in Diophantine theory. Some mathematicians believe it may be true but unprovable, now that Kurt Gödel has shown, in his famous undecidability proof, that arithmetic contains theorems that cannot be established inside the deductive system of arithmetic. (If Fermat's last theorem is Gödelian-undecidable, it would have to be true, because if it were false, it would be decidable by a single counterexample.) I earnestly ask readers not to send me proofs. I am not competent to evaluate them. Ferdinand Lindemann, the first to prove (in 1882) that pi is transcendent, once published a long proof of Fermat's last theorem that turned out, according to Bell, to have its fatal mistake right at the beginning. Dozens of other fallacious proofs have been published by leading mathematicians. When David Hilbert was asked why he never tackled the problem, his reply was: "Before beginning I should put in three years of intensive study, and I haven't that much time to squander on a probable failure."

The mathematics departments of many large universities return all proofs of Fermat's last theorem with a form letter stating that the paper will be evaluated only after an advance payment of a specified fee. Edmund Landau, the German mathematician, used a form letter that read: "Dear Sir/Madam: Your proof of Fermat's last theorem has been received. The first mistake is on page ___, line ___." Landau would then assign the filling in of the blanks to a graduate student.

Donald E. Knuth whimsically asks for a proof of Fermat's last theorem as the last exercise at the end of his preface to the first volume of his series *The Art of Computer Programming* (1968). His answer states that someone who read a preliminary draft of the book reported that he had a truly remarkable proof but that the margin of the page was too small to contain it.

Euler failed to prove Fermat's last theorem, but he made a more general conjecture that, if it is true, would include the truth of Fermat's last theorem as a special case. Euler suggested that no nth power greater than 2 can be the integral sum of fewer than n nth powers. As we have seen, it has long been known that the conjecture holds when n is 3, for this is merely Fermat's last theorem with powers of 3. It is not yet known whether or not $x^4 + y^4 + z^4 = a^4$ has a solution.

I close on a happier note. In 1966 Leon J. Lander and Thomas R. Parkin, using a computer, found the first counterexample to Euler's conjecture. It was first published in *Bulletin of the American Mathematical Society* (Vol. 72, 1966, page 1079) about two centuries after Euler made his guess, in an article of just five lines. The counterexample, the smallest when n is 5, is

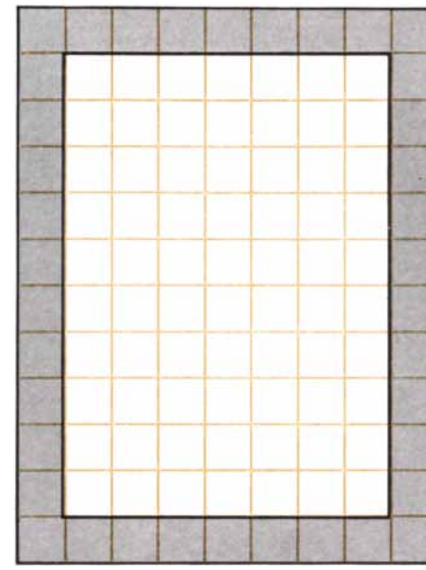
$$27^5 + 84^5 + 110^5 + 133^5 = 144^5 .$$

Last month's problem was to find the side of an equilateral triangle containing a point p that is three, four and five units from the triangle's corners. The following solution is from Charles W. Trigg, *Mathematical Quickies* (McGraw-Hill, 1967), answer to Problem 201. He credits the answer to a 1933 source. The broken lines of the illustration at the right are constructed so that PCF is an equilateral triangle and AE is perpendicular to PC extended left to E. Angle $PCB = 60$ degrees minus angle PCA = angle ACF. Triangles PCB and FCA are therefore congruent and $AF = BP = 5$. Because APF is a right triangle, angle $APE = 180 - 60 - 90 = 30$ degrees. From this we conclude that AE is 2 and EP is twice the square root of 3. This permits the equation

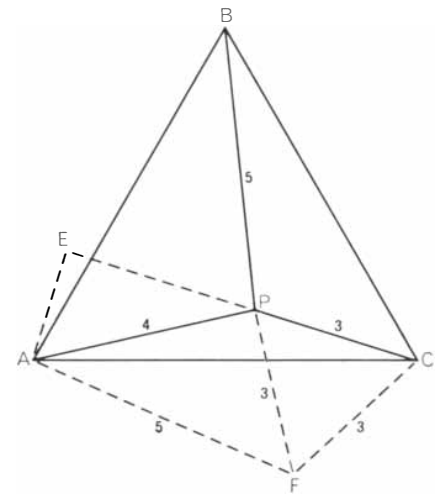
$$AC = \sqrt{2^2 + (3 + 2\sqrt{3})^2}$$

$$= \sqrt{25 + 12\sqrt{3}}\,,$$

which gives AC, a side of the original triangle, a value of 6.766+. For two ways to solve the general problem see L. A. Graham, *Ingenious Mathematical Problems and Methods* (Dover, 1959), answer to Problem 55.



A simple Diophantine problem



Solution to triangle problem

THE AMATEUR SCIENTIST

How to construct a molecular-beam apparatus and a mass spectrometer

Conducted by C. L. Stong

Many amateurs have built a cloud chamber and used it to detect the subatomic particles emitted by the isotopes in a speck of radium paint. Few, however, have undertaken other atomic experiments, primarily because of the hazards associated with radioactive materials. Radioactive substances are indeed hazardous. On the other hand, some of the most fascinating experiments in the field of atomic physics can be conducted with harmless substances. For example, molecules of common table salt can be split apart and formed into high-velocity beams that can be manipulated in interesting ways by electric and magnetic fields. Atoms in the beam can be separated into isotopes. The isotopes can be graded according to their mass and identified.

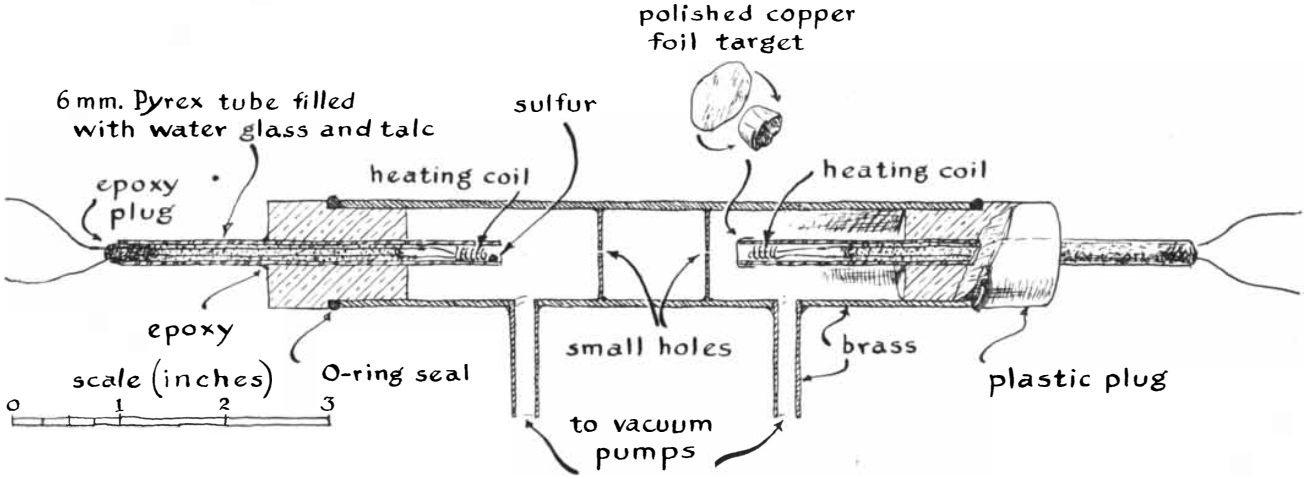
The apparatus required for doing such experiments is reasonably simple and can be built at home, much of it from cast-off odds and ends. For example, Roger Flood of Natick, Mass., recently devised an apparatus for generating a slender beam of sulfur molecules and detecting it chemically. His device was based on an apparatus developed in 1911 by the French physicist Louis Dunoyer for generating a beam of sodium molecules.

From a piece of glass tubing about eight inches long Dunoyer made a vessel in the form of three interconnecting bulbs. The vessel resembled an hourglass with three bulbs instead of two. A bulb at one end contained a small lump of sodium. When Dunoyer pumped the air from the vessel and heated the sodium, a dark spot appeared on the wall of the bulb at the opposite end. The spot was in perfect alignment with the small openings that interconnected the three bulbs. The position and size of the spot could be explained only on the assumption that molecules of evaporated sodium had traveled in straight lines through the openings to form a beam of particles that condensed on the glass.

Flood's version of the generator was made largely of metal. "My apparatus," he writes, "is identical in principle with Dunoyer's. It consists of a brass tube about half an inch in diameter and four inches long that is divided into three compartments by two brass disks that serve as partitions. A pinhole in the center of each disk interconnects the three compartments. The disks fit snugly in the tube and are soldered in place half an inch apart near the middle of the tube. The ends of the tube are closed by plastic plugs removed from a pair of UG-680/U coaxial-cable connectors. O rings inserted between the shoulders of the plugs and the ends of the pipe make the assembly vacuum-tight.

"The compartment at one end serves as an oven for evaporating sublimed sulfur. The perforated disks and the space between them function as a collimator: only the molecules that follow a straight path through the perforations are transmitted. The beam is detected in the third compartment, where molecules of sulfur react with a plate of hot copper to form a black deposit of cupric sulfide. This method of detecting a molecular beam is approximately a million times more sensitive than mechanical techniques of depositing a spot of sulfur thick enough to be visible to the unaided eye.

"The sulfur is evaporated by an electrically heated coil inside a tube of Pyrex glass. The tube is supported by one of the coaxial connectors [*see illustration below*]. The coil is made from 30-gauge Nichrome wire, whose ends are crimped to copper leads. The outer end of the



Roger Flood's molecular-beam apparatus

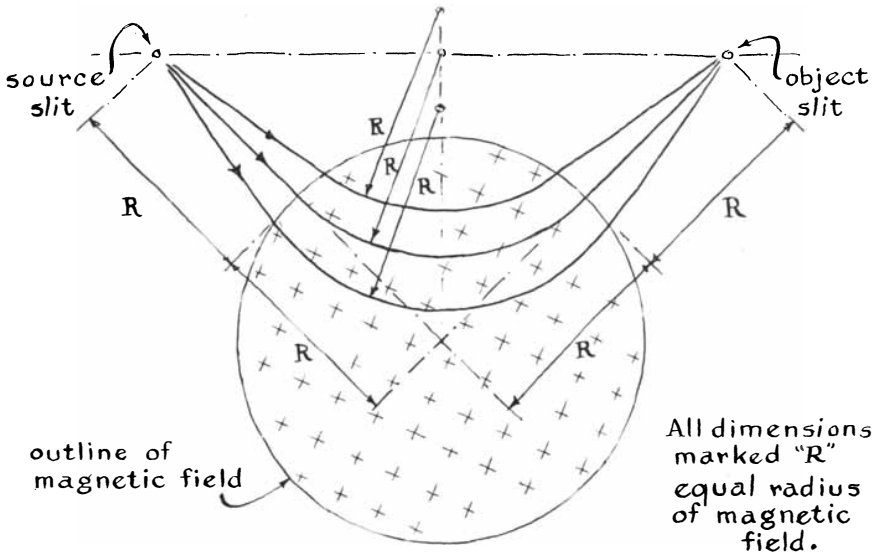
glass tube is sealed with epoxy cement and the interior is partly filled with a paste of water glass and talc that serves as insulation. The inner end of the tube is open. It is charged with a few milligrams of sulfur just before the apparatus is assembled for operation.

"A similar tube heats the copper target, which is a small piece of polished copper foil crimped in the form of a cup over the end of the glass. The proper operating temperatures are determined by experiment. Sufficient voltage should be applied to the vaporizing coil to melt sulfur in air within one minute. Similarly, sufficient voltage should be applied to the target coil to produce a thin coating of copper oxide in air within three minutes.

"The apparatus must be exhausted to a pressure of at least 10^{-4} torr. This pressure is just above the capability of mechanical air pumps. For the mechanical pump I used the compressor of a discarded home refrigerator, which I modified according to instructions in *The Scientific American Book of Projects for the Amateur Scientist* (Simon and Schuster, 1960). In series with the compressor I connected an inexpensive oil diffusion pump that is available from Morris & Lee Company, 1685 Elmwood Avenue, Buffalo, N.Y. 14207. The combination exhausts the apparatus to about 10^{-5} torr in 10 minutes.

"I apply a thin coating of vacuum grease to the O-ring seals. After the system has been exhausted to the limit of the pumps I heat the target for about five minutes and turn on the oven heater to vaporize the sulfur. A perceptible spot should form on the target within 15 minutes. I then turn off the oven heater but keep heating the target for another five minutes to ensure that the decaying beam continues to react with the copper. Air is admitted after the apparatus has cooled to room temperature, a procedure that prevents the target from oxidizing. The target is now examined. If all has gone well, a black spot with sharp edges is seen on the target. A fuzzy spot or random sputtering usually indicates that air is leaking into the system or that the pumps are not working properly.

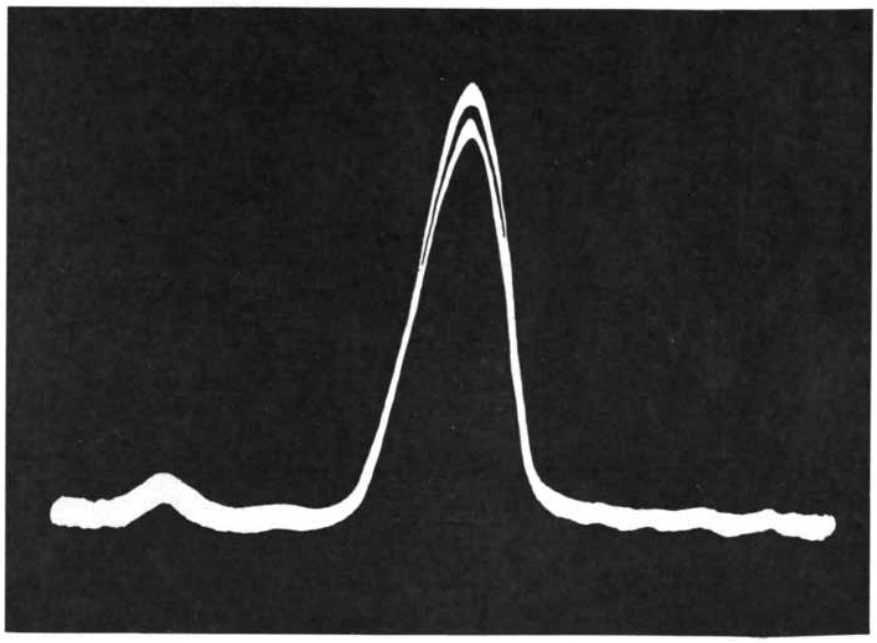
"Molecular beams of many chemical elements can be detected by their reaction with appropriate target materials, but I have experimented with only two other combinations. A beam of hydrogen reduces yellow molybdenum trioxide to blue molybdenum dioxide. Oxygen reacts with lead monoxide, which is pale yellow, to form lead dioxide, which is brown. I admit the gases to the heater tube through a small leak and operate the pumps continuously to maintain the desired low pressure."



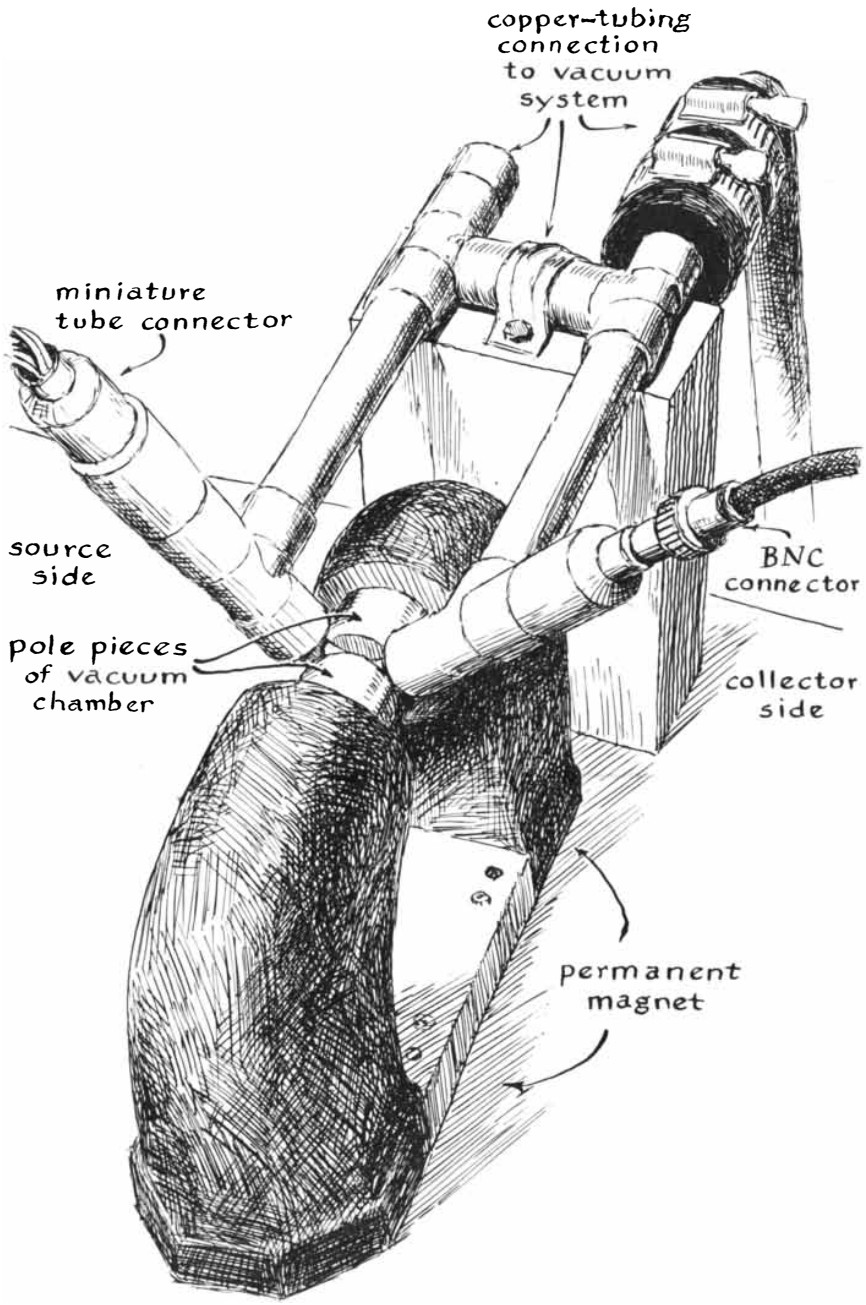
Focusing system of the mass spectrometer

By replacing the collimating apertures of Flood's apparatus with a pair of slits and passing the beam first through an electric field and then through a magnetic field, atoms constituting the beam can be separated into isotopes. The resulting apparatus is appropriately called a mass spectrometer because it sorts and separates isotopes according to their mass. In the December 1963 issue of the *American Journal of Physics* the late John W. Dewdney of Harvard University described the construction of a mass spectrometer that he assembled from copper tubes, darning needles, razor blades, a miniature incandescent lamp and other inexpensive parts. In 1963 the complete instrument cost less than $5, not counting the magnet and vacuum pumps.

The substance to be analyzed is vaporized by the incandescent filament of a small lamp bulb. Adjacent to the filament is a pair of metal plates and a pair of facing razor blades spaced to form a narrow slit. A potential difference of about 100 volts is applied to the plates



Oscilloscope trace depicting potassium isotopes



General view of the mass spectrometer

and the filament. The resulting electric field accelerates the ions and focuses them on the slit between the blades.

Ions pass through the slit as a diverging beam and enter a magnetic field placed at a right angle across the beam. The field creates a deflecting force that always acts at right angles to the direction in which the ions move. The force continues to act at right angles even though the path of the ions curves continuously. As a result the ions move in a circular arc. Overlapping arcs of equal radius intersect at two points. Hence the diverging beam of ions, which escape through a slit and enter a uniform magnetic field, converges at a distant point. In effect the magnetic field acts as a lens that focuses ions from a source slit onto an object slit [*see top illustration on preceding page*].

In Dewdney's mass spectrometer the magnet bends the path of the ions through an arc of 90 degrees and brings them to focus on an object slit. The object slit is also made of razor blades. Ions that pass through the object slit strike a collector electrode, to which they transfer charge. A continuous stream of ions that reaches the collector electrode induces a continuous current in the output circuit. The intensity of the current is a measure of the number of ions that impinge on the collector.

The magnetic field does not deflect all ions equally. Ions of low mass curve more sharply than those of higher mass. (In much the same way, although for different reasons, a simple lens bends short waves of light more sharply than long waves.) Moreover, ions that move at low velocity are deflected more than those that move at high velocity. The velocity of ions emerging from the source slit is determined in part by the potential difference between the filament and the focusing plates. By changing this potential one can alter the influence of the magnetic field on the path of the ions. In effect the focal length of the "magnetic lens" can be adjusted by adjusting the accelerating voltage.

Dewdney altered the accelerating voltage continuously by connecting an alternating-current transformer in series with the source of accelerating voltage. This arrangement automatically focuses ions of differing mass sequentially on the object slit and causes corresponding currents to appear sequentially in the output current. The current can be applied, after appropriate amplification, to the vertical plates of an oscilloscope. The same continuously varying potential can be applied to the horizontal plates of the oscilloscope. The combination causes the oscilloscope to display a trace in the form of a jagged line, each peak of which corresponds to a species of isotope in the beam. With this technique Dewdney made a graph of the isotopes of potassium [*see bottom illustration on preceding page*]. The graph displays two peaks. The shorter one shows the relative abundance in the beam of potassium of mass 41, and the larger peak displays the more abundant isotope of mass 39.

The vacuum chamber of the spectrometer and the ports through which air is exhausted are assembled from copper plumbing parts. The chamber consists of a pair of 1/2-inch copper *T*'s; the crossarms are coupled by a 90-degree elbow of 1/4-inch diameter through reducers that were sawed off flush with the end of the *T*'s. The legs of the *T*'s are coupled to the vacuum pumps by 1/2-inch copper tubing [*see illustration on this page*]. All joints are soldered.

The functional parts of the instrument, which include the filament, focusing plates, slits and collector electrode, are supported by No. 3 steel darning needles pushed through No. 1 rubber stoppers that fit the open ends of the *T*'s. Dewdney warned that the installation of the darning needles is the most

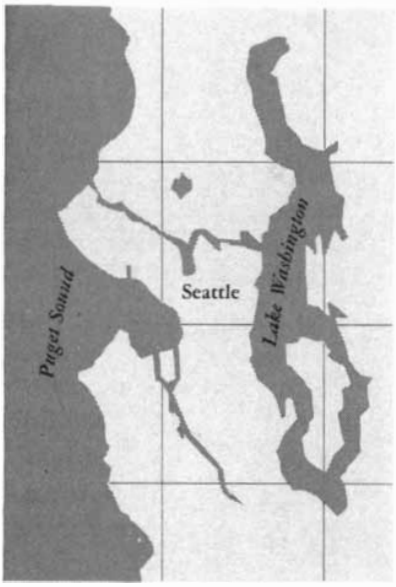
Blueprint for success: how a major metropolitan area solved water pollution.

The 1970's seem destined to be known as Ecology's Decade. But an important chapter was written over a decade ago in Washington State, when a remarkable combination of self-starting citizen action and legislative leadership joined together to save the unique marine environment of the Seattle metropolitan region.

Lake Washington is the largest lake in the United States which is completely surrounded by a metropolitan area. By the mid 50's, the pollution of the lake had reached a point of crisis. For one thing, a fast-growing population in and around Seattle had discovered that the lake and Puget Sound were convenient places for dumping effluents. At the same time, the widespread use of detergents had intensified the problem of fertilization.

As the phosphorous-rich detergents became part of the effluent, they had the effect of fertilizing the lake for burgeoning colonies of algae. Unlike bacteriological pollution, which was being dealt with by sewage treatment plants, lake fertilization brought its own problems—foul odors, loss of the lake's clarity, and in late summer, an incredible graveyard of algae gardens. In terms of recreational use, Lake Washington was dying. To the west, raw sewage was turning Puget Sound into a coffin for marine life.

With time running out, a deeply concerned, area-wide committee of volunteer citizens submitted state-enabling legislation for establishing a special metropolitan government. "Metro" was designed to cut across traditional political and geographical boundaries of urban and suburban



Seattle and was empowered to set up a comprehensive sewage system for the entire drainage basin.

Approved by the State legislature in 1957, and by a vote of the people in 1958, Metro began work on a ten-year, $130 million program which included circling the lake with large interceptors, abandoning all ten treatment plants on its shores and diverting all sewage from the drainage basin.

Following ground-breaking in 1961, the construction program proceeded without interruption. It utilized new tunneling procedures, such new concepts as computerized control and monitoring of the pumping stations, and automated, continuous water sampling to verify the design validity. Special engineering designs were precisely geared to the ecological conditions existing in the area, and featured a unique subsurface outfall that prevents treated sewage from rising and returning to the beach. The program was completed at less cost than was estimated and ahead of schedule.

Today, Lake Washington is the largest lake in the world to be completely rehabilitated from serious sewage damage. Its phosphorous content has fallen from a high of 70 parts per billion in 1963 to a level it had not enjoyed since 1933—only 22 parts phosphorous per billion. The minimum transparency of the lake has risen from a mere two and one half feet in 1965 to more than 10 feet in early 1970.

In recognition of the citizen effort that led to saving the area's waters, Metro received an All-America City award. But the rehabilitation of the lake and Puget Sound is only one of several major steps currently being taken to save Washington State's environment. The important point is that rhetoric and token anti-pollution programs have been rejected in place of a comprehensive plan and genuine action.

Washington State is a fertile field for exciting breakthroughs in many areas. In technology, aerospace, forestry, oceanography, bio-engineering and other fields, the smart money is betting on Washington State.

Daniel J. Evans
Governor

STATE OF WASHINGTON

For business location information, write: Daniel B. Ward, Director, Department of Commerce and Economic Development, Olympia, Washington 98501

Of the three fastest memories, we make two.

The hottest new thing in memories these days is plated wire. We suspect that a number of outfits are working on them. But we're already producing them.

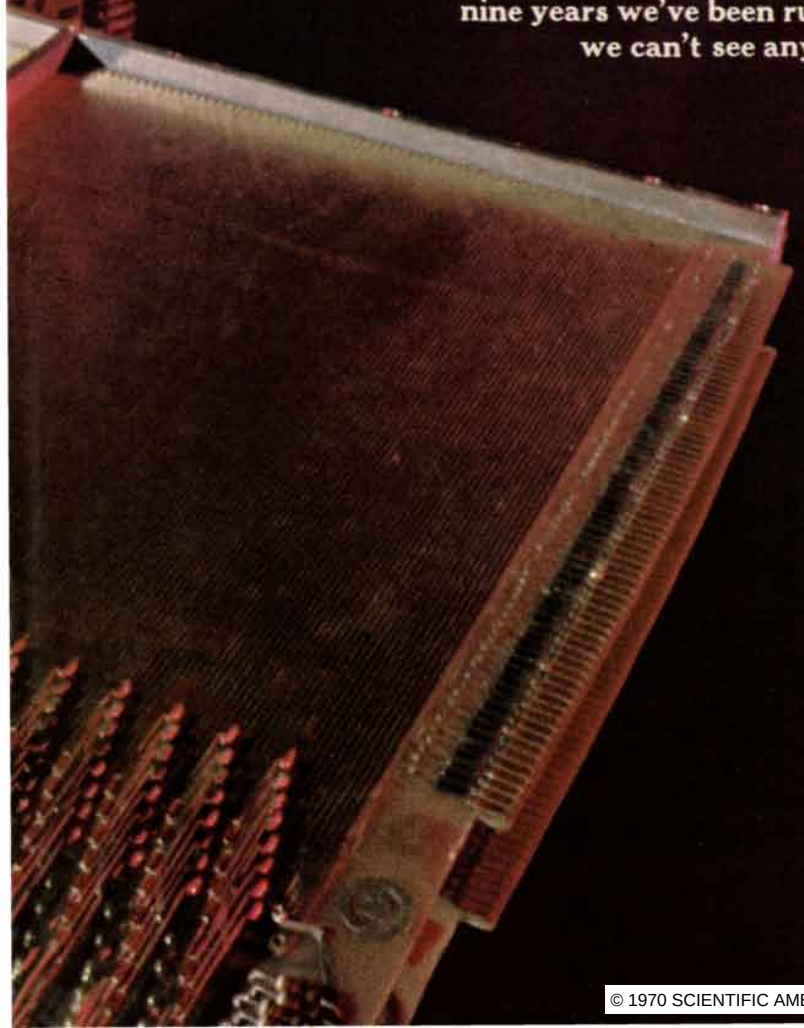
The one at the bottom is designed for use in nondestructive readout memory systems having read cycle times of 150 nanoseconds or less and write cycle times of 300 nanoseconds.

The other one is our new 14 mil core memory. With a 250 nanosecond full cycle time, it's the most agile core memory around. If you don't believe it, just ask.

While you're at it, ask about our printed circuits, too. You'll find out we've made more complex printed circuits than anyone else. Including flexible circuits. And multilayer boards with electrical characteristics so finely controllable they can be used as electronic components.

If you also happen to be in the market for control computers, we'll tell you all about MAC 16. We make all the major parts that go into it, so that alone makes us unique among mini computer manufacturers.

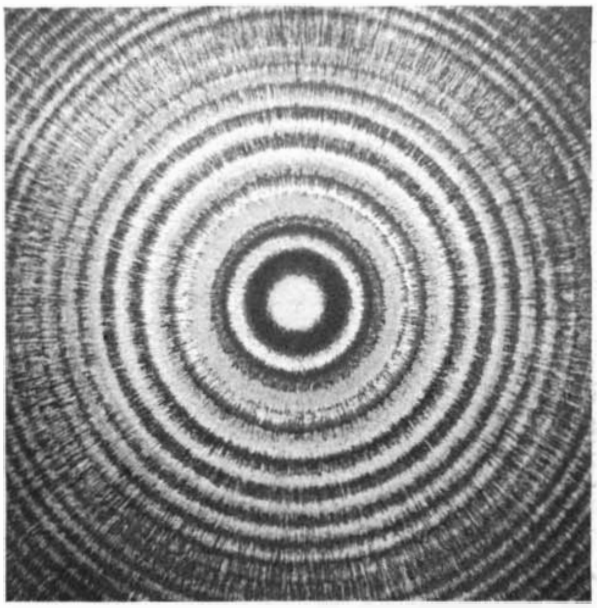
Competition in the data products business is tough. For nine years we've been running hard. All of a sudden, we can't see anyone in front of us.



Lockheed Electronics
Data Products Division
A Division of Lockheed Aircraft Corporation
6201 E. Randolph St., Los Angeles, California 90022

"LASER LIGHT"

A SCIENTIFIC AMERICAN FILM



Almost everyone has heard of lasers, but relatively few people have seen them in action. The Editors of SCIENTIFIC AMERICAN now present "LASER LIGHT," a 16-millimeter sound film about lasers: what they are, how they work, the marvelously pure and curiously scintillating light they produce, how they are being used and how they may be used in the near future. The film is in color and lasts 37½ minutes. It is now available for sale or rent.

A few highlights of the film are:

- **Computer-generated animation explaining stimulated emission and resonant optical cavities.**
- **Ripple-tank and oscilloscope demonstrations explaining the wave principles underlying laser action and holography.**
- **Holograms, their three-dimensionality dramatically evoked by the moving camera.**
- **A 600-foot, 8.8-kilowatt laser in action.**
- **Tunable lasers.**
- **A television picture transmitted by laser beam.**
- **The laser chalkline for the San Francisco Bay tunnel.**
- **Laser interferometry.**
- **Gas, solid and organic-liquid lasers.**
- **An experiment on the use of holography in a computer memory.**
- **Original musical score.**

"LASER LIGHT" is recommended for general audiences with an interest in science and technology, and for use in conjunction with the teaching of physics and optics. The film is accompanied by a selection of five SCIENTIFIC AMERICAN articles on lasers and holography, written by leading authorities in these fields.

The sale price per print is $375, the rental price $37.50 for a booking of three days. If the film is purchased after rental, the rental price will be deducted. If rental booking is desired, kindly specify date. Write Motion Picture Department, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017

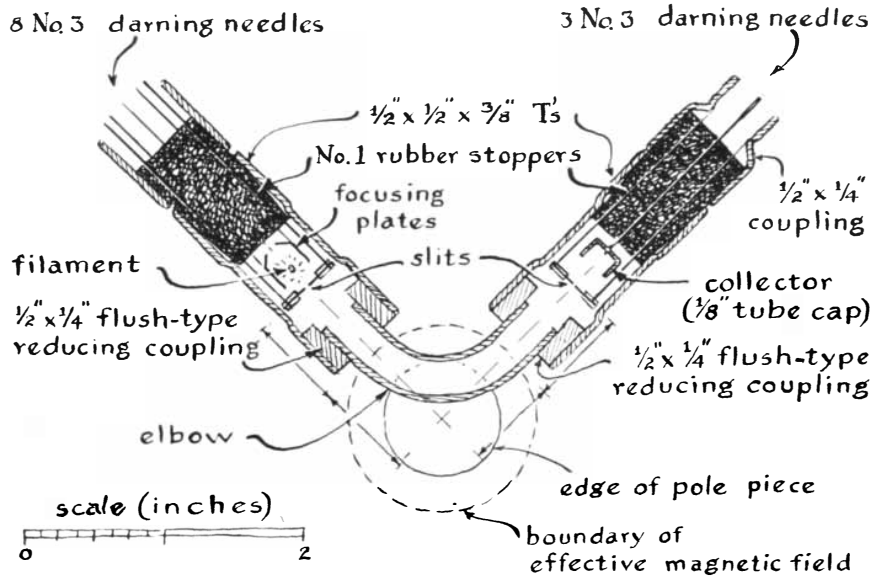
difficult part of making the instrument, and I agree! The needles must be positioned as accurately as possible and parallel to the axes of the stoppers. Dewdney made a jig for guiding the needles into the rubber. It is a thick cap of metal drilled with holes that make a sliding fit with the needles. The small end of the stopper is pushed into the cap. The needles are pushed through the holes and the rubber. One must hope they remain parallel.

I did the job by chucking the needles in a drill press, standing the stoppers on a block of wood that rested on the bed of the press and lowering the chuck until the points emerged from the rubber and penetrated the wood. Each needle was then pushed farther through the rubber by hand. Eight needles evenly spaced in a circle 3/8 inch in diameter are needed for the source assembly. Three needles evenly spaced across the diameter of the second stopper support the object slit and the collector electrode [*see top illustration at right*].

The filament used for evaporating specimen materials is removed from a low-voltage incandescent lamp, such as a No. 41 pilot lamp or a No. 63 automobile lamp. File a scratch about halfway around the glass at the upper edge of the base and, while wearing rubber gloves, grip the bulb with one hand, the base with the other and pull. Do not attempt to "bend" the glass. The filament will remain intact when the glass breaks if the lamp has not been used. With pliers crack the remainder of the bulb from the leads, but do not remove the colored glass bead that ties the leads together.

To mount the filament to the stopper assembly, thread the connecting leads through the eyes of the needles and pull the needles into the rubber until the tips of the eyes are flush with the surface. Clip off the excess wire and press the colored bead into contact with the rubber at the center of the stopper, being careful to preserve the shape of the filament. The focusing electrodes can be made of any thin sheet metal, such as brass shim stock. Mount the electrodes as shown in the accompanying illustration [*bottom of this page*].

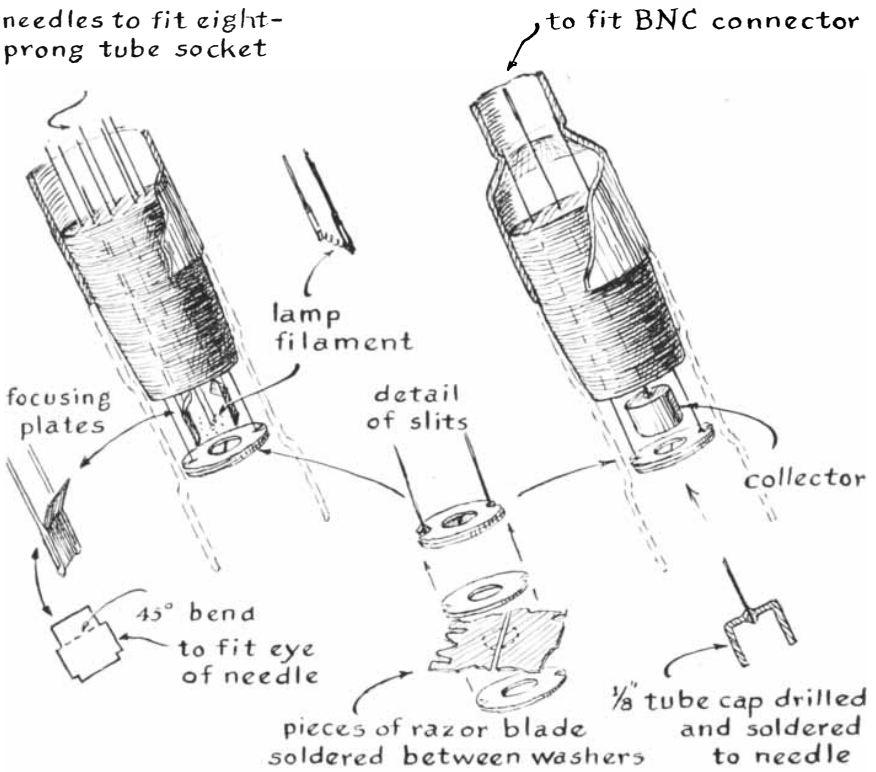
Make the slits out of a double-edged razor blade broken into four rectangular pieces, each with a cutting edge. Clean both sides of each piece, coat the sides with soldering flux and sandwich each pair of pieces, with the cutting edges facing, between No. 12 brass washers of 1/2-inch diameter. Clamp the sandwiches with alligator clips. Adjust the space between the cutting edges to form a slit .005 inch wide. Use another razor blade as a feeler gauge to make the adjustment. Solder the assembly while it is clamped. Break off portions of the soldered blades that extend beyond the edge of the washers.



Details of the mass spectrometer's vacuum chamber

Drill a hole that makes a snug fit with the needles in the center of a 1/8-inch copper pipe cap, turn the cap so that the open end faces away from the small end of the stopper and solder the tip of the middle needle into the hole. This completes the collector electrode. In the same way drill a pair of holes spaced 3/8 inch apart near the edges of one of the slits and solder it to the remaining pair of needles. This completes the objective slit.

The space between the collector electrode and the slit can be adjusted readily by sliding the needles through the rub-



Elements of the source and object assemblies

ber. The dimension is not critical but should be approximately 1/8 inch. The source assembly is similarly mounted. If the needles are numbered clockwise from 1 through 8, needles 1 and 5 support the filament; 3 and 7, the source slit; 2 and 4, one focusing electrode, and 6 and 8, the remaining focusing electrode.

The magnetic field is supplied by a permanent alnico magnet provided with pole pieces in the shape of right cylinders. The pole pieces can be cut from mild-steel shafting and should be faced squarely by a lathe. They should have a diameter of 7/8 inch. The air gap between the poles should be made as narrow as possible, just wide enough to admit the 90-degree elbow of the vacuum chamber. The radius of the circular magnetic field should equal the radius of the 90-degree elbow.

The magnetic field spreads outward as it emerges from the magnet, and its radius is accordingly longer than the radius of the pole pieces. Dewdney found experimentally that the diameter of the pole pieces is about 65 percent less than the diameter of the effective field. This information should be kept in mind if the dimensions of the spectrometer are modified.

Any alnico magnet can be used that develops a magnetic field in the air gap of 3,000 gauss or more. The magnet should be positioned as illustrated. Magnets that can be modified for this application are sold by the Edmund Scientific Co., Barrington, N.J. 08007.

Electrical connections to the source are made by pushing a miniature eight-hole vacuum-tube socket over the needles. If you wish, you can force the large end of the rubber stopper into a 1/2-inch copper nipple, so that 5/8 inch of the small end of the stopper protrudes. The nipple provides a base for supporting the socket mechanically. The stopper that supports the object slit and collector can be equipped with a 1/2-inch to 1/4-inch reducer to serve as a mechanical base for a Type BNC connector that couples the collector electrode to the output circuit.

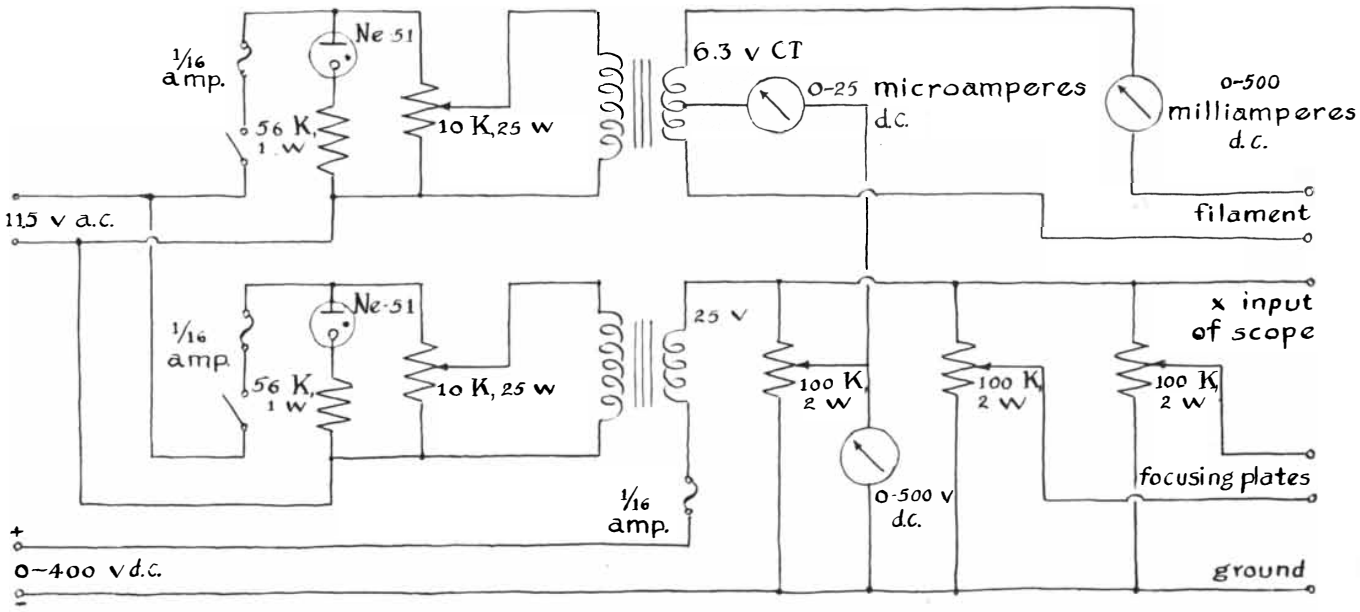
The electrical power supply provides a six-volt alternating current for heating the lamp filament and a pulsating direct voltage at a peak potential of up to 425 volts to the filament, the focusing plates and the horizontal plates (x axis) of the oscilloscope. The pulsations are developed by connecting the secondary winding of a 115-to-25-volt transformer in series with an external 400-volt direct-current power supply [*see illustration below*]. The potential applied to both the oscilloscope and the focusing plates is controlled by potentiometers that also determine the polarity of the focusing plates with respect to each other and to the filament.

The procedure for adjusting and operating the instrument is determined in part by the nature of the substances to be analyzed. Alkaline metals ionize readily and are convenient for introductory experiments. This group of metals includes lithium, sodium, potassium, rubidium and cesium. Dewdney used the chlorides of these elements, particularly potassium chloride.

For an initial experiment dissolve as much potassium chloride as possible in distilled water at room temperature, place a drop on the filament and let it dry. Apply a film of vacuum grease to both rubber stoppers and insert them in the T's. Rotate the stoppers so that the slits are approximately parallel to the lines of magnetic force in the air gap. Exhaust the vacuum chamber to the limit of the pumps, at least to 10^{-4} torr. Turn on the power supply and adjust the filament to a potential of approximately 150 volts and the focusing electrodes to about 50 volts. The filament is now at a *positive* potential of approximately 100 volts with respect to the focusing plates and is of the same polarity as the ions.

Substances such as the alkali metals, which have small ionization potentials with respect to the work function of tungsten, emit positive ions. Substances of ionization potential higher than the work function of tungsten emit negative ions. The polarity of the filament must correspond to the polarity of the ions. (The like charges repel the ions, thereby speeding them toward the slit.)

Doubtless the oscilloscope will display a straight or nearly straight horizontal trace. Rotate the object slit back and forth through a few angular degrees. Stop when one or more small peaks appear in the horizontal trace and then increase and decrease the potential of the focusing plates to maximize the peaks. Interesting salts for experimentation include lithium sulfate, sodium sulfate, barium chloride and cesium chloride.



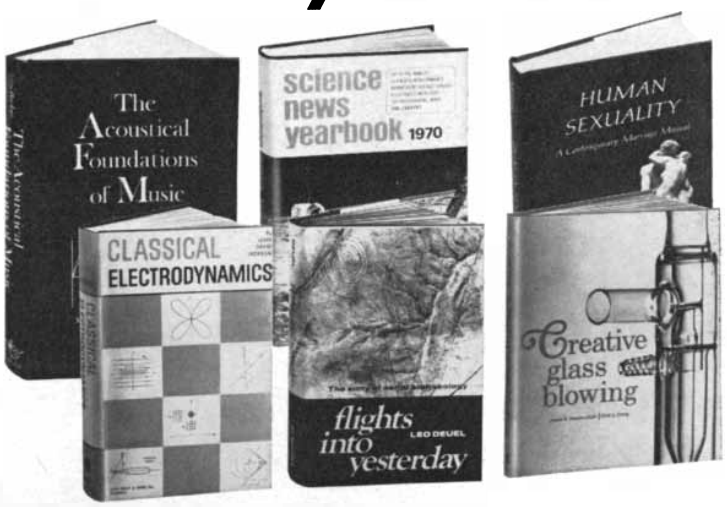
Power supply of the mass spectrometer

Contour strip planting of field crops in Monroe County, Wisconsin. USDA–SOIL CONSERVATION SERVICE

Citrus grove planted on contour lines at Santa Paula, California. USDA–SOIL CONSERVATION SERVICE

Only one book club helps you keep up with the whole spectrum of science and save on the books you want.

The Library of Science is the one. Our editors search out hundreds of manuscripts and pre-publication proofs to offer you only the books of highest interest and greatest significance. Great books. Fascinating books. Books you must know and own. And they're yours at impressive savings. Discounts on books range to 30%, often more. And every fourth book you buy earns you a free bonus book. Why not choose three of the books you want right now?



Take 3 of these important books

(values to $47.65)

all for only $3.95

with a short trial membership in

The Library of Science

(retail prices shown.)

32150. THE ACOUSTICAL FOUNDATIONS OF MUSIC. *John Backus.* A delightfully erudite book about musical sound: its properties, production, behavior, and reproduction. **$7.95**

32530. ALGEBRA. *Saunders MacLane and Garrett Birkhoff.* The "functional" approach to algebra — new concepts, new vocabulary. **$11.95**

34090. APPLIED NUMERICAL METHODS. *Brice Carnahan, H. A. Luther and James O. Wilkes.* Important new book on computer solutions to problems in science, engineering, applied mathematics. 624 pages. **$14.95**

35000. THE AUTOBIOGRAPHY OF BERTRAND RUSSELL, 1944-1969. The third and final volume of this distinguished mathematician and philosopher's own story — told in his usual witty, lucid style. **$8.95**

36280. A BIOGRAPHICAL DICTIONARY OF SCIENTISTS. *Edited by Trevor I. Williams.* 600 pages on more than 1,000 scientists in fields from agriculture to zoology. **$9.95**

37530. CHEMICAL EVOLUTION. *Melvin Calvin.* The Nobel laureate's major work on the biochemical evolution of life. **$9.00**

39770. COMPLETE RUSSIAN COURSE FOR SCIENTISTS / RUSSIAN-ENGLISH TRANSLATORS DICTIONARY. *A dual selection.* Beresford's intensive language course for beginners, ideal for self-tuition. *Plus* Zimmerman's unique dictionary, a perfect complement. **$17.95**

Trial Membership Application

2-910

The Library of Science
Riverside, N.J. 08075

Upon acceptance of this order, please enroll me as a trial member and send the 3 books whose numbers I have indicated below. Bill me just $3.95 for all three, plus shipping and handling. If not delighted, I will return them within ten days and my membership will be cancelled. As a trial member, I need accept only 3 more selections during the next 12 months, all at reduced member prices plus shipping and handling. I understand that savings range up to 30%, and occasionally more. Each month, I will receive advance reviews describing forthcoming selections, along with a convenient form for requesting alternate selections or no book at all. For every 4 selections purchased, I may choose a free bonus book. (Introductory offer counts as first selection.)

3 books—all for only $3.95 (write in numbers)

_____ _____ _____

Name_____

Address_____

City_____ State_____ Zip_____

Offer good in Continental U.S. and Canada only.

39850. COMPUTER SORTING. *Ivan Flores.* The definitive work on sorting methods and applications — from assembly design to list processing. **$12.95**

40860. CREATIVE GLASS BLOWING. *James E. Hammesfahr and Clair L. Stong.* For the scientist who needs laboratory equipment, as well as for the creative hobbyist. **$8.00**

44240. ELEMENTARY QUANTUM MECHANICS. *David S. Saxon.* From concepts to mathematical formulation of the twentieth century's most important theory. **$12.50**

48670. FLIGHTS INTO YESTERDAY. *Leo Deuel.* A brilliant scientific detective story about the archaeological discoveries resulting from the use of aerial photography. **$8.95**

52020. ARTIFACTS OF COLONIAL AMERICA. *Ivor Noël Hume.* For the serious archaeologist and the colonial period buff alike. A must for antique collectors. **$10.00**

54090. HUMAN SEXUALITY. A Contemporary Marriage Manual. *James Leslie McCrary.* The most comprehensive, enlightened work on sexual physiology and behavior now in print. **$9.75**

56350. INTRODUCTION TO GEOMETRY. Revised Edition. *H. S. M. Coexter.* The modern classic, revised, with a wealth of fresh material. Illustrated. **$10.95**

56430. INTRODUCTION TO STATISTICAL DATA PROCESSING. *Theodor D. Sterling and Seymour V. Pollack.* A thorough treatment of statistics and the intricacies of data collection and evaluation. 680 pages. **$12.50**

66060. OPTICAL PHYSICS. *Stephen G. Lipson and Henry Lipson.* Principles of optics, and their influence in other branches of physical theory. **$12.50**

66120. ORCHID FLOWERS: Their Pollination and Evolution. *L. van der Pijl and Calaway H. Dodson.* Ecology and adaptive evolution in action. Lavishly illustrated. **$12.50**

68730. PHYSICS OF THE EARTH. *Frank D. Stacey.* A sound, lucid work on the fundamentals of geophysics, incorporating today's new theories. **$11.95**

75720. SCIENCE NEWS YEARBOOK, 1970. *Science Service.* An invaluable reference on the latest developments in every area of science during the past year. **$9.95**

76900. SCIENTIFIC STUDY OF UNIDENTIFIED FLYING OBJECTS. *Edward U. Condon.* The complete text of the celebrated report commissioned by the U.S. Air Force. Illustrated; 967 pages. **$12.95**

78730. SO HUMAN AN ANIMAL. *René Dubos.* Must reading on the runaway technology that is destroying the environment. A Pulitzer Prize winner. **$6.95**

84130. THERMAL PHYSICS. *Charles Kittel.* A revolutionary treatment from a quantum viewpoint of this all-important field. **$10.95**

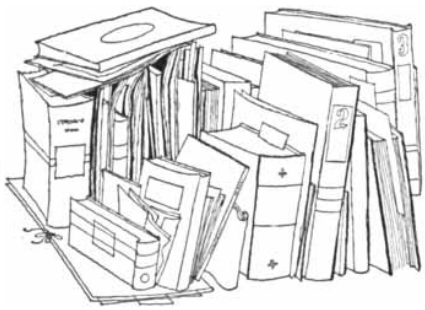
38310. CLASSICAL ELECTRODYNAMICS. *John David Jackson.* The pre-eminent work on electrodynamics, now in its fifth printing. **$13.95**

70350. PROBABILITY THEORY AND ITS APPLICATIONS, Volume I. Third Edition. *William Feller.* A lucid exposition of probability in terms of mathematical concepts, amply illustrated with problems and examples. **$10.95**

56140. INTRODUCTION TO OPERATIONS RESEARCH. *Frederick S. Hillier and Gerald J. Lieberman.* Basic methodology and techniques, emphasizing motivation and simplicity of explanation rather than technical details. **$14.75**

53060. A HISTORY OF MATHEMATICS. *Carl Boyer.* A monumental presentation of the growth and development of the mathematical sciences. Includes graded problems and exercises. **$10.95**

56330. AN INTRODUCTION TO SOLID STATE PHYSICS. Third Edition. *Charles Kittel.* The best book in its field, rewritten and brought up to date. **$12.95**



BOOKS

The realities of world agriculture: surveys by two concerned skeptics

by Philip Morrison

THE LONELY FURROW: FARMING IN THE UNITED STATES, JAPAN, AND INDIA, by Kusum Nair. The University of Michigan Press ($7.95). FOOD RESOURCES, CONVENTIONAL AND NOVEL, by N. W. Pirie. Penguin Books ($1.25).

Kusum Nair, an Indian woman journalist unusual both for her perspicacity and her sensitivity, has tested the generalizations of agricultural economists by comparative visits direct to the men and women who plant and plow in three diverse nations. Dr. Pirie, a biochemist and a literate and formidable debater, has written a brief, candid and extremely instructive summary of the yield and capabilities of farming as it is now practiced worldwide, with a critical account of potential new methods, following them from laboratory to fishing ground or field and all the way to the bowl of the hungry but wary buyer. His book complements Mrs. Nair's; it deals not with the farmer or his motives but with the statistics, the biology behind the data, the bureaucratic inertia that can belie them, and the fundamental physicochemical channels within which all growth must flow.

The two books subtly share a single mood; it is this that singles them out among the bumper crop of works on the food problem, both the economic and the biotechnical. They do not exhibit the detached, half-floating, assured commonplaces of many experts writing popular accounts. Nor do they display the polemic and the passion of those writers whose determination to prove a point outweighs their evidence. Both authors write in the vein of the concerned skeptic; they show that the experts are often demonstrably wrong. And both authors support their views with a large and useful supply of fresh fact and sharpened inference. The human insights one gains into the lives of Mrs. Nair's rice farmers, and the aptness of Dr. Pirie's use of the King James Version and the standard English authors, by themselves make these useful expositions also small works of art.

The stance Pirie takes is familiar and well argued. The ordinary but manifold means of agriculture contain the potential to double the world's fat and carbohydrate easily. Protein alone seems short. The evidence on which we base the protein requirement is scanty. The excretion of nitrogen is plainly a very indirect test of what is needed; there is no clear reason to believe that the rate of nitrogen loss on a protein-free diet is just what the doctor ought to order. In 1939 those who warned of protein shortage in beleaguered England were "stigmatized as obsessives." In the late 1950's experts of the United Nations Food and Agriculture Organization set the requirement at .5 gram of protein per day per kilogram of body weight. Pirie writes: "The estimates can be accurately called ungenerous." It is difficult to believe the members of the committees themselves subsisted on the diets they recommended for others. In 1965 the experts went up to .7 gram per kilogram. Perhaps we need the figure of one gram, which was the guess between the wars. The difference represents a world protein deficit of between 20 and 40 million tons per year. Now or by the end of the century we shall fall short by about a quarter of the expected total supply. Protein extraction from the residue of oilseed pressing and from the too bulky and too fibrous leaves of green plants and from coconuts are practical proposals. University of Oxford Famine Relief (Oxfam) gets the profits from Pirie's book "but [Oxfam] should be accompanied by political action. The most useful people now are those with ideas about a better plant, an unused animal, or a by-product that could be turned into food."

Mrs. Nair sat drinking coffee or tea or walked in the furrows of farmers at work across the world. That furrow is lonely, she says, because in the end the farmer alone makes the thousand decisions that determine yield. Of all the variables he is the most crucial and the least predictable. Her interviews make a compelling case. Of four Wisconsin dairymen only one struggles to innovate, improve, add. His "cows live better probably than the families" of the two poorest of the four. They are all fulltime commercial dirt farmers, starting out on their own farms, similar in size, in the early 1940's. They have the same schooling; they have the same access to advice; they sell the same standardized fluid milk at the same price. Yet one man is indifferent, another self-deluded, a third efficient but set against all change and growth. Why? Unevenness is the farmer's law. Even subtracting the factory farms of California, fewer than 5 percent of the most productive farmers produce more than the lowest 75 percent do.

In California's Yuba County "rice is an easy crop." It is heavily mechanized farming without transplanting; the flooded land is seeded, sprayed and fertilized from the air. Japanese rice farmers, however, get a larger yield, at about the same price, on poorer land in a worse climate. Japanese farmers transplant. They do so even though labor is getting scarcer. "I do not mind," says a Japanese farmer named Hatta. "Money is not of prime importance to me. I enjoy the work." When nearly a century ago the Japanese government established new farmers (they were surplus low-rank samurai) in the cold northern lands of Hokkaido, every foreign expert quite sensibly discouraged rice growing. It was even made illegal to plant rice. The new farmers persisted nonetheless. They transplanted twice, the first time to a cold frame. By that means and by the use of new varieties (not from the experiment stations but found by diligent farmers who even formed a society for exchanging seeds) they achieved a yield per acre no smaller than the yield in the southern parts of the country. They now use motors and small power tillers, but fertilizers, new varieties and the whole style of rice culture have not essentially changed from the skillful practice of Tokugawa days. Theirs is the purest Protestant ethic, although it knows neither

Calvin nor Knox; it is innocent of the doctrine of the elect but it too was born somehow out of feudalism.

The Madras rice farmers are not the same. "How much more do you think you could get if you did everything right?" The owner-cultivators who work one or two acres each reckon as much as 50 percent more. "Then why don't you?" "It's funny. But every year we decide we are going to do it. We *want* to do it. And then we don't." "But why?" "Oh! We say we'll do it tomorrow.... Or we have to go to court, or attend a marriage.... And so, for one reason or another, the various operations are not done as and when they should be."

The furrow is lonely, but it is patently also social. The uneven incentives will not be eased by the green revolution. The history of a land lies deeper than the rice roots.

HILBERT, by Constance Reid. Springer-Verlag ($8.80). *Extra Göttingen non est vita.* Outside Göttingen there is no life. So reads the arrogant motto on the wall of the rathskeller in that little hill town, where the yellow-brick university lies just outside the old wall. It is Gauss's university, as for scientists Cambridge must always be Newton's. In 1895, a century after Gauss had matriculated, David Hilbert, scion of an old Puritan family and son of a Königsberg judge, became a Göttingen professor. Until he retired in 1930 he stood as a worthy successor to Gauss, first among a dozen great mathematicians of Göttingen. One of Hilbert's powerful innovations—but only one—was his spectral theory. It was able to unite the treatment of simple mechanical vibrations with the most complicated problems of wave motions in a continuous medium. Exactly that mathematical theory turned out to lie beneath the outburst of ideas that in 1925 the Göttingen physicists—Werner Heisenberg, Max Born and Pascual Jordan (and elsewhere Erwin Schrödinger)—made into quantum mechanics, the deepest theory of matter men yet have. In 1924 Richard Courant had published his famous volume on mathematical physics (known as the "yellow peril" for its cover and its compactness), to which he also affixed Hilbert's name. Almost magically Courant-Hilbert prophesied the mathematics of quantum mechanics for three decades. All these men were students or grand-students of Hilbert's. So were Hermann Weyl and John von Neumann. Not Einstein. Einstein had worked in Zurich under Hermann Minkowski, whose death at 44 was a heavy blow to Hilbert, who had brought him to Göttingen. Minkowski and Hilbert had been students at the same time in the high schools of Königsberg. "Oh, that Einstein, always skipping lectures. ... I certainly never would have thought he could do it," said Minkowski when he read the great Einstein papers of 1905.

There is an elegiac, even mournful, quality to the tale, lively though it is. Mathematics died in Göttingen by the Nazi hand. Hilbert knew it. "How is mathematics in Göttingen now that it has been freed of the Jewish influence?" asked the Nazi minister of the elderly Prussian professor who sat next to him at a banquet. "There is really none any more," replied Hilbert. All his life he had used one estimate of men: their contribution to knowledge. "No prejudice—national, sexual or racial—had ever been allowed to enter into it." He had written a memorial for the French mathematician Jean Gaston Darboux while France and Germany were at war. He had found a post for Emmy Noether when no woman had ever been *dozent* at Göttingen. "After all, [it] is not a bathhouse," he said. Among his closest friends was the man he had brought as professor to Göttingen, the Jew Minkowski.

Irony entered even mathematics. A few months after Hilbert had retired the young logician Kurt Gödel (who had never met or written to Hilbert) published his famous papers that were a catastrophe for Hilbert's lifelong scheme for the foundations of mathematics. "[Hilbert's] objective was to prove the consistency of the axioms of classical mathematics on the basis of evidence just as concrete and immediately convincing as elementary arithmetic." So writes Gödel himself, who saw Hilbert's program not so much as a failure as a first step leading to the new science of metamathematics.

This is a remarkable kind of biography. The author is not a mathematician, but she has brought Hilbert and his times alive, leaving the indispensable mathematical background to lengthy citations from many mathematicians, including Hilbert himself. The result is not easy reading, but it is definitive. Those who can understand will find the subject rather fully covered, but those who cannot read the mathematicians will find no easy road to the kernel of Hilbert's thought.

David Hilbert's mind was brave, honest and magisterial. His epitaph, in his own phrases, is the affirmation of his own life, now shadowed by the darkness in which men and their thoughts have been plunged since fire flickered in the Reichstag ruins and ravaged the streets of Hiroshima. "Wir müssen wissen. Wir werden wissen." We must know; we shall know. That "contagious optimism" is the spirit of Hilbert.

The photographs of people and places are fascinating. There is one of Hilbert in a floppy Panama, checked ascot and tousled jacket, which belongs to a "group of portraits of professors which were sold as postcards in Göttingen" in 1912. Last year, it is true, one could find for sale in Harvard Square the poster-sized photograph of an M.I.T. psychologist. But it was for television debate and not for the lecture hall that our contemporary was so admired.

ASPECTS OF PALYNOLOGY, edited by Robert H. Tschudy and Richard A. Scott. Wiley-Interscience ($24.95). HOW TO KNOW POLLEN AND SPORES, by Ronald O. Kapp. Wm. C. Brown Company Publishers ($3.50). The Swedish paleobotanist Lennart von Post realized during World War I that the peat samples from ancient bogs hold a rich assembly of pollen grains, whose complex sculptured surfaces roughly identify the parent plant, and whose census reveals the succession of plant life as the glaciers approached and receded. He founded palynology, the science of pollen and spores. The name was coined in 1944 from the Greek for "fine dust." The first of these two books surveys the entire field in 18 chapters by a gaggle of experts (whose affiliations are somehow omitted). Its aim is to cover the field of fossil pollen, with only occasional references to the pollen produced by contemporary plants.

There are several chapters on the overall problems: nomenclature, sampling, chemical composition. These are followed by eight chapters on the spores and pollens of the various geological epochs. There is one chapter on the past 40,000 years, a period when the samples are so rich that they repay elaborate statistical analysis. There are other texts on this specialized subject. There is one chapter on the other microfossils in this size range, many of unknown nature, which may confuse the problem. (There was in *Nature* not long ago a worldwide appeal to identify a tiny sculpture, which was recognized by several experts as a microextrusion manufactured by a leafhopper!)

How to Know Pollen and Spores, one of a familiar series of illustrated nature keys, deals with the grains routinely found in North American samples from

living plants. This netsuke-like "fine dust" consists of the carved, knobbed, spiny or mesh-covered external shell of the once living pollen or spore, wind-blown or insect-borne into the sample. The objects are often strikingly beautiful and even more frequently strangely intricate. They are by now tediously well classified. The grains are made of a waxy nonbiodegradable natural plastic, so that they can be concentrated by dissolving away nearly everything else: the other organic tissue with acetic anhydride and sulfuric acid, the siliceous matter in powdered rock or clay with concentrated hydrofluoric acid! The pollen walls come through all of this, modified to be sure but still bearing details on the one-micron scale that the expert eye can perceive. That eye needs of course a good light microscope, or nowadays a transmission or scanning electron microscope. There are plenty of photographs and many more drawings in both of these books.

In the 1930's the micropaleontologists entered applied science. The oil companies realized that an exquisite correlation of the strata through which they drilled could be carried out by studying the little limy microfossils of the foraminifera. Forams are to be found almost everywhere in drill cores, if you do not go back more than a couple of hundred million years and search only marine sediments.

The subscience has burgeoned. Those tiny seashells are measured in hundreds of microns. In the 1950's the same oil seekers found that the pollen and spores (which are an order of magnitude smaller than forams, typically some tens of microns in diameter) work to the same end. These smaller particles, however, embrace all of geologic time, although they do favor freshwater or dry-land environments.

Today this new science is the largest part of paleobotany. These two books, jagged with long Greek taxonomic names and subtle distinctions of form and kinship, are perhaps a little too dusty for the general reader. There is one special chapter: James Schopf of the U.S. Geological Survey writes on the first vestiges of life. He has drawn them all here; they belong to palynology, although they are not solely spores. They are found in the chalcedony of the cherts of three billion years ago as bacillar forms, not much different from modern ones. They do not seem to be contaminants. Some of the remarkable forms found in a few meteorites have been shown to be ragweed pollen in its less familiar germinated

state. "Not all, however, have been explained." Working at the 10-micron scale, experimenters examine a world where gravity is insignificant compared with surface forces. Any speck can drift onto a surface and stick to it. Writes Schopf: "There can never be an *absolute* safeguard against foreign small-particle contamination."

STUDIES IN EROTIC ART, edited by Theodore Bowie and Cornelia V. Christenson. Basic Books, Inc., Publishers ($15). A product of the Kinsey Institute for Sex Research, this volume is a handsome outgrowth of the collection of erotic materials assembled with the royalties from *Sexual Behavior in the Human Male*. The first three papers in the volume, each accompanied by a few dozen reproductions in black and white, present the most renowned examples of the natural history of human sexuality in ancient art. The first paper is a compendium from classical times. Painted pottery, carved lamps, large stone reliefs and a number of other mediums are included, mainly displaying themes such as Leda and the marvelous swan and assorted sexual encounters around the rim of a cup, usually vigorous and realistic figures against a ground of solid color, sometimes in the elegant surround of a Pompeiian villa. The second paper shows photographs of intricately modeled Peruvian pottery of pre-Columbian times, spanning two millenniums and several cultures. The Inca rule ended the style. These ceramics generally portray human genitalia in grotesquely prominent form; there are also scenes of active sexual contact, some of it between flayed cadavers.

The third paper is a collection of famous Japanese erotica, mostly woodcuts, with some other drawings. These are in the full style of the "floating world," mainly vivid and realistic in tone but with a good deal of deliberate exaggeration and some reference to the cruder strata of society. Each of these three sets of pictures, with the accounts by art historians and archaeologists, represents a partial equivalent of a museum visit to one of these famous widespread national collections. Almost every image presents that curious tension between the trite and the startling which is the high art of pornography.

There are two other papers that belong more fully to art history; here the works discussed are symbolic rather than explicit. The first explains why there is a leg missing from the figure of Christ in the magnificent "Pietà" by Michelangelo

at the Duomo in Florence. The explanation by Professor Leo Steinberg is compelling, a *tour de force* of iconographic studies. The last piece interprets rather simply the more pneumatic periods of Picasso's paintings and drawings. Altogether this is a book suited to the times, yet not too strange for any library that feels a deep self-assurance.

HARD TIMES: AN ORAL HISTORY OF THE GREAT DEPRESSION, by Studs Terkel. Pantheon Books ($8.95). Here is "a memory book," the verbatim account of some 160 taped interviews drawn out by the gifted 58-year-old journalist from Chicago. It holds more of myth than his first set of interviews, *Division Street*, which exhibited the inner life of today's Chicago. These men and women are looking back 30 years. The stories occupy a page or two each; they flow from Sally Rand and James Farley, from Saul Alinsky and Myrna Loy, from Russell Long and William Patterson, the man who found counsel for the Scottsboro boys. There are miners and heist men, professors and playboys, farmers and a court bailiff. "Everybody passed through the neighborhood was a member of the Workers Alliance.... When that one person came, he'd have about fifty people with him.... The men would connect those lights and... get gas pipe, and connect that stove back... so it don't look like you been out the door."

Emma Tiller, dispossessed black sharecropper's wife, "trained my own self to cook.... As long as you got your hand in the lion's mouth, you have to be easy till you get it out. Well, I got my hand out." She had guilefully abandoned her Wichita Falls lady, "the whiskiest folks you ever saw," when she was asked to cook and serve a banquet for 40 guests by herself. "All that work for that same $7. Not on your life." "She was to pay you $7 for that day?" asked the interviewer. "Seven bucks a week, honey." In 1932 the police and the Marines had declined to expel the Bonus Marchers from Anacostia; the 12th Infantry did it, in full battle dress, with bayonets, gas and four baby tanks, behind General MacArthur "riding a white horse.... He had a youngish major as an aide. His name was Dwight Eisenhower." E. Y. Harburg, later a songwriter but then in business, says: "Collapse, bango!... All I had left was a pencil." From it flowed "Brother, Can You Spare a Dime?" According to Harburg, "the Republicans... tried to ban it from the air. But it was too late. The song had already done its damage."

INDEX OF ADVERTISERS

JULY 1970

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

CONVERSION TO THE METRIC SYSTEM

THE METRIC SYSTEM IN THE UNITED KINGDOM: THE USE OF SI UNITS. British Standards Institution, 1965.

CHANGING TO THE METRIC SYSTEM: CONVERSION FACTORS, SYMBOLS AND DEFINITIONS. Pamela Anderton and P. H. Bigg. Her Majesty's Stationery Office, 1969.

GUIDE TO METRICATION. M. J. B. Jones. Pergamon Press, 1969.

GOING METRIC: THE FIRST 5 YEARS, 1965–69. Metrication Board 1970. Her Majesty's Stationery Office, 1970.

GLOBULAR-CLUSTER STARS

EVOLUTION OF STARS AND GALAXIES. Walter Baade. Harvard University Press, 1963.

STELLAR EVOLUTION: COMPARISON OF THEORY WITH OBSERVATION. Icko Iben, Jr., in *Science*, Vol. 155, No. 3764, pages 785–796; February 17, 1967.

ENERGY PRODUCTION IN STARS. Hans A. Bethe in *Science*, Vol. 161, No. 3841, pages 541–547; August 9, 1968.

STELLAR ABUNDANCES AND THE ORIGIN OF THE ELEMENTS. Albrecht O. J. Unsöld in *Science*, Vol. 162, No. 3871, pages 1015–1025; March 7, 1969.

THE MULTIPLE SCLEROSIS PROBLEM

THE FREQUENCY AND GEOGRAPHIC DISTRIBUTION OF MULTIPLE SCLEROSIS AS INDICATED BY MORTALITY STATISTICS AND MORBIDITY SURVEYS IN THE UNITED STATES AND CANADA. Leonard T. Kurland in *The American Journal of Hygiene*, Vol. 55, No. 3, pages 457–476; May, 1952.

MULTIPLE SCLEROSIS IN ISRAEL. Milton Alter, Lipman Halpern, Leonard T. Kurland, Bernard Bornstein, Uri Leibowitz and Joseph Silberstein in *Archives of Neurology*, Vol. 7, No. 4, pages 253–263; October, 1962.

THE PREVALENCE OF MULTIPLE SCLEROSIS IN AUSTRALIA. John M. Sutherland, John H. Tyrer and Mervyn J. Eadie in *Brain: A Journal of Neurology*, Vol. 85, Part I, pages 149–164; 1962.

MULTIPLE SCLEROSIS: A REAPPRAISAL. Douglas McAlpine, Charles E. Lumsden and E. D. Acheson. E. & S. Livingstone Ltd, 1968.

NERVE CELLS AND BEHAVIOR

NERVE, MUSCLE, AND SYNAPSE. Bernhard Katz. McGraw-Hill Book Company, 1966.

RESPONSE DECREMENT OF THE FLEXION REFLEX IN THE ACUTE CAT AND TRANSIENT RESTORATION BY STRONG STIMULI. W. A. Spencer, R. F. Thompson and D. R. Neilson, Jr., in *Journal of Neurophysiology*, Vol. 29, No. 2, pages 221–239; March, 1966.

CELLULAR NEUROPHYSIOLOGICAL APPROACHES IN THE STUDY OF LEARNING. Eric R. Kandel and W. Alden Spencer in *Physiological Reviews*, Vol. 48, No. 1, pages 65–134; January, 1968.

ANALYSIS OF RESTRICTED NEURAL NETWORKS. Donald Kennedy, Allen I. Selverston and Michael P. Remler in *Science*, Vol. 164, No. 3887, pages 1488–1496; June 27, 1969.

THE ROLE OF SYNAPTIC PLASTICITY IN THE SHORT-TERM MODIFICATION OF BEHAVIOR. E. R. Kandel, V. Castellucci, H. Pinsker and I. Kupfermann in *Short-Term Processes in Neural Activity and Behavior*, edited by R. A. Hinde and G. Horn. Cambridge University Press, in press.

NEGATIVE VISCOSITY

THE FLUX OF ANGULAR MOMENTUM IN ROTATING MODEL EXPERIMENTS. Victor P. Starr and Robert R. Long in *Geophysical Research Papers, No. 24: Contributions to the Study of Planetary Atmospheric Circulations*, edited by Robert H. White. Air Force Cambridge Research Center, 1953.

THE GENERAL CIRCULATION OF THE SOLAR ATMOSPHERE AND THE MAINTENANCE OF THE EQUATORIAL ACCELERATION. Fred Ward in *The Astrophysical Journal*, Vol. 141, No. 2, pages 534–547; February 15, 1965.

MEASUREMENTS OF EDDY FLUXES OF MOMENTUM IN THE SURFACE LAYER OF THE GULF STREAM. Ferris Webster in *Tellus*, Vol. 17, No. 2, pages 239–245; May, 1965.

PHYSICS OF NEGATIVE VISCOSITY PHENOMENA. Victor P. Starr. McGraw-Hill Book Company, 1968.

THE LOVE SONG OF THE FRUIT FLY

STIMULI PROVIDED BY COURTSHIP OF MALE *DROSOPHILA MELANOGASTER*. H. C. Bennet-Clark and A. W. Ewing in *Nature*, Vol. 215, No. 5101, pages 669–671; August 5, 1967.

THE WING MECHANISM INVOLVED IN THE COURTSHIP OF *DROSOPHILA*. H. C. Bennet-Clark and A. W. Ewing in *The Journal of Experimental Biology*, Vol. 49, No. 1, pages 117–128; August, 1968.

THE COURTSHIP SONGS OF *DROSOPHILA*. Arthur W. Ewing and H. C. Bennet-Clark in *Behaviour*, Vol. 31, Parts 2 and 3, pages 288–301; 1968.

NETWORK ANALYSIS

FLOWS IN NETWORKS. L. R. Ford, Jr., and D. R. Fulkerson. Princeton University Press, 1962.

THE THEORY OF GRAPHS AND ITS APPLICATIONS. Claude Berge. John Wiley & Sons Inc., 1962.

GRAPHS AND THEIR USES. Oystein Ore. Random House, 1963.

COMMUNICATIONS, TRANSMISSION AND TRANSPORTATION NETWORKS. H. Frank and I. T. Frisch. Addison-Wesley Publishing Co., Inc., to be published.

POPULATION TRENDS IN AN INDIAN VILLAGE

THE HEALTH SCIENCES AND INDIAN VILLAGE CULTURE. Carl E. Taylor in *Science and the Human Condition in India and Pakistan*, edited by Ward Morehouse. The Rockefeller University Press, 1968.

FIVE STAGES IN A PRACTICAL POPULATION POLICY. Carl E. Taylor in *International Development Review*, Vol. 10, No. 4, pages 2–7; December, 1968.

MATHEMATICAL GAMES

DIOPHANTUS OF ALEXANDRIA: A STUDY IN THE HISTORY OF GREEK ALGEBRA. Sir Thomas L. Heath. Dover Publications, Inc., 1964.

A SURVEY OF EQUAL SUMS OF LIKE POWERS. L. J. Lander, T. R. Parkin and J. L. Selfridge in *Mathematics of Computation*, Vol. 21, No. 99, pages 446–459; July, 1967.

DIOPHANTINE EQUATIONS. Louis Joel Mordell. Academic Press, Inc., 1969.

THE AMATEUR SCIENTIST

UNDERGRADUATE MASS SPECTROMETER. John W. Dewdney in *American Journal of Physics*, Vol. 28, No. 5, pages 452–456; May, 1960.

In this picture of Earth taken by NASA astronauts, we become dramatically aware that all mankind is on this trip through space together.

At this distance, subtle geopolitical boundaries are indistinguishable. Our planet is a mutually shared home in which the problems of one area become the problems of all.

From our new vantage point, we can also see things to help solve some of these problems. With this nation's existing space technologies we could monitor world crops for disease or locate new fisheries resources. We could open new frontiers in meteorology, education, navigation, forestry, communications. With benefits in nearly every area of human activity.

Think about the possibilities.

The Boeing Company.